

Chapter 1

Bird's Eye View

Well begun is half done.

Old Proverb

1.0 Purpose of the Chapter

Before starting on any journey, thoughtful people study a map of where they are going. This allows them to keep their bearings as they progress, and not get lost en route. This chapter is like such a map, a schematic overview of the terrain of quantum field theory (QFT) without the complication of details. You, the student, can get a feel for the theory, and be somewhat at home with it, even before delving into the “nitty-gritty” mathematics. Hopefully, this will allow you to keep sight of the “big picture”, and minimize confusion, as you make your way, step-by-step, through this book.

1.1 This Book’s Approach to QFT

There are two main branches to (ways to do) quantum field theory called

- the canonical quantization approach, and
- the path integral approach (also called the many paths or sum over histories approach).

This book is primarily devoted to the first of these. It is considered by many, and certainly by me, as the easiest way to be introduced to the subject. I have, however, provided a simplified, brief introduction to the path integral approach at the end of the book. Students wishing to make a career in field theory will eventually need to become well versed in both approaches.

1.2 Why Quantum Field Theory?

The quantum mechanics (QM) courses students take prior to QFT generally treat a single particle such as an electron in a potential (e.g., square well, harmonic oscillator, etc.), and the particle retains its integrity (e.g., an electron remains an electron throughout the interaction.) There is no general way to treat interactions between particles, such as that of a particle and its antiparticle annihilating one another to yield neutral particles such as photons (e.g., $e^- + e^+ \rightarrow 2\gamma$.) Nor is there any way to describe the decay of an elementary particle such as a muon into other particles (e.g. $\mu^- \rightarrow e^- + v + \bar{\nu}$, where the latter two symbols represent neutrino and antineutrino, respectively).

Here is where QFT comes to the rescue. It provides a means whereby particles can be annihilated, created, and transmigrated from one type to another. In so doing, its utility surpasses that provided by ordinary QM.

There are other reasons why QFT supersedes ordinary QM. For one, it is a relativistic theory, and thus more all encompassing. Further, as we will discuss more fully later on, the straightforward extrapolation of non-relativistic quantum mechanics (NRQM) to relativistic quantum mechanics (RQM) results in states with negative energies, and in the early days of quantum theory, these were quite problematic. We will see in subsequent chapters how QFT resolved this issue quite nicely.

Limitation of original QM: no transmutation of particles

QFT: transmutation included

*Energies < 0
RQM yes
QFT no*

*QFT example:
Bhabba scattering*

1.3 How Quantum Field Theory?

As an example of the type of problem QFT handles well, consider the interaction between an electron and a positron known as Bhabha scattering shown in Fig. 1. At event x_2 , the electron and

positron annihilate one another to produce a photon. At event x_1 , this photon is transmuted back into an electron and a positron. Antiparticles like positrons are represented by lines with arrows pointing opposite their direction of travel through time. The seemingly strange, reverse order of numbering here, i.e., $2 \rightarrow 1$, is standard in QFT.

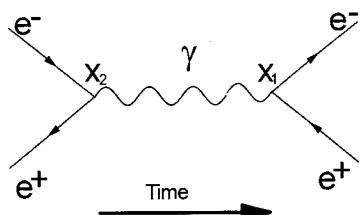


Fig 1. Bhabha Scattering

Note that we can think of this interaction as an annihilation (destruction) of the electron and the positron at x_2 accompanied with creation of a photon, and that followed by the destruction of the photon accompanied by creation of an electron and positron at x_1 . Unlike the electrons and positrons in this example, the photon here is not a “real” particle, but transitory, short-lived, and undetectable, and is called a “virtual” particle (which mediates the interaction between real particles.)

What we seek and what, as students eventually see, QFT delivers, is a mathematical relationship, called a transition amplitude, describing a transition from an initial set of particles to a final set (i.e., an interaction) of the sort shown pictorially via the Feynman diagram of Fig. 1. It turns out that the square of the absolute value of the transition amplitude equals the probability of finding (upon measurement) that the interaction occurred. This is similar to the square of the absolute value of the wave function in QM equaling the probability density of finding the particle.

QFT employs creation and destruction operators acting on states (i.e., kets), and these creation/destruction operators are part of the transition amplitude. We illustrate the general idea with the following grossly oversimplified transition amplitude, reflecting the interaction process of Fig. 1. Be cautioned that we have omitted a few more formal, and ultimately essential, ingredients in (1-1), in order to make it simpler, and easier, to grasp the fundamental concept.

$$\text{Transition amplitude} = \langle e^+ e^- | (\bar{\psi}_c \mathcal{A}_d \psi_c)_{x_1} (\bar{\psi}_d \mathcal{A}_c \psi_d)_{x_2} | e^+ e^- \rangle_{x_2}. \quad (1-1)$$

In (1-1), the ket $|e^+ e^- \rangle_{x_2}$ represents the incoming electron and positron at x_2 . The bra represents the outgoing electron and positron at x_1 . ψ_d is an operator that destroys an electron (at x_2); $\bar{\psi}_d$ an operator that destroys a positron (at x_2); ψ_c creates a positron (at x_1); and $\bar{\psi}_c$ creates an electron (at x_1). The \mathcal{A}_c is a photon operator that creates (at x_2) a virtual photon, and \mathcal{A}_d is an operator that destroys (at x_1) that virtual photon, with the lines underneath indicating that the photon is virtual and propagates from x_2 to x_1 . The mathematical procedure and symbolism (lines underneath) representing this virtual particle (photon here) process, as shown in (1-1), is called a contraction. When the virtual particle is represented as a mathematical function, it is known as the Feynman propagator or simply, the propagator, because it represents the propagation of a virtual particle from one event to another.

Feynman propagator

Note what happens to the ket part of the transition amplitude as we proceed, step-by-step, through the interaction process. At x_2 , the incoming particles (in the ket) are destroyed by the destruction operators, so at an intermediate point, we have

$$\text{transition amplitude} = \langle e^+ e^- | (\bar{\psi}_c \mathcal{A}_d \psi_c)_{x_1} (\mathcal{A}_c)_{x_2} K_2 | 0 \rangle, \quad (1-2)$$

Destruction operators leave vacuum ket times a numeric factor

where the destruction operators have acted on the original ket to leave the vacuum ket $|0\rangle$ (no particles left) with a purely numeric factor K_2 in front of it. The value of this factor is determined by the formal mathematics of QFT.

In the next step after (1-2), the virtual photon propagator, due to the creation operator \mathcal{A}_c , creates a virtual photon at x_2 that then propagates from x_2 to x_1 , where it is annihilated by \mathcal{A}_d . This process leaves the vacuum ket still on the right along with an additional numeric factor, which comes out of the formal mathematics, and which we designate below as K_γ .

Section 1.4 From Whence Creation and Destruction Operators?

$$\text{transition amplitude} = \langle e^+ e^- | (\bar{\psi}_c \psi_c)_{x_1} K_1 K_2 | 0 \rangle \quad (1-3)$$

Propagator action leaves only another numeric factor

Creation operators leave final state ket plus one more numeric factor

Calculated amplitude

Bracket of multiparticle state = 1 in QFT

Probability = |Amplitude|^2

QFT wave equation solutions are operators

The remaining creation operators then create an electron and positron out of the vacuum at x_1 . This leaves us with the newly created ket $|e^+ e^- \rangle_{x_1}$ plus a factor K_1 in front. The ket and the bra now represent the same state, i.e., the same particles at the same time and place x_1 , so their inner product (the bracket) is not zero (as it would be if they were different states). Nor are there any operators left, but only numeric quantities, so we can move them outside the bracket without changing anything. Thus, at x_1 and thereafter, we have

$$\begin{aligned} \text{transition amplitude} &= \langle e^+ e^- | K_1 K_\gamma K_2 | e^+ e^- \rangle_{x_1} = \langle e^+ e^- | \underbrace{S_{Bhabba}}_{\substack{\text{just a number} \\ \text{without operators}}} | e^+ e^- \rangle_{x_1} \\ &= S_{Bhabba} \underbrace{\langle e^+ e^- | e^+ e^- \rangle_{x_1}}_1 = S_{Bhabba} \end{aligned} \quad , \quad (1-4)$$

where we note the important point that *in QFT the bracket of a multiparticle state* (inner product of multiparticle state with itself) such as that shown in (1-4) is defined so it *always equals unity*.

The whole process can be pictured as simply an evolution, or progression, of the original state, represented by the ket, to the final state, represented by the bra. At each step along the way, the operators act on the ket to change it into the next part of the progression. When we get to the point where the ket is the same as the bra, the full transition has been made, and the bracket then equals unity. What is left is our transition amplitude.

Finally, the probability of the interaction occurring turns out to be

$$\text{probability of interaction} = S_{Bhabba}^\dagger S_{Bhabba} = |S_{Bhabba}|^2. \quad (1-5)$$

The quantity $K_1 K_\gamma K_2 = S_{Bhabba}$ arising in (1-4) depends on particle momenta, spins, and rest masses, as well as the inherent strength of the electromagnetic interaction, all of which one would rightly expect to play a role in the probability of an interaction taking place.

From the interaction probability, scattering cross sections can be calculated.

1.4 From Whence Creation and Destruction Operators?

In ordinary QM, the solutions to the relevant wave equation, the Schroedinger equation, are states (particles or kets.) Surprisingly, the *solutions to the relevant wave equations in QFT are not states (not particles.)* In QFT, it turns out that these solutions are *actually operators that create and destroy states.* Different solutions exist that create or destroy every type of particle and antiparticle. In this unexpected (and, for students, often strange at first) twist lies the power of QFT.

1.5 Overview: The Structure of Physics and QFT's Place Therein

Students are often confused over the difference (and whether or not there is a difference) between relativistic quantum mechanics (RQM) and QFT. The following discussion, summarized below in Wholeness Chart 1-1, should help to distinguish them.

1.5.1 Background: Poisson Brackets and Quantization

Classical particle theories contain rarely used entities call Poisson brackets, which, though it would be nice, are not necessary for you to completely understand at this point. (We will show their precise mathematical form in Chap. 2.) What you should realize now is that Poisson brackets are mathematical manipulations of certain pairs of properties (dynamic variables like position and momentum) that bear a striking resemblance to commutators in quantum theories. For example, the Poisson bracket for position X (capital letters will designate Cartesian coordinates in this book) and momentum p_X , symbolically expressed herein as $\{X, p_X\}$, is non-zero (and equal to one), but the Poisson bracket for Y and p_X equals zero.

Shortly after NRQM theory had been worked out, theorists realized that for each pair of quantum operators that had non-zero (zero) commutators, the corresponding pair of classical dynamic

Poisson brackets parallel quantum commutators

variables also had non-zero (zero) Poisson brackets. They had originally arrived at NRQM by taking classical dynamic variables as operators, and that led, in turn, to the non-zero commutation relations for certain operators (which result in other quantum phenomena such as uncertainty.) But it was soon recognized that one could do the reverse. One could, instead, take the classical Poisson brackets over into quantum commutation relations first, and because of that, the dynamic variables turn into operators. (Take my word for this now, but after reading the next section, do Prob. 6 at the end of this chapter.)

The process of extrapolating from classical theory to quantum theory became known as quantization. Apparently, for many, the specific process of starting with Poisson brackets and converting them to commutators was considered the more elegant way to quantize.

*Quantization:
Poisson brackets
become
commutators*

1.5.2 First vs. Second Quantization

Classical mechanics has both a non-relativistic and a relativistic side, and each contains a theory of particles (localized entities, typically pointlike objects) and a theory of fields (entities extended over space and time). All of these are represented in the first two rows of Wholeness Chart 1-1. Properties (dynamic variables) of entities in classical particle theories are *total* values, such as object mass, charge, energy, momentum, etc. Properties in classical field theories are *density* values, such as mass density, charge density, field amplitude at a point, etc. Poisson brackets in field theories are just like those for particle theories, except they entail *densities* of the respective dynamic variables, instead of total values.

With the success of quantization in NRQM, people soon thought of applying it to relativistic theory and found they could deduce RQM in the same way. Shortly thereafter they tried applying it to fields, the result being QFT. The term first quantization came to be associated with *particle* theories. The term second quantization became associated with *field* theories.

In quantizing, we also assume the classical Hamiltonian (total or density value) has the same quantum form. We can summarize all of this as follows.

*Branches of
classical
mechanics*

*1st quantization is
for particles; 2nd is
for fields*

First Quantization (Particle Theories)

- 1) Assume the quantum particle Hamiltonian has the same form as the classical particle Hamiltonian.
- 2) Replace the classical Poisson brackets for conjugate properties with commutator brackets (divided by $i\hbar$), e.g.,

$$\{X_i, p_j\} = \delta_{ij} \Rightarrow [X_i, p_j] = i\hbar\delta_{ij}. \quad (1-6)$$

In doing (1-6), the classical properties (dynamic variables) of position and its conjugate 3-momentum become quantum non-commuting operators.

Second Quantization (Field Theories)

- 1) Assume the quantum field Hamiltonian density has the same form as the classical field Hamiltonian density.
- 2) Replace the classical Poisson brackets for conjugate property densities with commutator brackets (divided by $i\hbar$), e.g.

$$\{\phi_r(\mathbf{x}, t), \pi_s(\mathbf{y}, t)\} = \delta_{rs}\delta(\mathbf{x} - \mathbf{y}) \Rightarrow [\phi_r(\mathbf{x}, t), \pi_s(\mathbf{y}, t)] = i\hbar\delta_{rs}\delta(\mathbf{x} - \mathbf{y}), \quad (1-7)$$

where π_s is the conjugate momentum density of the field ϕ_s , different values for r and s mean different fields, and \mathbf{x} and \mathbf{y} represent different 3D position vectors. In doing (1-7), the classical field dynamic variables become quantum field non-commuting operators (and this, as we will see, has major ramifications for QFT.)

Note that the specific quantization we are talking about here (both first and second) is called canonical quantization, because, in both the Poisson brackets and the commutators, we are using (in classical mechanics terminology) canonical variables. For example, p_x is called the *canonical*

*Our approach here:
canonical
quantization*

Section 1.6 Comparison of Three Quantum Theories

momentum of X . (It is sometimes also called the *conjugate momentum*, as we did above, or the *generalized momentum* of X .)

This differs from the form of quantization used in the path integral approach (see Sect. 1.1 on page 1) to QFT, which is known as functional quantization, because the path integral approach employs mathematical quantities known as *functionals*. (See Chap. XXX at the end of the book for a brief introduction to this alternative method of doing QFT.)

*Path integral
approach to QFT:
functional
quantization*

1.5.3 The Whole Physics Enchilada

All of the above two sections is summarized in Wholeness Chart 1-1. In using it, the reader should be aware that, depending on context, the term quantum mechanics (QM) can mean i) only non-relativistic (“ordinary”) quantum mechanics (NRQM), or ii) the entire realm of quantum theories including NRQM, RQM, and QFT. In the chart, we employ the second of these.

Note that because quantum field applications almost invariably involve photons or other relativistic particles, there is little real world application for a non-relativistic quantum field theory. Further, a relativistic theory will always suffice for any non-relativistic application. Hence, the QFT we restrict our study to is specifically relativistic.

As an aside, quantum theories of gravity such as superstring theory are not included in the chart, as QFT in its standard model form cannot accommodate gravity. Thus, the relativity in QFT is special, but not general, relativity.

Wholeness Chart 1-1. The Overall Structure of Physics

	<u>Non-relativistic</u>		<u>Relativistic</u>	
	Particle	Field	Particle	Field
Classical (non-quantum) mechanics	Newtonian particle theory	Newtonian field theory (continuum mechanics + gravity), e/m (quasi-static)	Relativistic macro particle theory	Relativistic macro field theory (continuum mechanics + e/m + gravity)
Properties (Dynamic variables) \Downarrow Operators	\Downarrow 1 st quantization \Downarrow		\Downarrow 1 st quantization \Downarrow	\Downarrow 2 nd quantization \Downarrow
Quantum mechanics	NRQM	No theory generally taught.	RQM	QFT (not gravity)

Conclusions: RQM is similar to NRQM in that both are particle theories. They differ in that RQM is relativistic. RQM and QFT are similar in that both are relativistic theories. They differ in that QFT is a field theory and RQM is a particle theory.

1.6 Comparison of Three Quantum Theories

NRQM employs the (non-relativistic) Schroedinger equation, whereas RQM and QFT must employ relativistic counterparts sometimes called relativistic Schroedinger equations. Students of QFT soon learn that each spin type (spin 0, spin $\frac{1}{2}$, and spin 1) has a different relativistic Schroedinger equation. For a given spin type, that equation is the same in RQM and in QFT, and hence, both theories have the same form for the solutions to those equations.

*Different spin
types → different
wave equations*

The difference between RQM and QFT is in the meaning of those solutions. In RQM, the solutions are interpreted as states (particles, such as an electron), just as in NRQM. In QFT, though it may be initially disorienting to students previously acclimated to NRQM, the solutions turn out not to be states, but rather operators that create and destroy states. Thus, QFT can handle transmutation of particles from one kind into another (e.g., muons into electrons, by destroying the

*Solutions:
RQM → states
QFT → operators*

original muon and creating the final electron), whereas NRQM and RQM can not. Additionally, the problem of negative energy state solutions in RQM does not appear in QFT, because, as we will see, the creation and destruction operator solutions in QFT create and destroy both particles and anti-particles. Both of these have positive energies.

QFT can be done without negative energies

*QFT:
multiparticle*

Additionally, while RQM (and NRQM) are amenable primarily to single particle states (with some exceptions), QFT more easily, and more compressively, accommodates multi-particle states.

In spite of the above, there are some contexts in which RQM and QFT may be considered more or less the same theory, in the sense that QFT encompasses RQM. By way of analogy, classical relativistic particle theory is inherent within classical relativistic field theory. For example, one could consider an extended continuum of matter which is very small spatially, integrate the mass density to get total mass, the force/unit volume to get total force, etc., resulting in an analysis of *particle dynamics*. The field theory contains within it, the particle theory.

*RQM
contained in
QFT*

In a somewhat similar way, QFT deals with relativistic states (kets), which are essentially the same states dealt with in RQM. QFT, however, is a more extensive theory and can be considered to encompass RQM within its structure.

*Calculate
expectation values
same way*

And in both RQM and QFT (as well as NRQM), operators act on states in similar fashion. For example, the expected energy measurement is determined the same way in both theories, i.e.,

$$\bar{E} = \langle \phi | H | \phi \rangle, \quad (1-8)$$

with similar relations for other observables.

These similarities and differences, as well as others, are summarized in Wholeness Chart 1-2. Comparison of Three Theories. The chart is fairly self explanatory, though we augment it with a few comments. You may wish to follow along with the chart as you read them (below).

The different relativistic Schroedinger equations for each spin type are named after their founders (see names in chart.) We will cover each in depth. At this point, you have to simply accept that in QFT their solutions are operators that create and destroy states (particles). We will soon see how this results from the commutation relation assumption of 2nd quantization (1-7).

With regard to phenomena, I recall wondering, as a student, why some of the fundamental things I studied in NRQM seemed to disappear in QFT. One of these was bound state phenomena, such as the hydrogen atom. None of the introductory QFT texts I looked at even mentioned, let alone treated, it. It turns out that QFT can, indeed, handle bound states, but elementary courses typically don't go there. Neither will we, as time is precious, and other areas of study will turn out to be more fruitful. Those other areas comprise scattering (including inelastic scattering where particles transmute types), decay of elementary particles, and vacuum fluctuations.

I also once wondered why spherical solutions to the wave equations are not studied, as they play a big role in NRQM, in both scattering and bound state calculations. It turns out that scattering calculations in QFT can be obtained to high accuracy with the simpler plane wave solutions. So, for most applications in QFT, they suffice.

Wave packets, as well, seem no where to be found in QFT. Like the other things mentioned, they too can be incorporated into the theory, but simple sinusoids (of complex numbers) serve us well in almost all applications. So, wave packets, too, are ignored in introductory (and most advanced) courses.

The last block in the chart points out the scope of each theory with regard to the four fundamental forces. Nothing there should be too surprising.

There are other things from earlier studies that seem to have been lost, as well, and we will mention these as we cross paths with them.

*Phenomena in
the 3 theories*

*QFT rarely
uses spherical
solutions*

*or wave
packets*

*QFT handles
various type
interactions*

Wholeness Chart 1-2. Comparison of Three Theories

	NRQM	RQM	QFT
Wave equation	Schroedinger	Klein-Gordon (spin 0) Dirac (spin $\frac{1}{2}$) Proça (spin 1) Special case of Proça: Maxwell (spin 1 massless)	Same as RQM at left
Solutions to wave equation	States	States	Operators that create and destroy states
Negative energy states?	No	Yes	No
Particles per state	Single*	Single*	Multi-particle
Expectation values	$\bar{\mathcal{O}} = \langle \phi \mathcal{O} \phi \rangle$	As at left, but relativistic.	As at left in RQM.
Phenomena:			
1. bound states	Yes, non-relativistic	Yes, relativistic	Yes, but usually not studied in introductory courses
2. scattering			
a. elastic	a. Yes	a. Yes	a. Yes
b. inelastic (transmutation)	b. No	b. Yes and no. (i.e., cumbersome, but can be done, though only for particle/antiparticle creation & destruction.)	b. Yes
3. decay			
a. composite particles	a. Yes (tunneling)	a. Yes	a. Yes
b. elementary particles	b. No	b. No	b. Yes
4. vacuum fluctuations	No	Yes and no. (Only for simple particle/antiparticle loops.)	Yes
Coordinates			
1. Cartesian (plane waves)	Free, 1D potentials, particles in "boxes"	As at left	Used primarily for free particles, particles in "boxes", and scattering.
2. Spherical (spherical waves)	Bound states and scattering.	As at left.	Not usually used in elementary courses.
Wave Packets	Yes	Yes	Yes, but rarely used. Not taught in intro courses.

Interactions:			
1. e/m	No, though can pseudo model for non relativistic cases	As at left	Yes
2. weak	No	No	Yes
3. strong	No*	No*	Yes
4. gravity	No	No	Not at present.

*Some caveats exist for this chart. For example, NRQM and RQM can handle certain multiparticle states (e.g. hydrogen atom), but QFT generally does it more easily and more extensively. And the strong force can be modeled in NRQM and RQM by assuming a Yukawa potential, though a truly meaningful handling of the interaction can only be achieved via QFT.

1.7 Major Components of QFT

There are four major components of QFT, and this book (after the first two foundational chapters) is divided into four major parts corresponding to them. These are:

1. Free (non-interacting) fields/particles

The field equations (relativistic Schroedinger equations) have no interaction terms in them, i.e., no forces are involved. The solutions to the equations are free field solutions.

2. Interacting fields/particles

In principle, one would simply add the interaction terms to the free field equations and find the solutions. As it turns out, however, doing this is intractable, at best (impossible, at least in closed form, is a more accurate word). The trick employed in interaction theory actually uses the free field solutions of 1 above, so those solutions end up being quite essential throughout all of QFT.

3. Renormalization

If you are reading this text, you have almost certainly already heard of the problem with infinities popping up in the early, naïve QFT calculations. The calculations referred to here are specifically those of the transition amplitude (1-4), where some of the numeric factors, if calculated straightforwardly, turn out to be infinite. Renormalization is the mathematical means by which these infinites are tamed, and made finite.

4. Scattering and decay calculations

The theory of parts 1, 2, and 3 above are put to practical use in determining scattering probabilities and cross sections, as well as decay probabilities (half lives, etc.) Particle decay is governed by the weak force, so we will not do anything with that in the present volume, which is devoted solely to quantum electrodynamics (QED), involving only the electromagnetic force.

1.8 Points to Keep in Mind

When the word “field” is used classically, it refers to an entity, like fluid wave amplitude, \mathbf{E} , or \mathbf{B} , that is spread out in space, i.e., has different values at different places. By that definition, the wave function of ordinary QM, or even the particle state in QFT, is a field. But, it is important to realize that in quantum terminology, the word “field” means an operator field, which is the solution to the wave equations, and which creates and destroys particle states. States (= particles = wave functions = kets) are *not* considered fields in that context.

In this text, the symbol e , representing the magnitude of charge on an electron or positron, is always positive. The charge on an electron is $-e$.

The four major parts of QFT

Terminology
“field” =
operator in
QFT

Symbol $e > 0$

1.9 Big Picture of Our Goal

The big picture of our goal is this. We want to understand Nature. To do so, we need to be able to predict the outcomes of particle accelerator scattering experiments and elementary particle half lives. To do this, we need to be able to calculate probabilities for scattering, and decay, to occur. To do that, we need to be able to calculate transition amplitudes for specific elementary particle interactions. And for that, we need first to master a fair amount of theory, based on the postulates of quantization.

We will work through the above steps in reverse. Thus, our immediate goal is to learn some theory in Parts 1 and 2. Then, how to calculate transition amplitudes, also in Part 2. Scattering and decay will take up Part 3, with necessary refinements in Part 4.

Our goal: predict scattering and decay seen in Nature

Steps to our goal

2nd quantization postulates → QFT theory → transition amplitude calculation → probability
→ scattering and decay results → confirmation of QFT

Steps to our goal

1.10 Summary of the Chapter

In this book, we will close each chapter with a summary, emphasizing its most salient aspects. However, the present chapter is actually a summary (in advance) of the entire book and all of QFT. So, you, the reader, can simply look back in this chapter to find appropriate summaries. These should include Sect. 1.1 (This Book's Approach to QFT), the transition amplitude relations of Eqs. (1-1) through (1-5), Sect. 1.5.2 (1st and 2nd Quantization), Wholeness Chart 1-1 (The Overall Structure of Physics), Wholeness Chart 1-2 (Comparison of Three Theories), and Sect. 9 (Big Picture of Our Goal).

1.11 Suggestions or Questions?

In this and other chapters, if you have questions or suggestions to make the material easier to learn, or if you find any errors, please let me know at rklauer@quantumfieldtheory.info. I will post answers to frequently asked questions at that site, so if you have a question, check there first. I will also post suggestions/corrections that are used as well, and if yours is one, I will list your name as a contributor there and in the next edition.

Robert D. Klauber
Pedagogic Aides to QFT at
www.quantumfieldtheory.info

1.12 Problems

As there is not much in the way of mathematics in this chapter, for most of it, actual problems are not really feasible. However, you may wish to try answering the questions in 1 to 5 below without looking back in the chapter. Doing Prob. 6 can help a lot in understanding first quantization.

1. Draw a Feynman diagram for a muon and anti-muon annihilating one another to produce a virtual photon, which then produces an electron and a positron. Using simplified symbols to represent more complex mathematical quantities (that we haven't studied yet), show how the probability of this interaction would be calculated.
2. Detail the basic aspects of first quantization. Detail the basic aspects of second quantization, then compare and contrast it to first quantization. In second quantization, what is analogous to position in first quantization? What is analogous to particle 3-momentum?
3. Construct a chart showing how non-relativistic theories, relativistic theories, particles, fields, classical theory, and quantum theory are interrelated.
4. For NRQM, RQM, and QFT, construct a chart showing i) which have states as solutions to their wave equations, ii) how to calculate expectation values in each, iii) which can handle

bound states, inelastic scattering, elementary particle decay, and vacuum fluctuations, iv) which can treat the following interactions: i) e/m, ii) weak, iii) strong, and iv) gravity.

5. What are the four major areas of study making up QFT?
6. Using the corresponding Poisson bracket relation $\{X, p_x\} = 1$, we deduce, from first quantization postulate #2, that, quantum mechanically, $[X, p_x] = i\hbar$. For this commutator acting on a function ψ , i.e., $[X, p_x] \psi = i\hbar \psi$, determine what form p_x must have. Is this an operator? Does it look like what you started with in elementary QM, and from which you then derived the commutator relation above? Can we go either way?

Then, take the eigenvalue problem $E\psi = H\psi$, and use the same form of the Hamiltonian H as used in classical mechanics (i.e., $p^2/2m + V$), with the operator form you found for p above. This last step is the other part of first quantization (see page 4).

Did you get the time independent Schrödinger equation? (You should have. ☺) Do you see how, by starting with the Poisson brackets and first quantization, you can derive the basic assumptions of NRQM, i.e., that dynamic variables become operators, the form of those operators, and even the time independent Schrödinger equation, itself? We won't do it here, but from that point, one could deduce the time dependent Schrödinger equation, as well.

Chapter 2

Foundations

*She was a fool, and so am I, and so is anyone
who thinks he sees what God is doing.*

Jonah in *Cat's Cradle*
by Kurt Vonnegut

2.0 Chapter Overview

In this chapter, we will cover the mathematical and physical foundations underlying quantum field theory to be sure you, the reader, are prepared and fit enough to traverse the rest of the book. The first cornerstone of these foundations is a new system of units, called natural units, which is common to QFT, and once learned, simplifies mathematical relations and calculations.

Topics covered after that comprise the notation used in this book, a comparison of classical and quantum waves, variational methods, classical mechanics in a nutshell, different "pictures" in quantum mechanics, and quantum theories in a nutshell. Whereas Chap. 1 was strictly an overview of what you will study, much of this chapter is an overview of what you have already studied, structured to make its role in our work more transparent. The rest is material you will need to know before we leap into the formal development of quantum field theory, beginning in Chap 3.

2.1 Natural Units and Dimensions

The Gaussian system (an extension of cgs devised for use in electromagnetism) has been common in NRQM, although standard international units (SI) [essentially, MKS for electromagnetism] are also used. Natural units are another set of units that arise "naturally" in relativistic elementary particle physics. QFT uses them almost exclusively, they are the units we employ in this book, and we will see how they arise below.

2.1.1 Deducing a System of Units

Convenient systems of units start with arbitrary definitions for units of certain fundamental quantities and derive the remaining units from laws of nature. To see how this works, assume we know three basic laws of nature and we want to devise a system of units from scratch. We will do this first for the cgs system and then for natural units.

The three laws are:

1. The distance L traveled by a photon is the speed of light multiplied by its time of travel. $L = ct$.
2. The energy of a particle of mass is equal to its mass times the speed of light squared. $E = mc^2$.
3. The energy of a photon is proportional to its frequency f . The constant of proportionality is Planck's constant \hbar . $E = \hbar f$ or re-expressed as $E = \hbar\omega$.

2.1.2 Deducing the cgs System

The cgs system takes its fundamental dimensions to be length, mass, and time. It then defines standard units of each of these dimensions to be the centimeter, the gram, and the second, respectively. With these standards and the laws of nature, dimensions and units are then derived for all other quantities science deals with.

*Natural units are
"natural" and
used in QFT*

*Any system of units:
defined units +
laws of nature →
additional derived
units*

For example, from law number one above, the speed of light in the cgs system is known to have dimensions of length/time and units of centimeters/second. Further, by measuring the time it takes for light to travel a certain distance we can get a numerical value of 3×10^{10} cm/sec.

From law number two, the dimensions of energy are mass-length²/time² and the units are gm-cm²/sec². We use shorthand by calling this an erg.

From number three, \hbar has dimensions of energy-time and units of gm-cm²/sec, or for short, erg-sec. It, like the speed of light, can be measured by experiment and is found to be 1.0545×10^{-27} erg-sec.

The point is this. We started with three pre-defined quantities (length, mass, and time) and derived the rest using the laws of nature. Of course, other laws could be used to derive other quantities ($F=ma$ for force, etc.). We only use three laws here for simplicity and brevity.

2.1.3 Deducing Natural Units

With natural units we do much the same thing as was done for the cgs system. We start with three pre-defined quantities and derive the rest. The trick here is that we choose different quantities and define *both* their dimensions and their units in a way that suits our purposes best.

Instead of starting with length, mass, and time, we start with c , \hbar , and energy. We then get even trickier. We take both c and \hbar to have numerical values of one. In other words, just as someone once took an arbitrary distance to call a centimeter and gave it a numerical value of one, or an arbitrary interval of time to call a second and gave it a value of one, we now take whatever amount nature gives us for the speed of light and call it one in our new system. We do the same thing for \hbar . (This, in fact, is why the system is called *natural*, i.e., because we use *nature's* amounts for these things to use as our basic units of measure and not some amount arbitrarily chosen by us.)

We then get even trickier still. We take c and \hbar to be dimensionless, as well. Since c (or any velocity) is distance divided by time, we find, in developing our new system, that length and time must therefore have the same units.

Note that the founders of the cgs system could have done the same type of thing if they had wanted to. If they had started with velocity as dimensionless they would have derived length and time as having the same dimensions, and we might now be speaking of time as measured in centimeters rather than seconds. Alternatively, they could have first decided instead that time and space would be measured in the same units and then derived velocity as a dimensionless quantity. The only difference in these two alternative approaches would have been in choice of which units were considered fundamental and which were derived. In any event this was not done, not because it was invalid, but because it was simply not convenient.

In particle physics, however, it does become convenient, and so we define $c=1$ and dimensionless. It is also convenient to define $\hbar=1$ dimensionless for similar reasons.

With energy, our third fundamental quantity, we stay more conventional. We give it a dimension (energy), and we give it units of mega-electron-volts, i.e., MeV = 1 million eV. (We know from other work "how much" an electron-volt is just as the devisors of the metric system knew "how much" one second is.) As with everything else, we do this because it will turn out to be advantageous.

Note now what happens with our three fundamental entities defined in this way. From law of nature number two with $c=1$ dimensionless, mass has the same units as energy and the same numerical value as well. So an electron with .511 MeV rest energy also has .511 MeV rest mass. Because mass and energy are exactly the same thing in natural units, this dimension has come to be referred to commonly as "mass" (i.e., M) rather than "energy" even though the units remain as MeV.

From law of nature number three with $\hbar=1$ dimensionless, the dimensions for ω are M (instead of sec^{-1} as in cgs), and hence time has dimension M^{-1} and units of (MeV)⁻¹. Similarly, from law number one, length has inverse mass dimensions and inverse MeV units as well. Units and dimensions for all other quantities can be derived from other laws of nature, just as was done in the cgs system.

So, by starting with different fundamental quantities and dimensions, we derive a different (more convenient for particle physics) system of units. Because we started with only one of our three

cgs: cm, gm, sec defined. Other units derived from laws of nature

Natural units: $\hbar = c = 1$ and energy defined. Other units derived from laws of nature.

Energy in natural units: electron volts (MeV convenient)

fundamental entities having a dimension, the entire range of quantities we will deal with will be expressible in terms of that one dimension or various powers thereof.

2.1.4 The Hybrid Units System

When doing theoretical work, natural units are the most streamlined, and thus, usually the quickest and easiest. They are certainly the most common. When carrying out experiments or making calculations that relate to the real world, however, it is often necessary to convert to units which can be measured most readily. In particle physics applications, one typically uses *centimeters, seconds, and MeV*. Note this is a hybrid system and is not quite the same as cgs. (Energy is expressed in ergs in cgs.) It is convenient though, since energy in natural units is MeV, and no conversion is needed for it. Converting other quantities is necessary, however, and there is a little trick for doing it.

*Hybrid units used in experiments:
cm, sec, MeV*

2.1.5 Converting from One System to Another

To do the conversion trick alluded to above, we first have to note two things: i) in natural units any quantity can be multiplied or divided by c or \hbar any number of times without changing either its numerical value or its dimensions, and ii) a quantity is the same thing, the same total amount, regardless of what system it is expressed in terms of.

To illustrate, suppose we determine a theoretical value for some time interval in natural units to be 10^{16} (MeV) $^{-1}$. What is its measurable value in seconds? To find out, observe that

$$t = 10^{16} \text{ (MeV)}^{-1} \times \hbar = 10^{16} \text{ (MeV)}^{-1} \text{ where } \hbar = 1, \text{ and all quantities are in natural units.}$$

But the above relation can be expressed in terms of the hybrid MeV-cm-sec system also. The actual amount of time will stay the same, only the units used to express it, and the numerical value it has in those units, will change. So let's simply change \hbar to its value in the hybrid system, $\hbar = 6.58 \times 10^{-22}$ Mev-sec. Then,

$$t = 10^{16} \text{ (MeV)}^{-1} \times \hbar = 10^{16} \text{ (MeV)}^{-1} \times 6.58 \times 10^{-22} \text{ MeV-sec} = 6.58 \times 10^{-6} \text{ sec.}$$

The same time interval is described as either 10^{16} (MeV) $^{-1}$ or 6.58×10^{-6} seconds depending on our system of units.

The moral here is that we can simply multiply or divide any quantity we like (which is expressed in natural units) by c and/or \hbar (expressed in MeV-cm-sec units) as many times as is necessary to get the units we know that quantity should have in the MeV-cm-sec system.

Multiply natural units by powers of \hbar and/or c to get hybrid units

2.1.6 Mass and Energy in the Hybrid and Natural Systems

As mentioned, the hybrid system is not the same as the cgs system, even though both use centimeters and seconds. In the cgs system, energy is measured in ergs and mass in grams. In the hybrid system, energy is measured in MeV and mass in unfamiliar, and never used, units. (See Wholeness Chart 2-1 below.) It may be confusing, but when experimentalists talk of mass, energy, length, and time, they like to use the hybrid system, *yet* they commonly refer to mass in MeV. For example, in high energy physics, the mass of the electron is commonly referred to as .511 MeV, rather than hybrid (unfamiliar) or cgs (gram) mass units. Hopefully, Wholeness Chart 2-1 will help to keep all of this straight.

*Mass is MeV in natural units.
Commonly
expressed the same
way even when
other units used.*

Though we have used MeV (1 million eV) for energy in hybrid and natural units throughout this chapter, energy is also commonly expressed in keV (kilo electron volts), GeV (giga electron volts = 1 billion eV), and TeV (tera electron volts = 1 trillion eV). It is, of course, simple to convert any of these to, and from, MeV.

Wholeness Chart 2-1. Conversions between Natural, Hybrid, and cgs Numeric Quantities

<u>Natural Units</u>		<u>Hybrid Units</u>		<u>cgs Units</u>	
$c = \hbar = 1$		$c = 2.99 \times 10^{10} \text{ cm/sec}$ $\hbar = 6.58 \times 10^{-22} \text{ MeV-sec}$ $\hbar c = 1.973 \times 10^{-11} \text{ MeV-cm}$		conversion factor $F = 1.602 \times 10^{-6} \text{ ergs/MeV}$	
Quantity, units of $(\text{MeV})^M$	M	Multiply \leftarrow value by \downarrow to get \rightarrow	in MeV-cm-sec	Multiply \leftarrow value by \downarrow to get \rightarrow	in cgs
energy	1	1	MeV	F	ergs
mass, m	1	$1/c^2$	$\text{MeV-sec}^2/\text{cm}^2$	F	$\text{erg-sec}^2/\text{cm}^2 = \text{gms}$
length	-1	$\hbar c$	cm	1	cm
time	-1	\hbar	sec	1	sec
velocity	0	c	cm/sec	1	cm/sec
acceleration, a	1	c/\hbar	cm/sec^2	1	cm/sec^2
force	2	ma factors = $1/c\hbar$	MeV/cm	F	$\text{ergs/cm} = \text{dynes}$
$\hbar (= 1)$	0	\hbar	MeV-sec	F	erg-sec
Hamiltonian	1	1	MeV	F	ergs
Hamiltonian density	4	$1/(\hbar c)^3$	MeV/cm^3	F	ergs/cm^3
Lagrangian	1	1	MeV	F	ergs
Lagrangian density	4	$1/(\hbar c)^3$	MeV/cm^3	F	ergs/cm^3
action S	0	\hbar	MeV-sec	F	erg-sec
fine structure constant	0	1	unitless	1	unitless
cross section	-2	$(\hbar c)^2$	cm^2	1	cm^2

2.1.7 Summary of Natural, Hybrid, and cgs Units

To summarize the three systems of units we have discussed.

cgs: cm, sec, gm fundamental, other quantities derived from laws of nature

hybrid: cm, sec, MeV fundamental, other quantities derived from laws of nature

natural: c, \hbar, MeV fundamental (c and \hbar unitless and unit magnitude; 1 MeV = an amount we know from other work), other quantities derived from laws of nature

Conversion of algebraic relations

cgs or hybrid to natural: Put $c = \hbar = 1$. e.g., $E = mc^2 \rightarrow m$; $p_x = \hbar k_x \rightarrow k_x$.

natural to cgs or hybrid: Easiest just to remember, or look up, relations. e.g., $E = m \rightarrow mc^2$.

Can instead insert factors of c and \hbar needed on each side to balance units. e.g., $E(\text{energy units}) = m(\text{energy-sec}^2/\text{cm}^2 \text{ units}) \times ?$, where $?$ must be c^2 .

Conversion of numeric quantities

natural to hybrid to cgs: go from left to right in Wholeness Chart 2-1.

cgs to hybrid to natural: go from right to left, dividing rather than multiplying.

Note in the chart, that the Lagrangian and Hamiltonian densities in cgs have energy/(length)³ dimensions. In natural units these become (energy)⁴ or (mass)⁴. The action is the integral of the Lagrangian density over space and time. In cgs this is energy-time; in natural units it is M^0 .

2.1.8 QFT Approach to Units

QFT starts with familiar relations for quantities from the cgs system, e.g., $p_x = \hbar k_x$, and then expresses them in terms of natural units, e.g., $p_x = k_x$. The theory is then derived, and predictions for scattering and decay interactions made, in terms of natural units. Finally, before comparing these predictions to experiment, they are converted to the hybrid system, which is the system experimentalists use for measurement.

In summary:

relations in cgs → same relations in natural units → develop theory in natural units → predict experiment in natural units → same predictions in hybrid (MeV-cm-sec) units.

How QFT uses different systems of units

The first arrow above is easy. Just set $c = \hbar = 1$. For the last arrow, use Wholeness Chart 2-1. All of the other arrows are what the remainder of this book is all about.

You may wonder if this conversion to natural units is really all that worthwhile, as its primary value seems to be in saving the extra effort of writing out c and \hbar in all our equations. You may have a point, but natural units are what everyone working in QFT uses, so you should resign yourself to getting used to them.

2.2 Notation

We shall use a notation in which we define contravariant components x^μ of the 4D position vector, where X_i are 3D Cartesian coordinates, as

$$x^\mu = \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} = \begin{bmatrix} ct \\ X_1 \\ X_2 \\ X_3 \end{bmatrix} = (ct, X_i)^T, \quad \mu = 0, 1, 2, 3 \quad i = 1, 2, 3 \quad c = 1 \text{ in natural units.} \quad (2-1)$$

Contravariant 4D position components for us = 3D Cartesian coordinates plus time

Contravariant components, and their siblings described below, are essential to relativity theory, and QFT is grounded in special relativity. To avoid confusion, whenever we want to raise a component to a power, we will use parenthesis, e.g., the contravariant z component of the position vector squared is $(x^3)^2$. From henceforth, we will use natural units, and not write c .

From special relativity, we know the differential proper time passed on an object (with $c=1$) is

$$(d\tau)^2 = (dt)^2 - dX_1 dX_1 - dX_2 dX_2 - dX_3 dX_3. \quad (2-2)$$

If we define covariant components of the 4D position vector as

$$x_\mu = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} t \\ -X_1 \\ -X_2 \\ -X_3 \end{bmatrix} = (t, -X_i)^T, \quad (2-3)$$

Covariant components have negative 3D Cartesian coordinates

then (2-2) becomes

$$(d\tau)^2 = dx^0 dx_0 + dx^1 dx_1 + dx^2 dx_2 + dx^3 dx_3 = \underbrace{dx^\mu dx_\mu}_{\text{summation convention}}, \quad (2-4)$$

where on the RHS, we have introduce the shorthand Einstein summation convention, in which repeated indices are summed, and which we will use throughout the book. If we do not wish to sum when repeated indices appear, we will underline the indices, e.g., $\underline{dx^\mu} dx_\mu$ means no summation.

Repeated indices means summation

We can obtain (2-3) by means of a matrix operation on (2-1), i.e.,

$$x_\mu = g_{\mu\nu} x^\nu = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}}_{g_{\mu\nu}} \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} t \\ X_1 \\ X_2 \\ X_3 \end{bmatrix}, \quad (2-5)$$

Getting covariant components from contravariant ones

where the matrix $\underline{g}_{\mu\nu}$ is known as the metric tensor. Its inverse, $\underline{g}^{\mu\nu}$, has the exact same form,

$$\delta_\alpha^\nu = g_{\alpha\mu} g^{\mu\nu} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}}_{g^{\mu\nu}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \quad (2-6)$$

Contravariant and covariant forms of the metric tensor

where $\underline{\delta}_\alpha^\nu$ is the Kronecker delta. With the metric tensor and its inverse, we can re-write (2-4) as

$$(d\tau)^2 = g_{\mu\nu} dx^\mu dx^\nu = g^{\mu\nu} dx_\mu dx_\nu. \quad (2-7)$$

Partial derivatives with respect to x^μ and x_μ , often designated by $\partial^\mu \phi = \phi^{,\mu}$ and $\partial_\mu \phi = \phi_{,\mu}$, are

$$\partial^\mu = \frac{\partial}{\partial x_\mu} = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x_i} \right)^T = \left(\frac{\partial}{\partial t}, -\frac{\partial}{\partial X_i} \right)^T \quad \text{and} \quad \partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x^i} \right)^T = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial X_i} \right)^T. \quad (2-8)$$

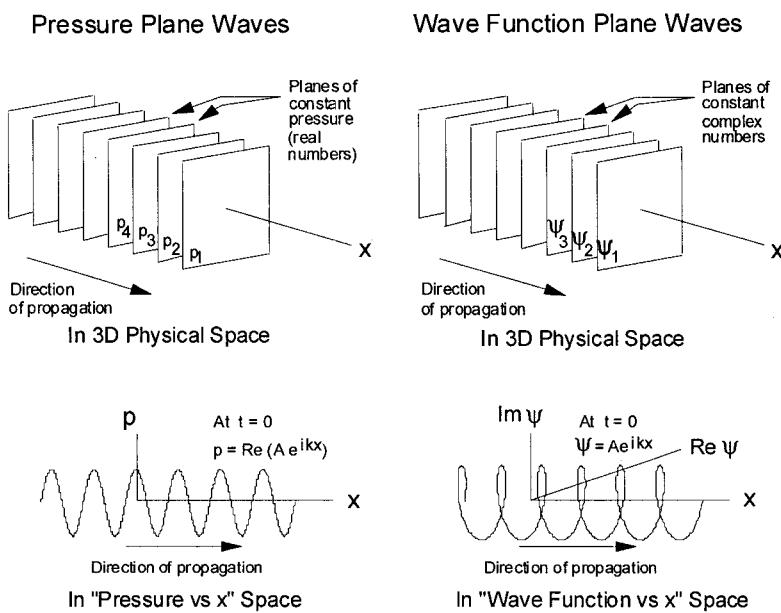
Note the spatial parts of x^μ and ∂^μ have opposite signs.

In general (see Prob. 4), we can raise or lower indices of any 4D vector w^μ using the (covariant) metric tensor and its inverse, the contravariant metric tensor, via $w^\mu = g^{\mu\nu} w_\nu$ and $w_\mu = g_{\mu\nu} w^\nu$.

Quantities for a single particle will be written in lower case, e.g., p_μ is the 4-momentum for a particle; for a collection of particles, in upper case, e.g., P_μ is 4-momentum for a collection of particles. Density values will be in script form, e.g., \mathcal{H} for Hamiltonian density.

Further, as one repeatedly sums p_μ and x^μ in QFT relations, we will employ the common streamlined notation $p_\mu x^\mu = px$ (the 4D inner product of 4 momentum and 4D position vectors.)

2.3 Classical vs Quantum Plane Waves



As we will be dealing throughout the book with quantum plane waves, the following quick review of them is provided.

Real vs complex (quantum) plane waves

Figure 2-1 illustrates the analogy between classical and quantum waves. Pressure plane waves, for example, can be represented as planes of constant *real* numbers (pressures) propagating through space. Particle wave function plane waves can be represented as planes of constant *complex* numbers (thus,

Figure 2-1. Classical vs Quantum Plane Waves

constant phase angle) propagating through space. Theoretically, the planes extend to infinity in the y and z directions. The lower parts of Figure 2-1 plot the numerical values of the waves on each plane vs. spatial position at a given instant of time. The complex wave has two components to plot; the real wave, only one. Plane wave packets for both pressure and wave function waves can be built up by superposition of many pure sinusoids, like those shown. (Though, as we will see, QFT rarely has need for wave packets.)

2.4 Review of Variational Methods

2.4.1 Classical Particle Theory

Recall, from classical mechanics, that, given the Lagrangian L for a particle, which is the kinetic energy minus the potential energy,

$$L = T - V = \sum_{i=1}^3 \frac{1}{2} m (\dot{x}^i)^2 - V(x^1, x^2, x^3) = \frac{\mathbf{p}^2}{2m} - V, \quad (2-9)$$

we can find the 3D equations of motion for the particle by the Euler-Lagrange equation, i.e.,

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right) - \frac{\partial L}{\partial x^i} = 0. \quad (2-10)$$

This, with (2-9), readily reduces to Newton's second law.

For a system of particles, we need only add an extra kinetic and potential energy term to (2-9) for each additional particle. For relativistic particles, we merely need to use relativistic kinetic and potential energy terms in (2-9), instead of Newtonian terms.

Recall also, that given the Lagrangian, we could find the Hamiltonian H , via the Legendre transformation,

$$H = p_i \dot{x}^i - L, \quad \text{where } p_i = \frac{\partial L}{\partial \dot{x}^i} = m \dot{x}^i. \quad (2-11)$$

p_i is the conjugate, or canonical, momentum of x^i . (Note that a contravariant component in the denominator is effectively equivalent to a covariant component in the entire entity, and vice versa.)

It is an important point that by knowing any one of H , L , or the equations of motion, we can readily deduce the other two using (2-9) through (2-11). That is, each completely describes the particle(s) and its (their) motion.

Equivalent entities

$$\text{Lagrangian } L \leftrightarrow \text{equations of motion} \leftrightarrow \text{Hamiltonian } H$$

Hence, when we defined first quantization in Chap. 1 as i) keeping the classical Hamiltonian and ii) changing Poisson brackets to commutators, we could just as readily have used the Lagrangian L or the equations of motion [for $x^i(t)$] for i) instead.

2.4.2 Pure Mathematics

We can apply the mathematical structure of the prior section to any kind of system, even some having nothing to do with physics. That is, if any system has a differential equation of motion (for example, an economic model), then one can find the Lagrangian for that system, as well as the Hamiltonian, the conjugate momentum, and more. So the mathematics derived for classical particles can be extrapolated and used to advantage in many other areas. Of course, one must then be careful in interpretation of the Hamiltonian, and similar quantities. The Hamiltonian, for example, will not, in general, represent energy, though many behavioral analogies (like conservation of H , etc.) will exist that can greatly aid in analyses of these other systems.

2.4.3 Classical Field Theory

Classical field theory is analogous in many ways to classical particle theory. Instead of the Lagrangian L , we have the Lagrangian density \mathcal{L} . Instead of time t as an independent variable, we have $x^\mu = x^0, x^1, x^2, x^3 = t, x^i$ as independent variables. Instead of a particle described by $x^i(t)$, we

*Definition of
classical mechanics
Lagrangian L*

*Governing equation
= Euler-Lagrange
equation*

*Legendre
transformation
 $H \leftrightarrow L$*

*L , H , and equations
of motion all tell us
the same thing*

*Variational math
can be applied to
many diverse fields*

have a field value described by $\phi(x^\mu)$ [or $\phi^r(x^\mu)$, where r designates different field types, or possibly, different spatial components of the same vector field (like \mathbf{E} or \mathbf{B} in electromagnetism).]

Particle Theory → Field Theory

$$L, H, \text{etc.} \rightarrow \mathcal{L}, \mathcal{H}, \text{etc.} \quad t \rightarrow x^\mu \quad x^i(t) \rightarrow \phi^r(x^\mu)$$

From these correspondences in variables, we can intuit the analogous forms of (2-9) through (2-11) [though we will derive the Euler-Lagrange equation afterwards] for fields. For the Lagrangian density, in terms of the kinetic energy density and potential energy densities of the field, we have

$$\mathcal{L} = \mathcal{T} - \mathcal{V}. \quad (2-12)$$

(Digressing here into the expressions for \mathcal{T} and \mathcal{V} in terms of the classical field ϕ would divert us away from our main purpose. In the next chapter we will see the form of these for a quantum field.)

The Euler-Lagrange equation for fields becomes

$$\boxed{\frac{d}{dx^\mu} \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}^r}_{,\mu} \right) - \frac{\partial \mathcal{L}}{\partial \phi^r} = 0.} \quad (2-13)$$

The Legendre transformation for the Hamiltonian density, with π_r being the conjugate momentum density of the field ϕ^r , is

$$\boxed{\mathcal{H} = \pi_r \dot{\phi}^r - \mathcal{L}, \quad \text{where } \pi_r = \frac{\partial \mathcal{L}}{\partial \dot{\phi}^r}.} \quad (2-14)$$

To see a real world example using (2-13), work through Prob. 6.

Compare (2-12) through (2-14) to (2-9) through (2-11), and note, that similar to particle theory, if we know any one of \mathcal{L} , \mathcal{H} or the equations of motion, we can readily find the other two. That is, they are equivalent, and in our first assumption of second quantization (see Chap. 1), we could take any one of the three (not just \mathcal{H} as we did in Chap. 1) as having the same form in quantum field theory as it did in classical field theory.

Derivation of Euler-Lagrange Equation for Fields

The fundamental assumption behind (2-13) is that the action of the field over an arbitrary 4D region Ω ,

$$S = \underbrace{\int_T \int_V \mathcal{L}(\phi, \phi_{,\mu}) d^3x dt}_{\mathcal{L}} = \int_{\Omega} \mathcal{L}(\phi, \phi_{,\mu}) d^4x, \quad (2-15)$$

where $d^4x = d^3x dt$ is an element of 4D volume, is stationary. More precisely, consider a virtual variation in ϕ of

$$\phi(x^\mu) \rightarrow \phi(x^\mu) + \delta\phi(x^\mu), \quad (2-16)$$

where the variation vanishes on the surface $\Gamma(\Omega)$ bounding the region Ω , i.e., $\delta\phi = 0$ on Γ . The “surface” here is actually three dimensional (rather than 2D), because it bounds a 4D region. This restriction on $\delta\phi$ is reasonable for a region Ω large enough so the field ϕ vanishes at its boundary.

For S to be stationary under the variation, we must have

$$\delta S = 0. \quad (2-17)$$

Using (2-17) in (2-15), we have

$$\delta S = \int_{\Omega} \left\{ \frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \delta\phi_{,\mu} \right\} d^4x = \int_{\Omega} \left\{ \frac{\partial \mathcal{L}}{\partial \phi} \delta\phi + \underbrace{\frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \frac{\partial}{\partial x^\mu} \delta\phi}_{\text{term } Z} \right\} d^4x. \quad (2-18)$$

With the last term on the RHS of (2-18), which we label “ Z ” here, re-written using

*Analogous entities
in particle and field
theories*

*Intuitive deduction
of field relations
from particle ones*

*\mathcal{L} , \mathcal{H} , and eqs of
motion all tell us
the same thing*

*Formal derivation
of Euler-Lagrange
eq for fields*

$$\frac{\partial}{\partial x^\mu} \left(\frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \delta\phi \right) = \left(\frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \right) \delta\phi + \underbrace{\frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \frac{\partial}{\partial x^\mu} \delta\phi}_{\text{term } Z}, \quad (2-19)$$

we can express (2-18) as

$$\begin{aligned} \delta S &= \int_{\Omega} \left\{ \frac{\partial \mathcal{L}}{\partial \phi} \delta\phi - \left(\frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \right) \delta\phi \right\} d^4x + \underbrace{\int_{\Omega} \frac{\partial}{\partial x^\mu} \left(\frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \delta\phi \right) d^4x}_{= \int_{\Gamma} n_\mu \left(\frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \delta\phi \right) d^3x}. \end{aligned} \quad (2-20)$$

The last term in the above relation can, via the 4D version of Gauss's divergence theorem, be converted into an integral over the surface Γ , as we show under the downward pointing bracket. In that integral, n_μ is the unit length vector normal to the surface Γ at every point on the surface, and it forms an inner product with the quantity in brackets by virtue of the summation over μ . Since we stipulated at the outset that $\delta\phi = 0$ on this surface, the last term in (2-20) must equal zero.

From (2-17), the first integral in (2-20),

$$\underbrace{\int_{\Omega} \left\{ \frac{\partial \mathcal{L}}{\partial \phi} - \left(\frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} \right) \right\} \delta\phi d^4x}_{\text{must } = 0} = 0, \quad (2-21)$$

for any possible variation of ϕ , i.e., for any possible $\delta\phi$ everywhere within Ω . The only way this can happen is if the quantity inside the brackets equals zero. But this is just (2-13) for one field. A similar derivation can be made for each additional type of field, i.e., for different values of r in (2-13), and thus, we have proven (2-13).

End of derivation

2.4.4 Real vs. Complex Fields

In classical theory we typically deal with real fields, such as the displacement at every point in a solid or fluid, or the value of the \mathbf{E} field in electrostatics. However, given our experience in NRQM, where complex wave functions were everywhere, so will we find that in QFT, quantum fields are commonly complex. Nothing in the above limited our derivation to real fields, so all of the relationships in this Sect. 2.4 are valid for complex fields, as well.

*Classical field real;
quantum fields
usually complex*

2.5 Classical Mechanics: An Overview

Wholeness Chart 2-2 is a summary of the key relations in all of classical physical theory (from the variational viewpoint.) The chart is intended primarily as an overview of past courses and as a lead in to quantum field theory, so a detailed study of it is not really warranted at this time. We have other fish to fry. I did say in the preface that we would focus on the essentials, and this chart is provided solely as i) a reference (which may aid some readers in studying for graduate oral exams), and ii) a lead in to technical details regarding Poisson brackets and second quantization.

*Variational
classical mechanics
overview in
Wholeness Chart
2-2*

The full theory behind Wholeness Chart 2-2 can be found in Goldstein (see Preface). The most important points regarding field theory, as represented in the chart, and which we will need to understand, are listed below.

Note that if the chart relationships are used for simple systems with only Cartesian coordinates, one need only take $x^i \rightarrow X_i$ everywhere and leave everything else as it is.

2.5.1 Key Concepts in Field Theory

1. Generalized coordinates do not have to be independent, and the Lagrangian can have second and/or higher coordinate derivatives. However, in most cases, including those of Wholeness Chart 2-2, the coordinates are independent and the Lagrangian only contains first derivatives.
2. The $x^i(t)$ for particles are not quite the same thing as the x^i for fields. The former are not independent variables, but functions of time t that represent the particle position at any given t .

*For us: q^i are
independent and
only 1st derivatives
in L, \mathcal{L}*

Wholeness Chart 2-2

	Mathematically	Non-relativistic Particle
Independent variable(s)	t	t
Coordinates	$q_i = q_i(t), i = 1, \dots, n$ (generalized)	$x^i = x^i(t), i = 1, 2, 3$ (contravariant)
Lagrangian density	see Fields columns	not applicable for particle
Lagrangian	$L = L(q_i, \dot{q}_i, t)$	$L = L(x^i, \dot{x}^i, t) = \sum_i \frac{1}{2} m(\dot{x}^i)^2 - V(x^i, t)$
Action	$S = \int L dt$	as at left
Euler- Lagrange equation (From $\delta S = 0$)	$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial q_i} = 0$	$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^i} \right) - \frac{\partial L}{\partial x^i} = 0$
Equations of motion for chosen coordinates	use explicit form for L in Euler-Lagrange equation	$m\ddot{x}^i = -\frac{\partial V}{\partial x^i}$ usually V not function of t
Conjugate momentum density ; total	see Fields columns ; $p_i = \frac{\partial L}{\partial \dot{q}_i}$	n/a ; $p_i = \frac{\partial L}{\partial \dot{x}^i} = m\dot{x}^i = -m\dot{x}_i$
Physical momentum density ; total	not relevant, purely math	n/a ; same as conjugate momentum
Alternative formulation	q_i, p_i and $L = L(q_i, p_i, t)$	x^i, p_i and $L = p^2/2m - V(x^i, t)$
Hamiltonian density; total	see Fields; $H = p_i \dot{q}_i - L$ (pure math)	n/a ; $H = p_i \dot{x}^i - L = p^2/2m + V$
Hamiltonian's Equations of Motion for conjugate variables	$\dot{p}_i = -\frac{\partial H}{\partial q_i} \quad \dot{q}_i = \frac{\partial H}{\partial p_i}$	$\dot{p}_i = -\frac{\partial H}{\partial x^i} = -\frac{\partial V}{\partial x^i} \quad \dot{x}^i = \frac{\partial H}{\partial p_i}$
Poisson Brackets, definition	for $u = u(q_i, p_i, t)$, $v = v(q_i, p_i, t)$ $\{u, v\} = \frac{\partial u}{\partial q_i} \frac{\partial v}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial v}{\partial q_i}$	for $u = u(x^i, p_i, t)$, $v = v(x^i, p_i, t)$ $\{u, v\} = \frac{\partial u}{\partial x^i} \frac{\partial v}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial v}{\partial x^i}$
Equations of motion in terms of Poisson brackets	i) for $v = H \quad \frac{du}{dt} = \{u, H\} + \frac{\partial u}{\partial t}$	i) for $v = H \quad \frac{du}{dt} = \{u, H\} + \frac{\partial u}{\partial t}$
i) any variable	ii) for i) plus $u = q_i$ or p_i	ii) for i) plus $u = x^i$ or p_i
ii) conjugate variables	$\dot{p}_i = \{p_i, H\} = -\frac{\partial H}{\partial q_i}; \quad \dot{q}_i = \{q_i, H\} = \frac{\partial H}{\partial p_i}$	$\dot{p}_i = \{p_i, H\} = -\frac{\partial H}{\partial x^i}; \quad \dot{x}^i = \{x^i, H\} = \frac{\partial H}{\partial p_i}$
Poisson Brackets for conjugate variables	$\{q_i, p_j\} = \delta_{ij} \quad \{q_i, q_j\} = \{p_i, p_j\} = 0$	$\{x^i, p_j\} = \delta^i_j \quad \{x^i, x^j\} = \{p_i, p_j\} = 0$

Summary of Classical (Variational) Mechanics

Non-relativistic Fields	Relativistic Particle	Relativistic Fields
$x^i, t \quad i = 1, 2, 3$	t	$x^\mu \quad \mu = 0, 1, 2, 3$
$\phi^r(x^i, t) \quad r = \text{field type} = 1, \dots, n$	$x^i = x^i(t), \quad i = 1, 2, 3$	$\phi^r(x^\mu) \quad r = \text{field type} = 1, \dots, n$
$\mathcal{L} = \mathcal{L}(\phi^r, \dot{\phi}^r, \partial_i \phi^r, x^i, t)$	not applicable for particle	$\mathcal{L} = \mathcal{L}(\phi^r, \partial_\mu \phi^r, x^\mu)$
$L = \int \mathcal{L} d^3x$	$L(x^i, v^i, t) = -m\sqrt{1-v^2} - V$	$L = \int \mathcal{L} d^3x$
$S = \int L dt = \int \mathcal{L} d^3x dt$	$S = \int L dt$	$S = \int L dt = \int \mathcal{L} d^3x dt$
$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}^r} \right) + \frac{d}{dx^i} \left(\frac{\partial \mathcal{L}}{\partial \phi_{,i}^r} \right) - \frac{\partial \mathcal{L}}{\partial \phi^r} = 0$	$\frac{d}{dt} \left(\frac{\partial L}{\partial v^i} \right) - \frac{\partial L}{\partial x^i} = 0$	$\frac{d}{dx^\mu} \left(\frac{\partial \mathcal{L}}{\partial \phi_{,\mu}^r} \right) - \frac{\partial \mathcal{L}}{\partial \phi^r} = 0$
\mathcal{L} above in Euler-Lagrange equation	$\frac{d}{dt} \left(\frac{\partial L}{\partial v^i} \right) = -\frac{\partial V}{\partial x^i}; \quad V(x^i, v^i)$	\mathcal{L} above in Euler-Lagrange equation
$\pi_r = \frac{\partial \mathcal{L}}{\partial \dot{\phi}^r}; \quad \Pi_r = \int \pi_r d^3x$	n/a; $p_i = \frac{\partial L}{\partial v^i} = \frac{mv^i}{\sqrt{1-v^2}} - \frac{dV}{dv^i}$	$\pi_r = \frac{\partial \mathcal{L}}{\partial \dot{\phi}^r}; \quad \Pi_r = \int \pi_r d^3x$
$\phi_i = \pi_r \frac{\partial \phi^r}{\partial x^i}; \quad p_r = \int \phi_i d^3x$	n/a; = conjugate momentum	$\phi_i = \pi_r \frac{\partial \phi^r}{\partial x^i}; \quad p_r = \int \phi_i d^3x$
$\mathcal{L} = \mathcal{L}(\phi^r, \pi_r, \partial_i \phi^r, x^i, t)$	$L = L(x^i, p_i, t)$	$\mathcal{L} = \mathcal{L}(\phi^r, \pi_r, \partial_i \phi^r, x^i, t)$
$\mathcal{H} = \pi_r \dot{\phi} - \mathcal{L}; \quad H = \int \mathcal{H} d^3x$	n/a; $H = p_i v^i - L = T + V$	$\mathcal{H} = \pi_r \dot{\phi} - \mathcal{L}; \quad H = \int \mathcal{H} d^3x$
same form as Relativistic Fields	$\dot{p}_i = -\frac{\partial H}{\partial x^i} = -\frac{\partial V}{\partial x^i} \quad \dot{x}^i = \frac{\partial H}{\partial p_i}$	$\dot{\pi}_r = -\frac{\delta \mathcal{H}}{\delta \dot{\phi}^r} \quad \dot{\phi}^r = \frac{\delta \mathcal{H}}{\delta \pi_r}$ where $\frac{\delta}{\delta \phi^r} = \frac{\partial}{\partial \phi^r} - \frac{d}{dx^i} \left(\frac{\partial}{\partial \phi_{,i}^r} \right)$
same form as Relativistic Fields	same form as Non-relativistic Particle, but different meaning for p_i	for $u = u(\phi^r, \pi_r, t)$, $v = v(\phi^r, \pi_r, t)$ $\{u, v\} = \left(\frac{\delta u}{\delta \phi^r} \frac{\delta v}{\delta \pi_r} - \frac{\delta u}{\delta \pi_r} \frac{\delta v}{\delta \phi^r} \right) \delta(\mathbf{x} - \mathbf{y})$
same form as Relativistic Fields	same form as Non-relativistic Particle	i) for $v = H \quad \dot{u} = \frac{du}{dt} = \{u, H\} + \frac{\partial u}{\partial t}$ ii) for i) plus $u = \phi^r$ or π_r $\dot{\pi}_r = \{\pi_r, H\} = \frac{\delta H}{\delta \phi^r}; \quad \dot{\phi}^r = \{\phi^r, H\} = \frac{\delta H}{\delta \pi_r}$
same form as Relativistic Fields	same form as Non-relativistic Particle	$\{\phi^r, \pi_s\} = \delta^r_s \delta(\mathbf{x} - \mathbf{y}); \quad \{\phi^r, \phi^s\} = \{\pi_r, \pi_s\} = 0$

The latter are independent variables, and not functions of time, but fixed locations in space upon which the value for the field (and other things like energy density) depends. The field and related density type quantity values also depend on the other independent variable, time.

$x^i(t)$ for particles; x^i independent of time for fields

r label = different field types or different components of field

3. Different values for the r label for fields can represent

- i) completely different fields, as well as
- ii) different components in spacetime of the same vector field.

4. In general, the Hamiltonian does not have to represent energy, and can be simply a quantity which obeys all of the mathematical relations shown in the chart. However, in the application of analytical mechanics, it proves immensely useful if the Hamiltonian is, in fact, energy (or an energy operator.) Similarly, in general, the Lagrangian does not have to equal kinetic energy minus potential energy (i.e., $T - V$), and can simply be a quantity which gives rise via the Lagrange equation to the correct equation(s) of motion (called field equations for fields.)

Fortunately, in field theory, the Lagrangian density can be represented as kinetic energy density minus potential energy density, and the Hamiltonian density turns out to be total energy density. These correspondences carry over to quantum field theory.

In our work,
 L always = $T - V$;
 H always = $T + V$

5. For fields,

$$\frac{\partial \phi}{\partial t} = \frac{d\phi}{dt} = \dot{\phi} \quad (2-22)$$

This is generally not true for other quantities. For an explanation of this, see Box 2-1.

For fields, partial and total time derivatives are the same thing

Box 2-1. Time Derivatives and Fields

Any field, say ϕ , is a function of space and time, i.e., $\phi = \phi(x^i, t)$, where x^i is an independent variable representing a coordinate (non-moving) point in space upon which field quantities depend.

Note that the total time derivative is

$$\frac{d\phi}{dt} = \frac{\partial \phi}{\partial x^i} \frac{dx^i}{dt} + \frac{\partial \phi}{\partial t}$$

But since x^i is an independent variable like time, and hence is not a function of time, its time derivative above is zero. Thus,

$$\frac{d\phi}{dt} = \frac{\partial \phi}{\partial t} = \dot{\phi}$$

So the partial time derivative and the total time derivative of a field are one and the same thing, and both are designated with a dot over the field.

Note that quantities other than fields do not, in general, have this property. (See the Poisson bracket blocks in the fields section of Wholeness Chart 2-2.) It is necessary, therefore, when talking about time derivatives of quantities other than the fields themselves, to specify precisely whether we mean the total or partial derivative with respect to time.

The conclusions reached here apply in both the relativistic and non-relativistic field cases.

6. There are two kinds of momenta, conjugate and physical. In some cases these are the same, but in general they are not. For fields, each of these can be either total momentum or momentum density. Box 2-2 derives the relations between conjugate and physical momentum densities.

2 kinds of momenta.
Each kind can be total or density

7. Key difference between the particle and field approaches.

For a single particle, particle position coordinates are the generalized coordinates and particle momentum components are its conjugate momenta. For fields, each field is itself a generalized coordinate and each field has its own conjugate momentum (density). As noted, this field conjugate momentum (density) is different from the physical momentum (density) that the field possesses.

Generalized coords
Particle: x^i
Field: ϕ^r

Box 2-2. Conjugate and Physical Momentum Densities

The relationship between physical momentum density and conjugate momentum density for fields is not so intuitive. It can be derived using the Hamiltonian density expression for fields given in Wholeness Chart 2-2,

$$\mathcal{H} = \pi_r \dot{\phi} - \mathcal{L}.$$

Mathematically, this expression would not necessarily have to be energy density. If on physical grounds, we require that it be energy density, then it must also be expressible as

$$\mathcal{H} = \mu_i v^i - \mathcal{L},$$

where μ_i is physical momentum density at a point in the field and v^i is the velocity of the field at the same point. For a fluid where the field value is displacement, this is the velocity of the fluid particle at the point. For an electromagnetic \mathbf{E} field, this would be the velocity at which one would travel in order to see a constant \mathbf{E} value.

Equating the above two equations, we see

$$\mu_i v^i = \pi_r \dot{\phi}^r \rightarrow \mu_i \frac{\partial x^i}{\partial t} = \pi_r \frac{\partial \phi^r}{\partial t},$$

where we note carefully that our x^i here is the position coordinate of a point fixed relative to the field and thus is time dependent, like particle position time dependence. Further, the total derivative $v^i = dx^i/dt$ equals the partial derivative with respect to time $\partial x^i/\partial t$, since $x^i(t)$ in the present case is only a function of time.

From the above equation, by cancelling the ∂t on each side, we see

$$\mu_i = \pi_r \frac{\partial \phi^r}{\partial x^i}.$$

The partial derivative of ϕ^r with respect to either of our definitions of x^i (time dependent or independent) is the same because by definition partial derivative means we hold everything else (specifically time here) constant. Thus, the above relation holds in field theory when we consider the x^i as independent variables.

8. Note that it is common in QFT to refer to the field conjugate momentum density as simply the conjugate momentum, the Hamiltonian density as merely the Hamiltonian, and the Lagrangian density as the Lagrangian. This may be unfortunate, but you will learn to live with gleaning the exact sense of these terms from context.

The word “density” often dropped in field theory

9. The relativistic particle summary, as outlined in Wholeness Chart 2-2, is not, in the strictest sense, formulated covariantly. It describes relativistic behavior, but position and momentum are (non-Lorentz covariant) three vectors, and the Lagrangian and Hamiltonian are not world scalars (world scalars are invariant under Lorentz transformation.) Alternative approaches are possible using proper time for the independent variable and world vector (four vector) quantities for generalized coordinates and conjugate momenta. (Goldstein and Jackson [see Preface] show two different ways to do this.) In those treatments the Lagrangian and Hamiltonian are world scalars though the Hamiltonian does not turn out to be total energy. The approach taken here has been chosen because, in it, we have the advantage of having a Hamiltonian that represents total energy. Further, the parallel between relativistic particles and the usual treatment of relativistic fields becomes much more transparent.

Several ways to formulate variational relativistic theory

10. Some comment is needed on the several different equations of motion that one runs into.

A differential equation of motion is generally an equation that contains derivative(s) with respect to time of some entity, and has as its solution that entity expressed as an explicit function of time (and for fields, space, as well.) For example, $F^i = m\ddot{x}^i$ is the equation of motion for a particle, with $x^i(t)$ as its solution. There are in general two kinds of entities for which we have equations of motion. One is the generalized coordinates themselves. The other is any function of those coordinates, generally expressed as u or v in the next to last row of Wholeness Chart 2-2. (The

Eqs of motion exist for i) generalized coordinates, and ii) functions of those coordinates

first class is really a special case of the second, where, for example, u might equal the generalized coordinate itself.)

In Wholeness Chart 2-2, the equations of motion for generalized coordinates are expressed in three different but equivalent ways: the Lagrange equations formulation, the Hamilton's equations formulation, and the Poisson bracket formulation. These are all different expressions for describing the same behavior of the generalized coordinates of a given system via different differential equations. For any particular application, one of these formulations may have some advantage over the others.

The other class of equation of motion for a function of generalized coordinates, say u , can be expressed for the purely mathematical case (the others are analogous) as

$$\frac{du(\dot{q}_i, p_i, t)}{dt} = \frac{\partial u}{\partial q_i} \dot{q}_i + \frac{\partial u}{\partial p_i} \dot{p}_i + \frac{\partial u}{\partial t}. \quad (2-23)$$

Using Hamilton's equations for the time derivatives of q_i and p_i yields

$$\frac{du}{dt} = \underbrace{\frac{\partial u}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial H}{\partial q_i}}_{\text{Poisson bracket definition for } u \text{ and } H} + \frac{\partial u}{\partial t} = \{u, H\} + \frac{\partial u}{\partial t}, \quad (2-24)$$

Poisson bracket definition used in equation of motion

which is effectively the same equation of motion as (2-23), for the same coordinate u , expressed instead in terms of a Poisson bracket. See the first line of the next to last row block in Wholeness Chart 2-2.

Summary of Forms of Differential Equations of Motion

For generalized coordinates (all three below are equivalent)

1. Lagrangian into Euler-Lagrange equation
2. Hamilton's equations of motion
3. Poisson bracket notation for 2 above

Forms for differential equations of motion

For a function of those generalized coordinates (both below are equivalent)

1. Total time derivative expressed as partial derivatives (see (2-23), not shown in Wholeness Chart 2-2.)
2. Total time derivative expressed in terms of Poisson bracket notation (see (2-24), also shown in Wholeness Chart 2-2.)

11. The field equations (equations of motion) for relativistic fields keep the exact same form in any inertial frame of reference*, i.e., they are Lorentz invariant. Components of four vectors in any of the equations can change from frame to frame, but the relationship between these components expressed in the field equation must remain inviolate. Four vectors transform via the Lorentz transformation of course, and are termed Lorentz covariant. Four scalars (world scalars) are invariant under a Lorentz transformation and look exactly the same to any observer. (e.g., Rest mass m of a free particle is a four scalar, where $m^2 = p^\mu p_\mu$. Another observer in a different (primed) frame could measure a different four momentum p'_μ , but would find the same rest mass via $p'^\mu p'_\mu = m^2$.)

Lorentz invariance (scalars and form of equations) and covariance (vectors and tensors)

Note the result of demanding that the Euler-Lagrange equation (i.e., the field equation) (2-13) be Lorentz invariant. We know that, within that equation, x^μ, ϕ^r , and derivatives of x^μ are Lorentz covariant or invariant. So, in order for the whole equation to be Lorentz invariant, the Lagrangian density \mathcal{L} must be invariant, i.e., a world scalar.

\mathcal{L} is a Lorentz invariant scalar

* To be completely accurate, this is true strictly for Einstein synchronization, the synchronization convention of Lorentz transformations. If you are not a relativity expert, please don't worry about this fine point.

Since d^4x is also a Lorentz (world or four) scalar (i.e., four volume is the same in any Lorentz coordinate system, just as 3D volume is the same in any Cartesian system), the action S (see Chart 2-2) must be a Lorentz scalar as well. Note though that the total Lagrangian L is *not* a four scalar since d^3x is not a four scalar. Neither is the Hamiltonian or the Hamiltonian density. To see this, do Prob. 6.

L, H, and H̄ are not Lorentz scalars.

End of Key Concepts in Field Theory points

Box 2-3. Unitary Transformations in Quantum Theories

A *unitary transformation* is called unitary because it has unit length, i.e., when it operates on (transforms) a state vector, the magnitude of the state vector is unchanged. It is the complex space analogue of an *orthogonal transformation* in Cartesian coordinate space, which, when acting on a (real number) vector in that space, rotates the vector but does not stretch or compact it. A unitary transformation can be thought of as “rotating” a (complex number) state vector in Hilbert space (the complex space where each coordinate axis is an eigenvector) without changing the length of the vector. In NRQM, the square of the absolute value of the state vector is the square of its length, and this is the probability density for measuring the particle. This means a unitary transformation of a state vector leaves the probability of detecting the particle unchanged.

Recall, from classical mechanics, that an orthogonal transformation represented by a matrix \mathbf{A} has an inverse equal to the transpose of that matrix, i.e., $\mathbf{A}^{-1} = \mathbf{A}^T$. In the complex space of state vectors, a unitary transformation U has an analogous form for its inverse, the complex conjugate transpose, i.e., $U^{-1} = U^\dagger$ and $U^\dagger U = 1$. Thus, $|U| = 1$.

An example of a unitary matrix, consider $U = e^{-iHt}$, where H is the Hamiltonian operator. By inspection, one knows its length = 1 in complex space, and also $U^\dagger U = 1$. So it is a unitary transformation.

Wholeness Chart 2-3. Unitary vs Orthogonal Transformations

	3D Cartesian Space (Real)	Hilbert Space (Complex)
Magnitude conserving transformation	Orthogonal $\mathbf{A} = \text{matrix}$	Unitary $U = e^{iX}$
Effect on vector	rotates in real space	“rotates” in complex space
Physical effect	vector length unchanged	probability unchanged
Inverse	$\mathbf{A}^{-1} = \mathbf{A}^T$	$U^{-1} = U^\dagger$

How an exponential operator works

Do a Taylor expansion of $U = e^{-iHt}$ above about t , when U is operating on an energy eigenstate., i.e.,

$$U|\psi_E\rangle = e^{-iHt}|\psi_E\rangle = \left(1 - itH - \frac{1}{2}t^2H^2 + \dots\right)|\psi_E\rangle = \left(1 - itE - \frac{1}{2}t^2E^2 + \dots\right)|\psi_E\rangle = e^{-iEt}|\psi_E\rangle$$

So an operator in the exponent has the same effect in the exponent as it would if acting in the usual non-exponential way on an eigenstate state. This conclusion can be generalized to any state.

Note: Although it is common to write $U = e^{-iHt}$, it is implied that H (if you think of it as $i\partial/\partial t$) does not act on t . To be proper, the t should be placed before the H , as we did in the expansion above, but it usually is not done that way.

2.6 Shrödinger vs Heisenberg Pictures

In quantum theory, there are different methods by which one can describe state and operator behavior that all result in the same measurable quantity. That is, the underlying math differs, but the predictions one would make for experimentally measurable dynamic variables remain the same.

These different, but equivalent, ways are called different pictures and apply in the same way to all branches of quantum theory (NRQM, RQM, and QFT.) Most QM courses more elementary than this one use what is known as the Shrödinger picture, and that is, no doubt, what you unconsciously thought in terms of, when you did NRQM. We will review that, and then introduce what is called the Heisenberg picture, which helps immensely in QFT with developing theory and doing calculations. Note carefully, before we start, that these terms *do not* refer to the Shrödinger wave approach vs the Heisenberg matrix approach to QM. Everything we do will comprise the wave approach, not the matrix approach, but there are two distinct pictures within that approach, i.e.,

Shrödinger Wave Approach

1. Shrödinger picture
2. Heisenberg picture.

Heisenberg Matrix Approach

*Different pictures
in quantum theory*

We will review the Shrödinger picture and develop the Heisenberg picture in terms of NRQM, though the final results will be applicable to any branch of QM, including QFT.

2.6.1 The Shrödinger Picture

In QM, one has i) states (wave functions, particles, kets, state vectors), and ii) operators (such as momentum, the Hamiltonian, and the like), which act on those states. The real world value corresponding to any such operator that one would expect to measure in an experiment, i.e., the average value over many trials, is called the expectation value. The expectation value for any operator is typically designated with a bar over the operator and is found via the statistical relationship (with normalized wave function ψ)

$$\bar{\mathcal{O}} = \int \psi^* \mathcal{O} \psi d^3x = \langle \psi | \mathcal{O} | \psi \rangle. \quad (2-25)$$

The time derivative of the expectation value (2-25) (being what we would expect to measure in experiment for the rate of change of the corresponding dynamic variable) is

$$\frac{d\bar{\mathcal{O}}}{dt} = \frac{d}{dt} \langle \psi | \mathcal{O} | \psi \rangle = \left\langle \frac{\partial \psi}{\partial t} \middle| \mathcal{O} \right| \psi \rangle + \langle \psi \left| \frac{\partial \mathcal{O}}{\partial t} \right| \psi \rangle + \langle \psi \left| \mathcal{O} \right| \frac{\partial \psi}{\partial t} \rangle. \quad (2-26)$$

*Operator
expectation value =
“expected” or
mean measurement*

In the Shrödinger picture, the solutions to the Shrödinger equation

$$i \frac{\partial \psi_S}{\partial t} = H \psi_S \quad \text{or} \quad i \frac{\partial}{\partial t} |\psi\rangle_S = H |\psi\rangle_S \quad (2-27)$$

*Calculating
expectation value*

are the states ψ_S (or $|\psi\rangle_S$), which are time dependent. The subscript S indicates the Shrödinger picture (S.P.). In that picture, the operators are usually not time dependent. For example, using the familiar momentum operator $p_1^S = -i\partial/\partial x^1$ for the S.P. in the x^1 direction, with

$$\psi_S = A e^{-i(Et-\mathbf{p}\cdot\mathbf{x})} = |\psi\rangle_S \quad A^\dagger A = 1, \quad (2-28)$$

*In S.P., NRQM eq
of motion of state
(Shrödinger eq)*

(2-25) is

$$\bar{p}_1 = \int A^\dagger e^{i(Et-\mathbf{p}\cdot\mathbf{x})} \left(-i \frac{\partial}{\partial x^1} \right) A e^{-i(Et-\mathbf{p}\cdot\mathbf{x})} d^3x =_S \langle \psi | p_1^S | \psi \rangle_S, \quad (2-29)$$

An example

where the state is time dependent, but the operator p_1^S is not. That is, since the latter has no t in it,

$$\frac{dp_1^S}{dt} = \frac{\partial p_1^S}{\partial t} = 0. \quad (2-30)$$

*In S.P., NRQM eq of
motion of momentum
operator (p_1^S
constant in time)*

Equation (2-26) for p_1 is then

$$\frac{d\bar{p}_1}{dt} = \frac{d}{dt}\langle\psi|p_1^S|\psi\rangle = {}_S\left\langle\frac{\partial\psi}{\partial t}\left|p_1^S\right|\psi\right\rangle_S + {}_S\langle\psi|\frac{\partial p_1^S}{\partial t}|\psi\rangle_S + {}_S\langle\psi|p_1^S\left|\frac{\partial\psi}{\partial t}\right\rangle_S, \quad (2-31)$$

where we leave in the zero quantity of (2-30), because we will want to generalize this result to all operators, including those rare cases where S.P. operators are time dependent (such as the Hamiltonian when $V = V(t)$.) Using the Shrödinger equation (2-27) and its complex conjugate for the ket and bra time derivatives, respectively, in (2-31), we get

$$\frac{d\bar{p}_1}{dt} = {}_S\langle\psi| -i\left[p_1^S, H\right]|\psi\rangle_S + {}_S\langle\psi|\underbrace{\frac{\partial p_1^S}{\partial t}}_{=0}|\psi\rangle_S. \quad (2-32)$$

Eq of motion of momentum expectation value

(2-32) generalized to any operator in Chart 2-4

Recall the old NRQM adage that the expectation value of any operator without explicit time dependence that commutes with the Hamiltonian is conserved (its time derivative is zero.) Note that (2-27), (2-30), and (2-31)/(2-32) are equations of motion for the state, momentum operator, and momentum expectation value, respectively, in the Shrödinger picture. These are generalized to any state and operator in Wholeness Chart 2-4.

Note further that the partial time derivative $\partial/\partial t$ in the Shrödinger equation (2-27) acting on the ket is equivalent to the full time derivative d/dt by the same logic as that in Box 2-1. That is, the ket, or wave function, here is mathematically the same as a classical field, functionally dependent on the independent variables, x^i and t . So we can write the equation of motion for a state (i.e., the Shrödinger equation) with either a partial or total time derivative.

2.6.2 The Heisenberg Picture

The Shrödinger picture states and operators can be transformed to states and operators having different form via what is known as a unitary transformation (see Box 2-3). The particular unitary transformation for this is

$$U = e^{-iHt/\hbar}, \quad (2-33)$$

where states and operators transform as

$$\begin{aligned} U^\dagger |\psi\rangle_S &= |\psi\rangle_H & U^\dagger \mathcal{O}^S U &= \mathcal{O}^H \\ U |\psi\rangle_H &= |\psi\rangle_S & U \mathcal{O}^H U^\dagger &= \mathcal{O}^S. \end{aligned} \quad (2-34)$$

Transforming between Schrödinger and Heisenberg pictures

Note the effect of the first relation in (2-34) on our sample ket (2-28),

$$U^\dagger |\psi\rangle_S = e^{iHt} A e^{-i(Et - \mathbf{p} \cdot \mathbf{x})} = A e^{iEt} A e^{-i(Et - \mathbf{p} \cdot \mathbf{x})} = A e^{i\mathbf{p} \cdot \mathbf{x}} = |\psi\rangle_H. \quad (2-35)$$

We find that the state, which was time dependent in the S.P., is *time independent* in the Heisenberg picture (H.P.). This statement is generally true for any state. (Think through it, if you like, for a more general wave function state of several terms.)

Thus, the equation of motion for a state in the S.P. (2-27), becomes, in the H.P.,

$$\frac{d|\psi\rangle_H}{dt} = 0. \quad (2-36)$$

In H.P., eq of motion of state (state is constant in time)

Now take the time derivative of the second relation in the top row of (2-34), we have

$$\begin{aligned} \frac{d}{dt}(U^\dagger \mathcal{O}^S U) &= (iH) \underbrace{e^{iHt/\hbar} \mathcal{O}^S e^{-iHt/\hbar}}_{\mathcal{O}^H} + e^{iHt/\hbar} \underbrace{\left(\frac{\partial \mathcal{O}^S}{\partial t}\right)}_{\text{defined as } \partial \mathcal{O}^H / \partial t} e^{-iHt/\hbar} + \underbrace{e^{iHt/\hbar} \mathcal{O}^S e^{-iHt/\hbar}}_{\mathcal{O}^H} (-iH) \\ &= \frac{d\mathcal{O}^H}{dt} = -i[\mathcal{O}^H, H] + \underbrace{\frac{\partial \mathcal{O}^H}{\partial t}}_{=0 \text{ in this book}}. \end{aligned} \quad (2-37)$$

In H.P., eq of motion of operator

We will not be considering any operators that are time dependent in the S.P., so for us, the last term above will always be zero. Nonetheless, even in this case, we see that in the H.P., an operator time derivative can be non-zero, and thus, the operator, time dependent.

Because $H (= H^S$ by definition) commutes with itself, U and U^\dagger commute with H , so using $\mathcal{O}^S = H^S = H$ in the second relation on the top line of (2-34),

$$H = H^S = H^H. \quad (2-38)$$

Hamiltonian H has same form in S.P. and H.P.

Finally, for (2-32) expressed in terms of a general operator ($p_1^S \rightarrow \mathcal{O}^S$), we find, after inserting $UU^\dagger = 1$ where needed, that

$$\begin{aligned} \frac{d\bar{\mathcal{O}}}{dt} &= {}_S\langle\psi|UU^\dagger(-i[\mathcal{O}^S, H])UU^\dagger|\psi\rangle_S + {}_S\langle\psi|UU^\dagger\frac{\partial\mathcal{O}^S}{\partial t}UU^\dagger|\psi\rangle_S \\ &= {}_H\langle\psi|(-i[\mathcal{O}^H, H])|\psi\rangle_H + {}_H\langle\psi|\frac{\partial\mathcal{O}^H}{\partial t}|\psi\rangle_H. \end{aligned} \quad (2-39)$$

From which we see that the expectation value of an operator has exactly the same form in both pictures. This means that whichever picture we choose to work in, although the states and operators will be different, the predictions for quantities we can measure (dynamic variables) will be the same. So we can choose whichever system is easier to work with mathematically. For NRQM, this was the S.P. For QFT, as we will see, it is the H.P.

Eq of motion of expectation value has same form in S.P. and H.P.

Wholeness Chart 2-4. Shrödinger vs. Heisenberg Picture Equations of Motion

	States	Operators	Expectation Values
Shrödinger Picture	Time dependent $i\frac{d}{dt} \psi\rangle_S = H \psi\rangle_S$ (Shrödinger eq)	Usually time independent $\frac{d\mathcal{O}^S}{dt} = \frac{\partial\mathcal{O}^S}{\partial t} = 0$ usually	$\frac{d\bar{\mathcal{O}}}{dt} = {}_S\langle\psi -i[\mathcal{O}^S, H] + \frac{\partial\mathcal{O}^S}{\partial t} \psi\rangle_S$ $ \psi\rangle_S$ changes in time; \mathcal{O}^S usually const in time
Transform via $U = e^{-iHt/\hbar}$ \Downarrow	$U^\dagger \psi\rangle_S = \psi\rangle_H$	$U^\dagger\mathcal{O}^S U = \mathcal{O}^H$	$\frac{d\bar{\mathcal{O}}}{dt}$ invariant under the transformation
Heisenberg Picture	Time independent $\frac{d \psi\rangle_H}{dt} = 0$	Often time dependent $\frac{d\mathcal{O}^H}{dt} = -i[\mathcal{O}^H, H] + \frac{\partial\mathcal{O}^H}{\partial t}$ usually $= 0$	Same as Shrödinger picture above with sub and superscript $S \rightarrow H$ $ \psi\rangle_H$ const in time; \mathcal{O}^H often changes in time
Hamiltonian		$H^H = H^S = H$	
Key Relation	In S.P., the state eq of motion	In H.P., the operator eq of motion	In both pictures, expectation value and its equation of motion are the same, equally key.

2.6.3 Visualizing Schrödinger and Heisenberg Pictures

One can think of the S.P. as quantum waves (wave functions or kets) moving and evolving in time, but operators as constant (generally) in time. The H.P., by contrast, can be thought of as quantum waves frozen in time (static wave functions or time independent kets), with operators being the thing that moves and evolves. Either way, the expectation value (2-40) (what we would measure on average over many measurements) is the same, and so is its equation of motion.

$$\bar{\mathcal{O}} = {}_S\langle\psi|\mathcal{O}^S|\psi\rangle_S = {}_H\langle\psi|\mathcal{O}^H|\psi\rangle_H. \quad (2-40)$$

*S.P.: particle waves move, operators (usually) do not.
H.P.: waves frozen, operators evolve.
Measured values same in both.*

2.7 Quantum Theory: An Overview

Wholeness Chart 2-5, Summary of Quantum Mechanics, overviews the fundamental branches of quantum theory in much the same way that Wholeness Chart 2-2 overviews the fundamental branches of classical theory. These correspond to, and elaborate on, the bottom and top parts, respectively, of Wholeness Chart 1-1 in Chap. 1. (We will temporarily leave \hbar in our relations even though, in our units, it equals one, so that you, the reader, can see precisely where it comes into those, rather key, relations.)

Note particularly, that in Wholeness Chart 2-5, all relations and quantities are expressed in the Heisenberg picture. If it were expressed in the Shrödinger picture, then many quantities (i.e., operators) such as H , p_i , and the like would have to be expressed as expectation values. In the H.P., the equation of motion for an operator (see H.P. row in Wholeness Chart 2-4) has the same time dependence as the expectation value for that operator (the bra and ket in are constant in time in the right most block in that row.) That is, in the H.P. the operator equation of motion is the same as that of the expectation value. And the state (ket) equation of motion, which was quite critical in the S.P. (it is the Shrödinger equation), becomes rather meaningless, as the state is constant in time. So we can ignore the states in the H.P. summary of Wholeness Chart 2-5 and write the equations of motion in terms of the operators.

2.7.1 Classical vs. Quantum: Much is the Same

Note that everything in the first 12 blocks in each column of Chart 2-5 is the same as that in Chart 2-2, from the independent variables used through Hamilton's equations of motion. For example, the Hamiltonian H has the same form for a particle in quantum mechanics as it does for a classical particle. (Recall from Chap. 1, this was criterion number one for first quantization.)

2.7.2 Poisson Brackets vs. Commutators: Something is Different

However, note that the equation of motion for a dynamic variable, represented by u , changes from (2-24) in classical non-relativistic particle theory to

$$\frac{du}{dt} = \frac{-i}{\hbar} [u, H] + \frac{\partial u}{\partial t} \quad (2-41)$$

in NRQM. Equation(2-41), which you should have seen before in your NRQM studies, was discovered independently by early quantum theorists. Yet it was striking to everyone how closely it parallels its classical counterpart (2-24). The only difference is that the Poisson brackets have become commutators (with a factor of $-i/\hbar$ in front.)

Similarly, the Poisson bracket relations for conjugate variables in classical theory (last line, third column in Wholeness Chart 2-2) parallel the commutators (last line, third column of Wholeness Chart 2-5) discovered early on in the development of NRQM.

So, the classical non-relativistic particle and the NRQM theories mimic one another, with one difference. All relations remain the same except that the commutators of quantum theory correspond to Poisson brackets of classical theory (times an imaginary factor of \hbar .)

2.7.3 Quantization and the Correspondence Principle

According to the *correspondence principle*, in the macroscopic limit, our quantum relations must reduce to the usual classical relations. But in comparing the last two blocks in the third columns (NR particle and NRQM) of Wholeness Charts 2-2 and 2-5, this can only be true if

$$\underbrace{\{x^i, p_j\}}_{\substack{\text{classical} \\ \text{dynamic} \\ \text{variables}}} = \delta^i{}_j = \frac{-i}{\hbar} \underbrace{[x^i, p_j]}_{\substack{\text{quantum} \\ \text{operators}}}. \quad (2-42)$$

So the correspondence principle provides us with a key part of our method for quantization. That is, in going from classical theory to NRQM, we must take

$$\{x^i, p_j\} = \delta^i{}_j \xrightarrow{\text{1st quantization}} [x^i, p_j] = i\hbar \delta^i{}_j \quad (2-43)$$

Chart 2-5 summarizes QM

Chart 2-5 is in terms of H.P.

First 12 rows:
Classical NR
particle of Chart 2-2 same as NRQM of Chart 2-5.

Last 2 rows:
Classical NR
particle has
Poisson brackets;
NRQM has
commutators

Classical NR
particle theory
becomes NRQM if
Poisson brackets
converted to
commutators

Wholeness Chart 2-5

	Comments	Non-relativistic Quantum Mechanics
Independent variables through Hamilton's equations of motion		Same form as top 12 blocks of Wholeness Chart 2-2
Commutator brackets, definition		for $u = u(x^i, p_i, t)$, $v = v(x^i, p_i, t)$ $[u, v] = uv - vu$
Equations of motion in terms of commutator brackets i) any dynamic variable ii) conjugate variables	Correspondence principle: Classical \rightarrow Quantum $\{u, v\} \rightarrow \frac{-i}{\hbar}[u, v]$	i) for $v = H$ $\frac{du}{dt} = \frac{-i}{\hbar}[u, H] + \frac{\partial u}{\partial t}$ ii) for i) plus $u = x^i$ or p_i $\dot{p}_i = \frac{-i}{\hbar}[p_i, H] = -\frac{\partial H}{\partial x^i}; \dot{x}^i = \frac{-i}{\hbar}[x^i, H] = \frac{\partial H}{\partial p_i}$
Uncertainty principle		$[x^i, p_j] = i\hbar\delta^i_j \quad [x^i, x^j] = [p_i, p_j] = 0$

Of course, as noted in Chap. 1, we also keep the same form of the Hamiltonian (or equivalently, the Lagrangian) as we had classically.

2.7.4 Extrapolation to Field Theory

Shortly after understanding this, one gets the idea that perhaps the same thing can be done with field theory. So we try it. We postulate the same sort of bracket correspondence and see where it takes us. Does it indeed lead to a good theory, one that predicts the phenomena we observe? Very quickly we find that it does, and that new theory has come to be called *quantum field theory*. This means for going from our classical theory of fields to the quantum theory of fields is called second quantization, i.e.,

$$\{\phi^r(\mathbf{x}, t), \pi_s(\mathbf{y}, t)\} = \delta^r_s \delta(\mathbf{x} - \mathbf{y}) \xrightarrow{\text{2nd quantization}} [\phi^r(\mathbf{x}, t), \pi_s(\mathbf{y}, t)] = i\hbar \delta^r_s \delta(\mathbf{x} - \mathbf{y}) \quad (2-44)$$

where again, we keep the same form of the Hamiltonian (or equivalently, the Lagrangian) as we had classically. That is, as we develop QFT, we will use the same independent variables, the same sense for the Hamiltonian density as an energy density, the same Legendre transformation, the same Euler-Lagrange equation into which we will plug our Lagrangian density, the same conjugate momenta definitions, etc.

The delta function in $\mathbf{x} - \mathbf{y}$ in (2-44) ensures that we are only considering the field and its conjugate momentum density at the same point in space. We will see the role this plays in the mathematical development of the theory later.

Both of the processes (2-43) and (2-44) are formally called canonical quantization. They are canonical because it is the canonically conjugate variables - the generalized coordinates and their conjugate momenta - which are the center of attention. The term quantization arises because the metamorphosis of brackets, in going from the classical to quantum realm, changes the Poisson bracket relation for the canonical variables into the commutator, which is the mathematical basis of the uncertainty principle. The uncertainty principle is often called the quantum principle, hence the name *quantization*.

Quantization then, in a nutshell, is a means for deducing the governing quantum equations from knowledge of the classical macroscopic ones. We will begin to use it in the next chapter to develop our theory.

We guess: Classical relativistic field theory should become QFT if Poisson brackets converted to commutators

Quantization is a means for deducing quantum theory from classical theory.

Summary of Quantum Mechanics (Heisenberg Picture)

Non-relativistic Quantum Fields	Relativistic QM	Quantum Field Theory
	Same form as top 12 blocks of Wholeness Chart 2-2	Same form as top 12 blocks of Wholeness Chart 2-2
No theory generally used.	Same form as Non-relativistic Quantum Mechanics section, but different meaning for p_i	for $u = u(\phi^r, \pi_r, t)$, $v = v(\phi^r, \pi_r, t)$ $[u, v] = [uv - vu]$
	See Non-relativistic Quantum Mechanics section	i) for $v = H$ $\dot{u} = \frac{du}{dt} = \frac{-i}{\hbar}[u, H] + \frac{\partial u}{\partial t}$ ii) for i) plus $u = \phi^r$ or π_r $\dot{\pi}_r = \frac{-i}{\hbar}[\pi_r, H] = \frac{\delta H}{\delta \phi^r}; \quad \dot{\phi}^r = \frac{-i}{\hbar}[\phi^r, H] = \frac{\delta H}{\delta \pi_r}$
	See Non-relativistic Quantum Mechanics section	$[\phi^r, \pi_s] = i\hbar \delta^r_s \delta(\mathbf{x} - \mathbf{y}); \quad [\phi^r, \phi^s] = [\pi_r, \pi_s] = 0$

2.8 Chapter Summary

The bottom right hand block of Wholeness Chart 2-5, Summary of Quantum Mechanics, contains the essence of this chapter. A quantum field and its own conjugate momentum density do not commute, whereas all other pairings of fields and momentum density do commute. This is one postulate at the basis of QFT (see (2-44).) The other postulate comprises keeping the same form for the Lagrangian density (or equivalently, either the Hamiltonian density or the field equations of motion) as in the classical realm. These postulates are known as second quantization. (I guess we've said this enough. ☺)

Natural units and their relation to other types of units, summarized in Wholeness Chart 2-1 and Sect. 2.1.7, comprise another key concept in the chapter. In natural units, $c = \hbar = 1$ (dimensionless), and all quantities are expressed in units of powers of MeV.

Other fundamental concepts include certain field relations in the right most column of Wholeness Chart 2-2, which apply in the quantum realm. These are i) the Euler-Lagrange equation for fields, ii) the definition of conjugate momentum density, and iii) the Legendre transformation for fields. (Note that we will do virtually nothing with Hamilton's equations, so you need not worry about them.)

Unitary transformations, designated often by U , are quite important in QFT and are summarized in Box 2-3. When acting on a state vector, unitary transformations do not change the length in complex space of the state, the square of which is probability density. Thus, unitary transformations conserve probability. Importantly, $U^{-1} = U^\dagger$.

Quantum theories can be expressed in two different pictures, called the Schrödinger and Heisenberg pictures, summarized in Wholeness Chart 2-4. In the S.P., states are time dependent, but operators usually are not. The H.P. is the opposite. For it, states are static (fixed in time) and operators often time dependent. The key equation of motion in the S.P. is the state equation of motion (the Schrödinger equation). The key equation of motion in the H.P. is the operator equation of motion. (There is, since the state is constant, effectively, no H. P. state equation of motion.) The H.P. is closer to the classical perspective in that the focus in both is on dynamic variables/operators such as H, p_i , etc., which may vary in time. (And there is no state equation of motion in the classical world, since, for it, there is no such thing as a state.) QFT is easier to develop in the H.P., so we will be using it, rather than the S.P.

2.9 Problems

1. Pretend you are scientist in the pre MKS system days, with knowledge of Newton's laws. Units of meters for length, kilograms for mass, and seconds for time have been proposed. What units would force be measured in? Would it be appropriate to give the units for force the shortcut name "newton"? Could you have, alternatively, chosen units for other quantities than length, mass, and seconds as fundamental, and derived units for the remaining quantities? Could you have chosen the speed of sound as one of your basic units and selected it as equal to one and dimensionless? If so, and time in seconds was another basic unit, what units would length have?
2. The fine structure constant α in the Gaussian system (cgs with electromagnetism) is $e^2/4\pi\hbar c$, dimensionless, and approximately equal to 1/137. Without doing any calculations and without looking at Wholeness Chart 2-1, what are its algebraic expression, its dimensions, and its numerical value in natural units? Why can you find the dimensions and numerical value so easily? Does charge have dimensions in natural units? Without looking up the electron charge in Gaussian units, calculate the charge on the electron in natural units. (Answer: .303.)
3. Suppose we have a term in the Lagrangian density of form $m^2 \phi^2$, where m has dimensions of mass. What is the dimension M , in natural units, of the field ϕ ?
4. a) Derive $x^\alpha = g^{\alpha\beta} x_\beta$. [Hint: Use (2-5) and (2-6), or alternatively, use the matrix form of the contravariant metric tensor along with column vectors in terms of Cartesian coordinates] Note that this relation and (2-5) hold in general for any 4D vector, not just the position vector.
 b) Express $\partial^\mu \partial_\mu$ in terms of i) contravariant and covariant 4D components, and ii) in terms of time t and Cartesian coordinates X_i .
 c) Then find $\partial^\mu \partial_\mu (x^\alpha x_\alpha)$, where physical length of the interval of x^α is $\sqrt{x^\alpha x_\alpha}$, i) by expressing all terms in t and X_i , and ii) solely using 4D component notation. (For the last part, note, from a), that $\partial x^\alpha / \partial x_\beta = g^{\alpha\beta}$ and from (2-5), $\partial x_\alpha / \partial x^\beta = g_{\alpha\beta}$.)
5. Obtain your answers to the following questions by inspection of the final equation in Box 2-2, and then ask yourself whether or not your conclusions feel right intuitively.
 If ϕ^r were a sinusoid, how would the physical momentum density of a short wavelength wave compare to that of a longer one?
 If conjugate momentum density, in some way, measures the rate at which a field changes its values, how does physical momentum density depend on the speed at which a field changes?
6. Consider a classical, non-relativistic field of dust particles in outer space that are so diluted they do not exert any measurable pressure on one another. There is no gravitational, or other, potential density, i.e., $\mathcal{V}(x^i) = 0$. The density of particles is $\rho(x^i)$, which for our purposes we can consider constant in time. The displacement of the field (movement of each dust particle at each point) from its initial position is designated by the field value $\phi^r(x^i)$. $r = 1, 2, 3$, here, as there is a component of displacement, measured in length units, in each of the three spatial directions. ϕ^r and x^i are both measures of length, but the x^i are fixed locations in space, whereas the ϕ^r are displacements of the particles, in three spatial directions, relative to their initial positions.
 What is the kinetic energy density in terms of the field displacement ϕ^r (actually, it is in terms of the time derivatives of ϕ^r and $\dot{\phi}_r$)? What is the Lagrangian density for the field? Use (2-13) to find the differential equation of motion for the displacement ϕ^r . You should get $\rho\ddot{\phi}_r = 0$. Is this just Newton's second law for a continuous medium with no internal or external force?
7. Without looking back in the chapter, write down the Euler-Lagrange equation for fields. This is a good thing to memorize.
8. Why are the Hamiltonian and the Hamiltonian density not Lorentz scalars? If they are to represent energy and energy density, respectively, does this make sense? (Does the energy of an object or a system have the same value for all observers? Do you measure the same kinetic

energy for a plane passing overhead as someone on board the plane would?) Energy is the zeroth component of the four momentum p_μ . Does one component of a four vector have the same value for everyone?

9. (Do this problem only if you have extra time and want to understand relativity better.) Construct a column like those shown in Wholeness Chart 2-2 for the Relativistic Particle case, but do the entire summary in terms of relativistically covariant relationships. (That is, start with world (proper) time τ and fill in the boxes using 4D momentum, etc.)
10. Consider the unitary operator $U = e^{-iHt}$, where H is the Hamiltonian, and a non-energy eigenstate ket, $|\psi\rangle = C_1 |\psi_{E_1}\rangle + C_2 |\psi_{E_2}\rangle$. What is $U|\psi\rangle$?
11. Consider the unitary operator $U = e^{-iH(t-t_0)}$ and $|\psi_E\rangle = |e^{-i(Et_0 - \mathbf{p} \cdot \mathbf{x})}\rangle$, an energy eigenstate at time t_0 . What is $U|\psi_E\rangle$? Does U here act as a translator of the state in time? That is, does it have the effect of moving the state that was fixed in time forward in time, and turning it into a dynamic entity rather than a static one? If we operate on this new dynamic state with U^\dagger , would we turn it back into a static state? Is that not what we do when we operate on a Schrödinger picture state to turn it into a (static) Heisenberg picture state? (Earlier in the chapter we took $t_0 = 0$ to make things simpler.)

*Robert D. Klauber
www.quantumfieldtheory.info*

Path Integrals in Quantum Theories: A Pedagogic First Step

As I have mentioned elsewhere ([Quantum Field Theory: A Pedagogic Intro](#))¹, I strongly believe it far easier, and more meaningful, for students to learn QFT first by the canonical quantization method, and once that has been digested, move on to the path integral (many paths) approach. Hopefully, the material below will help such students, as well as those who are forced to begin their study of QFT via path integrals.

1 Background Math: Examples and Definitions

1.1 Functionals

Functionals form the mathematical roots of Feynman's many paths approach to quantum theories. A functional is simply a function of another function.

Example: Kinetic energy $T = \frac{1}{2}mv^2$ where $v = v(t)$. T is a function of v , and v is a function of t .

Definition(s): A functional is 1) a function of a function, OR equivalently, 2) a function of a dependent variable, OR equivalently, 3) a mapping of a function to a number.

Symbolism:

$$F[x(t)] \text{ or } F[x(t), \dot{x}(t)] \quad (1)$$

The square bracket notation is common, but not always used. Mathematically, x and t represent any function and its independent variable, though in physical problems, they are typically spatial position and time, respectively. Functionals are often dependent on the derivative(s) of a function, as well as the function itself, as in the RHS of (1). Total energy, with potential energy dependent on $x(t)$ and kinetic energy dependent on $\dot{x}(t)$, is one example. Additionally, a functional could also be a direct (rather than indirect as in (1)) function of t , i.e.

$$F[x(t), \dot{x}(t), t]. \quad (2)$$

1.2 Functional Derivative

Definition: A functional derivative is simply the derivative of a functional (F above) with respect to a function upon which it depends (x above).

Symbolism:

$$\frac{\delta F}{\delta x(t)} \text{ or } \frac{\delta F}{\delta x} \quad (3)$$

The δ notation is common, though the partial derivative symbol ∂ is often used instead.

1.3 Functional Integral

Definition: A functional integral is the integral of a functional with respect to a function upon which it depends.

Symbolism:

$$\int_{x_a}^{x_b} F \delta x \text{ or } \int_{x_a}^{x_b} F \delta x(t) \quad (4)$$

In the literature, one may find use of the usual differential symbol d instead of δ .

2 Different Kinds of Integration with Functionals

The value of a functional F of a physical system, such as a particle, is dependent on where it is in space and time, i.e., $x(t)$ and t in (1) are then considered spatial position and time. Further, one can integrate a functional F in different ways over its path in space and time, or over projections of that path. Several of these are depicted in Table 1 below. The first three kinds of integration shown below are fairly self explanatory. We comment on the fourth after the table.

Table 1. Some Ways to Integrate Functionals

	Type of Integration	Graphically	Math	Comment & Use in Physics
1.	Area over the path in $x(t)$ vs. t space	$F[x(t)]$ path in $x-t$ space	$\int_{s_a}^{s_b} F ds$ where s is distance along path	No real physical application.
2.	Projection of the area in 1 onto the $F-t$ plane	$F[x(t)]$ projection onto $F-t$ plane	$\int_{t_a}^{t_b} F dt$	If $F=L$, the Lagrangian, then this integral = S , the action. Classically, $S =$ minimum (or stationary) for physical paths
3.	Projection of the area in 1 onto the $F-x(t)$ plane	$F[x(t)]$ projection onto $F-x$ plane	$\int_{x_a}^{x_b} F \delta x(t)$	This is the usual definition of "functional integral" This is starting point for 4, below
4.	Simultaneous integration over all possible paths in 3	$F[x(t)]$ 4 of an infinite number of paths	$\int_{x_a}^{x_b} F \mathcal{D}x(t)$	QM & QFT Feynman path integral approach. \mathcal{D} symbol implies a sum of the integrals of all paths in 3, not just the classical path

The fourth way to integrate above is not simple, nor is its purpose at all obvious at this point. We devote entire sections below to explaining its origin, its value, and means to evaluate it. So, for now, just let it float easily through your head and don't bother straining to understand it.

Alternative nomenclature: Because functional integrals are integrated over particular paths (in x - t space in above examples), they are often also referred to as path integrals.

3 The Transition Amplitude

3.1 General Wave Functions (States)

Recall from QM wave mechanics, that for a general normalized wave function ψ equal to a superposition of energy eigenfunction waves (which are each also normalized),

$$\psi = A_1\psi_1 + A_2\psi_2 + A_3\psi_3, \quad (5)$$

A_1 is the amplitude of ψ_1 , so the probability of finding ψ_1 upon measuring is

$$A_1^* A_1 = |A_1|^2. \quad (6)$$

If we were to start with ψ initially, and measure ψ_1 later, the wave function would have collapsed, i.e., underwent a transition to a new state. (6) would be the transition probability.

Definition: The transition amplitude is that complex number, the square of the absolute magnitude of which is the probability of measuring a transition from a given initial state to a specific final state.

Symbolism: The transition amplitude is often written as

$$U(\psi_i, \psi_f; T), \quad (7)$$

implying an initial state ψ_i , a final state ψ_f , and an elapsed time between measurements of the two of T .

This terminology carries over to QFT when particles change types. For example, the probability that an electron and a positron would annihilate to create two photons would be the square of the absolute value of the transition amplitude between the initial (e^- , e^+) and final (2γ) states. (Almost all of QFT is devoted ultimately to determining the transition amplitudes for the different possible interactions between particles.)

Schroedinger Approach Amplitudes

We can't get into explaining it here (for those who may not know it already), but the Schroedinger approach to QM leads to an expression of the transition amplitude of form

$$U(\psi_i, \psi_f; T) = \underbrace{\langle \psi_f |}_{\substack{\text{final state} \\ \text{measured} \\ \text{at } T+t_a}} e^{-iHT/\hbar} \underbrace{| \psi_i \rangle}_{\substack{\text{initial state} \\ \text{at } t_a}}, \quad (8)$$

evolved state
at $T+t_a$

where H is the Hamiltonian operator.

Alternative nomenclature: The transition amplitude U is sometimes called the propagator (though *not* the “Feynman propagator” of QED) because it is the contribution to the wave function at f at time T from that at i at time 0. It “propagates” the particle from i to f .

3.2 Position Eigenstates

When the particle has a definite position, e.g., x_i , the wave function is an eigenstate of position, and the ket is written $|x_i\rangle$. The transition amplitude for measuring a particle initially at x_i , and finally at x_f , would take the form

$$U(x_i, x_f; T) = \langle x_f | \underbrace{e^{-iHT/\hbar}}_{\text{evolved state } \psi} | x_i \rangle. \quad (9)$$

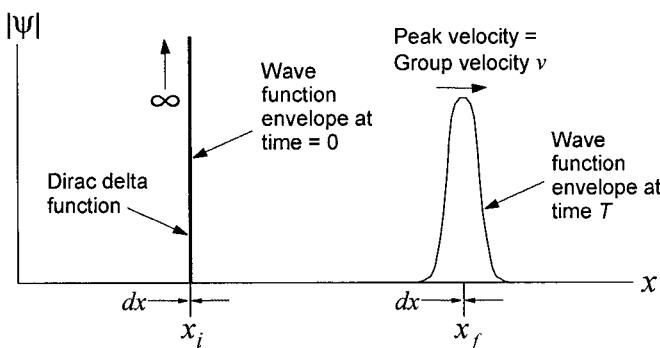


Figure 1. Propagation of a Position Eigenstate Quantum Wave

In wave mechanics notation, $|x_i\rangle$ and $|x_f\rangle$ are both delta functions of form $\delta(x - x_i)$ and $\delta(x - x_f)$, the first of which is represented schematically on the left in Figure 1. As the initial state evolves into ψ , however, it, like wave packets generally do, spreads, and its peak diminishes (see wave function envelope on right side of Figure 1.) The amplitude for measuring the particle at time T at x_f , i.e., for measuring a delta function $|x_f\rangle$ that collapsed from ψ , is (9).

We can re-write (9), in wave mechanics notation as

$$U(x_i, x_f; T) = \int_{-\infty}^{+\infty} \delta(x - x_f) \psi(x, T) dx = \psi(x_f, T). \quad (10)$$

Thus,

$$|U(x_i, x_f; T)|^2 = \psi^*(x_f, T) \psi(x_f, T) = \text{probability density at } x_f. \quad (11)$$

Modification to definition: Hence, the square of the absolute value of the transition amplitude for eigenstates of position is *probability density*, *not probability*, as was the case for energy eigenstate wave functions of form (5).

As we will see, the value found using the RHS of (9), i.e., that of the Schroedinger approach, is the same as the value found using Feynman’s many paths approach.

4 Expressing the Wave Function Peak in Terms of the Lagrangian

4.1 Background

One of Feynman’s assumptions for his path integral approach to QM and QFT was to express the wave function value at the peak of a wave packet (see Figure 1) in terms of the Lagrangian (exact relation shown at the end of this section 4). I have never seen much

justification for this in the literature, other than it is simply an assumption that works (so learn to live with it and move on!)

In the present section I have taken a different tack, by providing rational for why we could expect Feynman's form of the wave function peak to work. The logic herein may well parallel what went on in Feynman's mind as he was developing his path integral approach.

4.2 Deducing Feynman's Phase Peak Relationship

4.2.1 The Simplified, Heuristic Argument

In QM, the plane wave function solution to the Schrödinger equation,

$$\psi = Ae^{-i(Et - \mathbf{p} \cdot \mathbf{x})/\hbar}, \quad (12)$$

means the phase angle, at any given \mathbf{x} and t , is

$$\phi = -(Et - \mathbf{p} \cdot \mathbf{x})/\hbar. \quad (13)$$

If we have a particle wave packet, it is an aggregate of many such waves, so it is not in an energy or momentum eigenstate. However, it does have energy and momentum expectation values that correspond to the classical values for the particle. The wave packet peak travels at the wave packet group velocity, which corresponds to the classical particle velocity.

Now, imagine we approximate the wave packet with a (spatially short) wave function such as ψ , where E and \mathbf{p} take on the values of the wave packet expectation values for energy and momentum, respectively. If \mathbf{x} represents the position of the wave packet peak (the middle of our approximated wave function ψ), the time rate of change of phase at the peak is then

$$\frac{d\phi}{dt} = \frac{-(E - \mathbf{p} \cdot \mathbf{v})}{\hbar} = \frac{-T - V + \mathbf{p} \cdot \mathbf{v}}{\hbar}, \quad (14)$$

where \mathbf{v} is the velocity of the wave peak, T is kinetic energy, and V is potential energy. Non-relativistically,

$$T = \frac{1}{2}mv^2 \quad \mathbf{p} = m\mathbf{v}, \quad (15)$$

so, in terms of the classical Lagrangian L ,

$$\frac{d\phi}{dt} = \frac{T - V}{\hbar} = \frac{L}{\hbar}. \quad (16)$$

More formally, using the Legendre transformation

$$L = p_i \dot{q}_i - H \quad (L = \mathbf{p} \cdot \mathbf{v} - E \text{ here}), \quad (17)$$

directly in (14), we get (16).

Thus, from (16), the phase difference between two events the particle traverses is

$$\phi = \int \frac{L}{\hbar} dt = \frac{S}{\hbar}, \quad (18)$$

where S is the classical action of Hamilton. The classical path between two events is that for which the Hamiltonian action is least. Note that (18) is an integral of type 2 in Table 1.

Hence, the wave function at the peak could be written in terms of the Lagrangian as

$$\psi_{peak} = Ae^{i \int \frac{L}{\hbar} dt} = Ae^{\frac{iS}{\hbar}}. \quad (19)$$

This is the typical starting point assumption when teaching the Feynman path integral approach (still to be developed beginning in Section 5.)

In relativistic quantum mechanics (RQM) and quantum field theory (QFT), we get a solution form similar to (12) (differing only in the normalization factor A), and thus (14) is also true relativistically. Further, since (17) is true relativistically, as well, then so are (16), (18), and (19).

4.2.2 More Precise Argument

The precise expression for a QM particle wave packet², where overbars designate expectation (classical) values; v_g , the group (peak, classical) velocity; and $g(p)$, the momentum space distribution is

$$\psi(x, t) = e^{-\frac{i}{\hbar}(\bar{E}t - \bar{p}x)} \underbrace{\frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} g(p) e^{-\frac{i}{\hbar}(v_g t - x)(p - \bar{p})} e^{-\frac{i}{2\hbar m}(p - \bar{p})^2} dp}_{A(t) \text{ for } x=x_{peak}}. \quad (20)$$

real
=1 for $x=x_{peak}$,
i.e., for $x=v_g t$
time depend & complex

We are interested in the value of (20) at the peak, $\psi(x_{peak}, t)$, where $x_{peak} = v_g t$. To begin, note that with $x=x_{peak}$ inside the integral, the exponent of the second factor in the integrand equals zero, and so that factor equals one. The function $g(p)$ is typically a real, Gaussian distribution in $p - \bar{p}$, and independent of time. The third factor in the integrand is complex and time dependent.

Thus, with $x=x_{peak}$, the integral in (20) is a function (generally complex) only of time, which, along with the factor in front, we will designate as $A(t)$. Thus, for the entire history of the wave packet, the wave function value at the peak is

$$\psi(x_{peak}, t) = A(t) e^{-\frac{i}{\hbar}(\bar{E}t - \bar{p}x_{peak})}. \quad (21)$$

Except for the time dependence in $A(t)$, this is equivalent to (12), as the expectation values for E and p equal the classical values for the particle. So, with regard to the exponent factor in (21), all of the logic from (13) through (19) applies here as well. The final result is so important, we repeat it below, with L being the classical particle Lagrangian, T representing the time when the peak is detected, and phase at $t=0$ taken as zero. The RHS comes from (10).

$$\psi(x_{peak}, T) = A(T) e^{i \int_0^T \frac{L}{\hbar} dt} = A(T) e^{\frac{iS}{\hbar}} = U(x_i, x_f, T)$$

(22)

We evaluate $A(t)$ exactly in the Appendix.

Definition: Borrowing a term from electrical engineering, we will herein refer to $e^{i\phi}$ as a phasor.

5 Feynman's Path Integral Approach: The Central Idea

Feynman's remarkable idea takes a little getting used to. He reasoned that a particle/wave (such as an electron) traveling a path (world line in spacetime) between two events could actually be considered to be traveling along all possible paths (infinite in number) between those events.

Difficult as it may be, initially, to believe, we will see below that the result from superimposition of the phasors from all of these paths gives us the same result as if we used the standard QM theory of Schroedinger with a single wave. The two different approaches are equivalent.

Definition: Feynman's method is called the “path integral”, “many paths”, or “sum over histories” approach to QM (and QFT).

Note that the paths do not have to satisfy physical laws like conservation of energy, $\mathbf{F}=\mathbf{ma}$, least action, etc. Moreover, each possible path is considered equally probable.

We will lead into the formal mathematics of the many paths approach by first examining simple situations with a finite number of paths between two events.

6 Superimposing a Finite Number of Paths³

6.1 The Rotating Phasor

The phasor of (22) can be expressed in the complex plane as a unit length vector with angle ϕ relative to the real positive (horizontal) axis. As time evolves this vector rotates at the rate L/\hbar ,

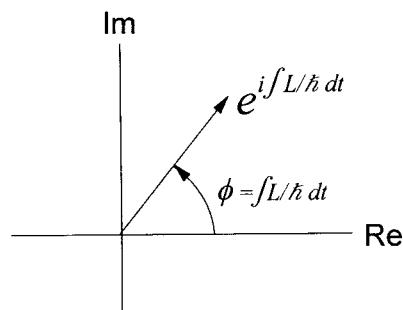


Figure 2. Rotating Phasor

i.e., the total phase $\phi = \int \frac{L}{\hbar} dt$. So we can picture the phasor as a unit length vector rotating like a hand on a clock in a 2D complex plane (though it is a counterclockwise rotation).

For the purposes of Feynman's approach, we can consider the particle as a wave packet with phase at the peak determined by (22), and our final measurement a position eigenstate measured at the packet peak. We then imagine a different wave packet following each one of the infinitely many paths between two specific events. We visualize the

phasor at the peak for each of these paths as a vector rotating in the complex plane as time passes (i.e., as the wave packet peak moves along the path), eventually having a particular value at the final event, the arrival place and time. Each path will have a different final phase.

6.2 Several Paths Graphically

Fig. 24 in Feynman's book *QED: The Strange Theory of Light and Matter*⁴, is an insightful, somewhat heuristic, illustration of the many paths concept for light. Since we wish to focus, for the time, on non-relativistic quanta, we employ a similar, and at least equally heuristic, illustration in Fig. 3 for an electron rather than a photon. In Fig. 3 an electron is emitted at event a, reflected, like light from a mirror, off of a scattering surface, and detected at point b. The scattering surface might be difficult to construct in practice, but one can imagine a surface densely packed with tightly bound negative charge.

We look at a representative 15 different paths for the electron, out the infinite number in the many paths approach, and label them with letters A to O. Each path takes the same time T . Note that path H is the classical path, having equal angles of incidence and reflection. Since it is the shortest, particle speed for that path is lowest.

The Lagrangian here is simply the kinetic energy, and this is constant, though different, for each path. Since speed is least for the classical path H, it has the smallest Lagrangian, and thus

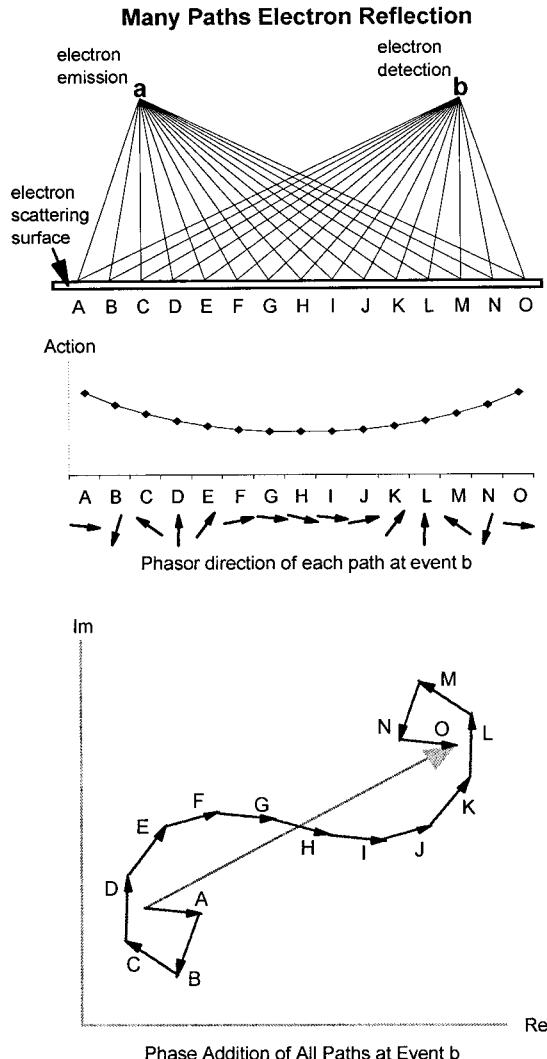


Figure 3. Graphical Justification for Many Paths Approach

greater contributions to the final sum. A similar effect would occur if the value for Planck's constant were smaller. As $\hbar \rightarrow 0$, all paths but H would tend to cancel out.

Footnote: I used to think that increasing mass, and thus getting closer to the classical situation, would bring the phase angle of the sum-of-all-paths phasor in directional alignment with H, the classical path phasor (or at least with U of (22).) However, this is not the case. The important thing in Feynman's approach is not the phase of the sum-of-all-paths phasor, but its *length*, which is proportional to $|U|$. And this length gets greater contribution from paths nearer H than from those further away.

Note that in order to get a graphically significant Fig. 3, I had to use a value for \hbar almost eight orders of magnitude greater than the actual value. Otherwise the phase angles between adjacent paths, for the relatively large spacing between paths of the figure, would have resulted in a seemingly random hodgepodge of phasors, and obscured, rather than illuminated, the real physics involved.

If you would like to experiment with changing values for mass, \hbar , and number of paths yourself, download the Excel spreadsheet [Many Paths Graphic Electron Reflection⁵](#). End footnote.

the least action. The other paths do not obey the usual classical laws, such as least action, equal angles of incidence and reflection, etc. But according to Feynman's approach, we have to include all of them.

From (22) and Fig. 2, we can determine the phasor $e^{iS/\hbar}$ of (22) for the particle/wave arriving at event b, for each path, where $S = LT = \frac{1}{2} mv^2 T$. The phasor direction in complex space for each path at the detection event b is depicted in the middle of Fig. 3.

The bottom part of Fig. 3 shows the addition of the final event phasors for all 15 paths. Note that the paths further from the classical path H tend to cancel each other out, because they are out of phase. Conversely, H and the paths close to H are close to being in phase, and thus, reinforce each other via constructive interference. So, the primary contributions to the phasor sum are from those paths close to the classical path.

If we were to increase the number of paths, the jaggedness of the curve formed by the 15 phasors would smooth out, but its basic overall shape would remain essentially the same. If we were to increase the Lagrangian, while keeping speed the same for each path (i.e., increasing mass of the particle), phasors now near the middle of the curve would shift towards the ends, and thus, be cancelled out via interference. In other words, increasing mass brings us closer to the classical case, and the paths closer to classical then make

Feynman intuited that the amplitude of the final phasor sum was extremely meaningful. That is, the square of its absolute value (i.e., the square of its length in complex space) was proportional (approximately, for a finite number of paths; exactly, for an infinite number) to the probability density for measuring the photon/particle at event b. What we mean by “proportional” should become clearer after the following three sections.

6.3 Many Paths Mathematically

Consider particle paths similar to those of Fig. 3, where the wave function peak for path number 1, with $A_1(T)$ as in (22), as

$$\psi_1^{\text{peak}} = A_1(T)e^{iS_1/\hbar}. \quad (23)$$

In the spirit of the prior section, one considers the phasor of (23) *without* $A_1(t)$ as representing the particle, AND that particle is considered to simultaneously travel many paths between events a and b. Then, the summation of the final phasors for each path is expressed mathematically as

$$e^{iS_1/\hbar} + e^{iS_2/\hbar} + e^{iS_3/\hbar} + \dots = A_b e^{i\phi_{\text{sum}}} \quad (24)$$

where A_b is the amplitude of the sum. As the number of paths approaches infinity, $|A_b|^2$ becomes proportional to the probability density of measuring the particular final state at event b. That is,

$$\lim_{N \rightarrow \infty} \sum_{j=1}^N e^{iS_j/\hbar} = A_b e^{i\phi_{\text{sum}}} \propto U(x_i, x_f, T) \quad |A_b e^{i\phi_{\text{sum}}}|^2 = |A_b|^2 \propto |U|^2 \text{ (probability density).} \quad (25)$$

We will learn how to evaluate the limit in (25).

6.4 Another Example

Consider a double slit experiment with a classical Huygen's wave analysis showing alternating fringes of light and dark, which via the classical interpretation is caused by constructive and destructive interference of light/electron waves.

By the Schroedinger wave approach, a single quantum wave travels through both slits, interferes with itself, either constructively or destructively, to result in a wave amplitude that varies with location along the receiving screen. The probability density (square of the amplitude absolute value) of finding a photon/electron also varies with that screen location. So as the quantum waves collapse, one at a time, on the screen, they tend to collapse more often in the

high probability (high magnitude amplitude) regions. These correspond to the bright fringe regions, which, with enough individual quanta collapsing on the screen, are seen by the human eye.

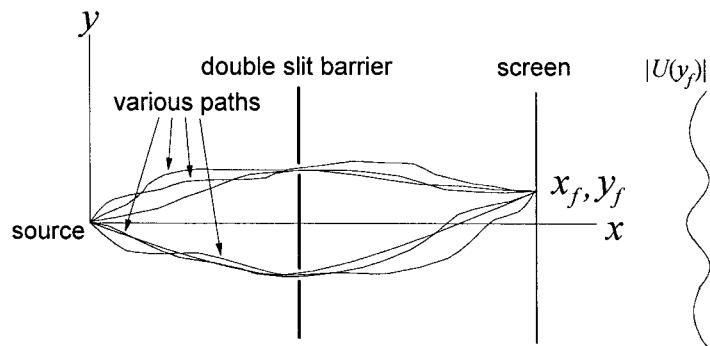


Figure 4. Double Slit Experiment in Many Paths Approach

In the many paths approach, for any particular spot on the screen, we would add the phases of every “possible” path from the emission point, through one slit, to that spot (x_f, y_f) , plus all paths through the other slit to the spot. See Figure 4.

The result would be proportional to the amplitude at the spot found in the Schroedinger approach. That is, the sum of all phasors at x_f, y_f (see (25)) yields

$$C \lim_{N \rightarrow \infty} \sum_{j=1}^N e^{iS_j/\hbar} = U(x_i, y_i; x_f, y_f; T), \quad (26)$$

where C is some constant.

We would then repeat that procedure for every other point on the screen. Since, for a fixed source at x_i, y_i , and a fixed x_f for the screen, the amplitude would be spatially only a function of y_f , and we could express it simply as $U(y_f)$.

6.5 Finding the Proportionality Constant: By Example

The square of the absolute value of the amplitude U is the probability density. So we can normalize U over the length of the screen, i.e.,

$$\int_{y_f=-\infty}^{y_f=+\infty} \left| C \lim_{N \rightarrow \infty} \sum_{j=1}^N e^{iS_j/\hbar} \right|^2 dy_f = \int_{y_f=-\infty}^{y_f=+\infty} |U(y_f)|^2 dy_f = 1, \quad (27)$$

and thus, once the value of the limit is determined, readily find the proportionality constant C .

7 Summary of Approaches

7.1 Feynman's Postulates

Richard Feynman was probably well aware of much of the foregoing when he speculated on the viability of the following three postulates for his many paths approach. Subsequent extensive analysis by Feynman and many others has validated his initial speculation.

The postulates of the many paths approach to quantum theories are:

1. The phasor value at any final event is equal to $e^{iS/\hbar}$ where the action S is calculated along a particular path beginning with a particular initial event.
2. The probability density for the final event is given by the square of the magnitude of a typically complex amplitude.
3. That amplitude is found by adding together the phasor values at that final event from all paths between the initial and final events, including classically impossible paths. The amplitude of the resultant summation must then be normalized relative to all other possible final events, and it is this normalized form of the amplitude that is referred to in 2.

Note two things.

First, there is no weighting of the various path phasors. The nearly classical paths are not weighted more heavily than the paths that are far from classical. That is, the different individual paths in the summation do not have different amplitudes (see (24) and Fig. 3). The correlation with the classical result comes from destructive interference among the paths far from classical, and constructive interference among the paths close to classical.

Second, time on all paths (all histories) must move forward. This is implicit in the exponent phase value of (19), where the integral of L is over time, with time moving forward. Our paths do not include particles zig-zagging backward and forward through time.

Footnote: Caveat: A famous quote by Freeman Dyson states that Feynman, while speculating on this approach, told him that one particle travels all paths, including those going backward in time. But the usual development of the theory (see Section 8) only includes paths forward in time. Perhaps all paths backward in time sum to zero and so are simply ignored. In such case, Dyson's quote would be accurate. But I don't know for sure. End footnote.

7.2 Comparison of Approaches to QM

Unifying Chart 1 summarizes the major similarities and differences between alternative approaches to QM.

Unifying Chart 1. Equivalent Approaches to Non-relativistic Quantum Mechanics

	<u>Schroedinger Wave Mechanics</u>	<u>Heisenberg Matrix Mechanics</u>	<u>Feynman Many Paths</u>
Probability Density of Position Eigenstates	$ \text{amplitude} ^2$		$ \text{amplitude} ^2$
Transition Amplitude	$U(x_i, x_f; T) = \langle x_f e^{-iHT/\hbar} x_i \rangle$	Same results as other two approaches.	$U(x_i, x_f; T) \propto \lim_{N \rightarrow \infty} \sum_{j=1}^N e^{iS_j/\hbar}$ $= \int_{x_i}^{x_f} e^{i \int_0^T \frac{L}{\hbar} dt} \mathcal{D}x(t)$
Comments	Above assumes normalized states.		RHS above must be normalized for $\alpha \rightarrow =$. We haven't done the integral part yet.

8 Finite Sums to Functional Integrals

8.1 Time Slicing: The Concept

After all of the foregoing groundwork, it is time to extend the phasor sum of a finite number of paths, such as we saw in Fig. 3 and (24), over into an infinite sum, or in other words, an integral. To do this, we first consider finite “slices” of time, for a finite number of paths in one spatial dimension, as shown in Fig. 5 where, for convenience, we plot time vertically and space horizontally. As opposed to our spatially 2D example in Fig. 3, different paths in Fig. 5 actually refer to the particle traveling the same direction x between i and f , though at varying velocities. The paths between each slice are straight lines, but there is no loss in generality, as one can take the time between slices $\Delta t \rightarrow dt$, and thus, any possible shape path can be included.

As noted earlier, for any single path, the

$$\text{phasor at } \mathbf{f} = \underbrace{e^{i \int_{t_i}^{t_f} \frac{L}{\hbar} dt}}_{\text{one path}} = e^{i S / \hbar}, \quad (28)$$

The amplitude U for the transition from \mathbf{i} to \mathbf{f} is proportional to the sum of (28) for all paths,

$$\text{sum of } \infty \text{ phasors at } \mathbf{f} = \lim_{N \rightarrow \infty} \sum_{j=1}^N e^{i S_j / \hbar}. \quad (29)$$

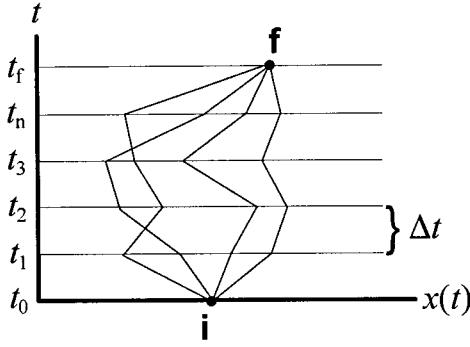


Figure 5. Time Slicing for Finite Number of Paths

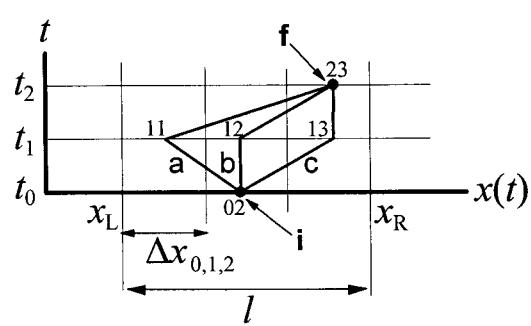


Figure 6. Space Slicing for Three Discrete Paths

8.2 Space Slicing: Simple Paths with Discrete Approximation

To evaluate (29), we next also discretize (“slice”) space, and consider a small number (three) of paths over a small number of discrete events in spacetime, as in Figure 6. We label the paths a , b , and c , and the events with two numbers, such that the first number represents the time slice, and the second the space slice. The continuous range of x values at time t_1 will be designated x_1 ; at t_2 , x_2 ; etc. We limit the spatial range for paths considered to $x_R - x_L = l$, where the number of paths $N = 3 = l / \Delta x_1$. Each path passes through the center of one Δx_1 segment.

We then assume the phase ϕ_{02} at \mathbf{i} is zero, and find the phasors at \mathbf{f} for each of the three paths by subsequently adding the phase difference between discrete events along a given path, as in the second column of Table 2 below. There, as elsewhere herein, clicking on an equation will enlarge it for easier viewing. Clicking again will contract it to its original size.

Note that in the last line of column two in Table 2, the Lagrangian L without subscript is assumed to be the L for the particular subpath being integrated, and this is common notation.

In column three, we approximate the integrals of L over t , such that, for example, for path a over an interval Δt ,

$$S_a \approx L_a^{\text{apprx}} \Delta t \quad (30)$$

where, for the first subpath,

$$L_a = \frac{1}{2} m \dot{x}^2 - V(x) \approx \frac{1}{2} m \left(\frac{x_{11} - x_{02}}{\Delta t} \right)^2 - V \left(\frac{x_{11} + x_{02}}{2} \right) \approx L_a^{\text{apprx}}(x_{11}, x_{02}) \quad (31)$$

Similar relations hold for the other subpaths, and are shown in Table 2.

Note that (31) is solely a function of x_{11} and x_{02} . The summation of all three paths in the last row of column three in Table 2 is solely a function of x_{02}, x_{23} , and the three intermediate event x values x_{11}, x_{12} , and x_{13} . Since x_{02} and x_{23} are the initial and final events, which are fixed and the same for all paths, the final summation approximation in Table 2 are really only functions of the three x_{1j} . It will, however, serve a future purpose if we keep x_{02} and x_{23} in the relationship for the time being.

Table 2. Adding Phasors at the Final Event for Three Discrete Paths

Path	Phasor at \mathbf{f}	Phasor at \mathbf{f} in Terms of Approx L
a	$e^{i_a \phi_{23}} = e^{i(\phi_{02 \rightarrow 11} + \phi_{11 \rightarrow 23})} = e^{i \int_{02}^{11} \frac{L_a}{\hbar} dt} e^{i \int_{11}^{23} \frac{L_a}{\hbar} dt}$	$\approx e^{\frac{i}{\hbar} \left\{ \frac{1}{2} m \left(\frac{x_{11} - x_{02}}{\Delta t} \right)^2 - V \left(\frac{x_{11} + x_{02}}{2} \right) \right\} \Delta t} e^{\frac{i}{\hbar} \left\{ \frac{1}{2} m \left(\frac{x_{23} - x_{11}}{\Delta t} \right)^2 - V \left(\frac{x_{23} + x_{11}}{2} \right) \right\} \Delta t}$ $= e^{\frac{i}{\hbar} f(x_{02}, x_{11})} e^{\frac{i}{\hbar} f(x_{11}, x_{23})}$
b	$e^{i_b \phi_{23}} = e^{i(\phi_{02 \rightarrow 12} + \phi_{12 \rightarrow 23})} = e^{i \int_{02}^{12} \frac{L_b}{\hbar} dt} e^{i \int_{12}^{23} \frac{L_b}{\hbar} dt}$	$\approx e^{\frac{i}{\hbar} \left\{ \frac{1}{2} m \left(\frac{x_{12} - x_{02}}{\Delta t} \right)^2 - V \left(\frac{x_{12} + x_{02}}{2} \right) \right\} \Delta t} e^{\frac{i}{\hbar} \left\{ \frac{1}{2} m \left(\frac{x_{23} - x_{12}}{\Delta t} \right)^2 - V \left(\frac{x_{23} + x_{12}}{2} \right) \right\} \Delta t}$ $= e^{\frac{i}{\hbar} f(x_{02}, x_{12})} e^{\frac{i}{\hbar} f(x_{12}, x_{23})}$
c	$e^{i_c \phi_{23}} = e^{i(\phi_{02 \rightarrow 13} + \phi_{13 \rightarrow 23})} = e^{i \int_{02}^{13} \frac{L_c}{\hbar} dt} e^{i \int_{13}^{23} \frac{L_c}{\hbar} dt}$	$\approx e^{\frac{i}{\hbar} \left\{ \frac{1}{2} m \left(\frac{x_{13} - x_{02}}{\Delta t} \right)^2 - V \left(\frac{x_{13} + x_{02}}{2} \right) \right\} \Delta t} e^{\frac{i}{\hbar} \left\{ \frac{1}{2} m \left(\frac{x_{23} - x_{13}}{\Delta t} \right)^2 - V \left(\frac{x_{23} + x_{13}}{2} \right) \right\} \Delta t}$ $= e^{\frac{i}{\hbar} f(x_{02}, x_{13})} e^{\frac{i}{\hbar} f(x_{13}, x_{23})}$
Sum of a, b, c	$= e^{i \int_{02}^{11} \frac{L_a}{\hbar} dt} e^{i \int_{11}^{23} \frac{L_a}{\hbar} dt} + e^{i \int_{02}^{12} \frac{L_b}{\hbar} dt} e^{i \int_{12}^{23} \frac{L_b}{\hbar} dt}$ $+ e^{i \int_{02}^{13} \frac{L_c}{\hbar} dt} e^{i \int_{13}^{23} \frac{L_c}{\hbar} dt}$ $= \sum_{j=1}^{N=3} e^{i \int_{02}^{1j} \frac{L}{\hbar} dt} e^{i \int_{1j}^{23} \frac{L}{\hbar} dt}$	$= e^{\frac{i}{\hbar} f(x_{02}, x_{11})} e^{\frac{i}{\hbar} f(x_{11}, x_{23})} + e^{\frac{i}{\hbar} f(x_{02}, x_{12})} e^{\frac{i}{\hbar} f(x_{12}, x_{23})}$ $+ e^{\frac{i}{\hbar} f(x_{02}, x_{13})} e^{\frac{i}{\hbar} f(x_{13}, x_{23})}$ $= \sum_{j=1}^{N=3} e^{\frac{i}{\hbar} f(x_{02}, x_{1j})} e^{\frac{i}{\hbar} f(x_{1j}, x_{23})}$

The final relationship in Table 2 is approximately proportional to the transition amplitude, i.e.,

$$U(i, f; T = t_f - t_i) \approx C \sum_{j=1}^{N=3} e^{\frac{i}{\hbar} f(x_{02}, x_{1j})} e^{\frac{i}{\hbar} f(x_{1j}, x_{23})} = C \sum_{j=1}^{N=3} e^{\frac{i}{\hbar} S^{appx}(x_{02}, x_{1j})} e^{\frac{i}{\hbar} S^{appx}(x_{1j}, x_{23})}, \quad (32)$$

where C is some constant, and what we designated as a function f in Table 2, in order to emphasize its independent variables, is actually an approximation to the action S .

Since U is *proportional* to the sum of the phasors, we can multiply the RHS of (32) by any constant we like and the proportionality still holds. To aid us in taking limits to get an integral, we multiply (32) by Δx_1 , and get

$$U(i, f; T) \approx C' \sum_{j=1}^{N=3} e^{\frac{i}{\hbar} S^{appx}(x_{02}, x_{1j})} e^{\frac{i}{\hbar} S^{appx}(x_{1j}, x_{23})} \Delta x_1, \quad (33)$$

where C' is a new constant. Taking the limit where $\Delta x_1 \rightarrow dx_1$ means taking the number of paths $N \rightarrow \infty$. And thus,

$$\begin{aligned} U(i, f; T) &\approx C' \lim_{N \rightarrow \infty} \sum_{j=1}^N e^{\frac{i}{\hbar} S^{appx}(x_{02}, x_{1j})} e^{\frac{i}{\hbar} S^{appx}(x_{1j}, x_{23})} \Delta x_1 \\ &= C' \int_{x_1=x_L}^{x_1=x_R} e^{\frac{i}{\hbar} S^{appx}(x_{02}, x_1)} e^{\frac{i}{\hbar} S^{appx}(x_1, x_{23})} dx_1 \approx C' \int_{x_1=x_L}^{x_1=x_R} e^{i \int_{02}^{23} \frac{L}{\hbar} dt} dx_1. \end{aligned} \quad (34)$$

where our discrete values x_{1j} have become a continuum x_1 , and it is implicit that the L of the last part of (34) is that over the appropriate path corresponding to each increment of dx_1 . (34) is still only approximately proportional to the amplitude because time is still discretized in Δt intervals and we limit the integration range to $x_L > x_1 > x_R$. Before extending those limits, however, we must consider a slightly more complicated set of paths.

8.3 From Simple Discrete Paths to the General Case

In Figure 7 we introduce one more time interval between the initial and final events, resulting in nine discrete paths.

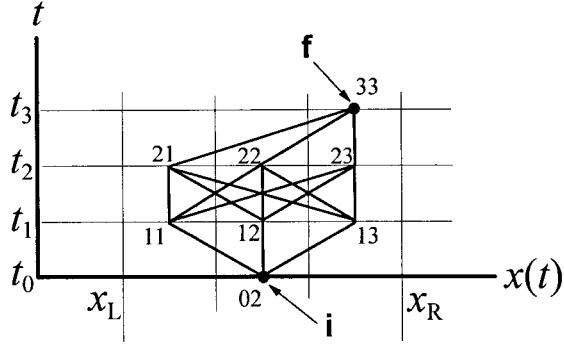


Figure 7. Nine Discrete Paths between Two Events

Repeating the logic from the previous section (use Table 2 as an aide), the phasor of the first path ($02 \rightarrow 11 \rightarrow 21 \rightarrow 33$) is simply

$$\begin{aligned} \underbrace{e^{i\phi_{33}}}_{\text{1st path only}} &= e^{i(\phi_{02 \rightarrow 11} + \phi_{11 \rightarrow 21} + \phi_{21 \rightarrow 33})} \\ &= e^{i \int_{02}^{11} \frac{L}{\hbar} dt} e^{i \int_{11}^{21} \frac{L}{\hbar} dt} e^{i \int_{21}^{33} \frac{L}{\hbar} dt}. \end{aligned} \quad (35)$$

We repeat this for the other eight paths, approximate L along subpaths as before, and take k below to indicate the k th Δx_2 segment. This results in a phasor summation from all paths at event f ($= 33$) [compare with last row, last column of Table 2 and (32)] proportional to the amplitude, i.e.,

$$U(i, f; T) \approx C \sum_{j=1}^{N=3} \sum_{k=1}^{N=3} e^{\frac{i}{\hbar} S^{apprx}(x_{02}, x_{1j})} e^{\frac{i}{\hbar} S^{apprx}(x_{1j}, x_{2k})} e^{\frac{i}{\hbar} S^{apprx}(x_{2k}, x_{33})}. \quad (36)$$

Note that (36) depends on the discrete values of both x_1 and x_2 . So, as we did with (33), we can multiply (36) by one or more constants without changing the proportionality. We choose to multiply by Δx_1 and Δx_2 . We follow by taking limits where $\Delta x_1 \rightarrow dx_1$ and $\Delta x_2 \rightarrow dx_2$ (i.e., $N \rightarrow \infty$), [compare with (34)] which results in

$$\begin{aligned} U(i, f; T) &\approx C' \lim_{N \rightarrow \infty} \sum_{j=1}^{N=3} \sum_{k=1}^{N=3} e^{\frac{i}{\hbar} S^{apprx}(x_{02}, x_{1j})} e^{\frac{i}{\hbar} S^{apprx}(x_{1j}, x_{2k})} e^{\frac{i}{\hbar} S^{apprx}(x_{2k}, x_{33})} \Delta x_1 \Delta x_2 \\ &= C' \int_{\substack{x_2=x_L \\ x_1=x_L}}^{\substack{x_2=x_R \\ x_1=x_R}} e^{\frac{i}{\hbar} S^{apprx}(x_{02}, x_{1j})} e^{\frac{i}{\hbar} S^{apprx}(x_{1j}, x_{2k})} e^{\frac{i}{\hbar} S^{apprx}(x_{2k}, x_{33})} dx_1 dx_2 \\ &\approx C' \int_{\substack{x_2=x_L \\ x_1=x_L}}^{\substack{x_2=x_R \\ x_1=x_R}} e^{i \int_{t_{02}}^{t_{33}} \frac{L}{\hbar} dt} dx_1 dx_2. \end{aligned} \quad (37)$$

We can readily generalize (37) to any number of time slices as

$$U(i, f; T = t_f - t_i) \approx C \int_{\substack{x_n=x_L \\ \dots \\ x_2=x_L \\ x_1=x_L}}^{\substack{x_n=x_R \\ \dots \\ x_2=x_R \\ x_1=x_R}} e^{i \int_{t_i}^{t_f} \frac{L}{\hbar} dt} dx_1 dx_2 \dots dx_n,$$

Approximation for Transition Amplitude

(38)

where, as before, it is implicit that L in the integral is for the particular path that crosses the respective t slices at x_1, x_2, \dots, x_n .

8.4 From Approximate to Exact

To get a precise, not approximate, relation for the RHS of (38) we would have to do two things.

1. Take the x range from l to infinity, i.e., $x_L \rightarrow -\infty$ and $x_R \rightarrow \infty$, and
2. Take $\Delta t \rightarrow dt$ for the same T (time between events.)

Doing this, (38) would become

$$U(i, f; T = t_f - t_i) = C \int_{\substack{x=x_i \\ \underbrace{\text{integ limits along with}} \\ \mathcal{D} \text{ symbol imply all paths between } i \text{ and } f}}^{x=x_f} e^{i \int_{t_i}^{t_f} \frac{L}{\hbar} dt} \mathcal{D}x$$

Exact Expression for Transition Amplitude

(39)

The symbol \mathcal{D} , as noted earlier, represents integration over all paths. With this, the integration limits designate the initial and final x values and do not imply a constraint on the x dimension during the integration (as was the case with (38).) In (39) we have, at long last, obtained the relation of integration type #4 in Table 1, where

$$F = e^{i \int_{t_i}^{t_f} \frac{L}{\hbar} dt}. \quad (40)$$

8.5 Practicality and Calculations

Practically, for the first approximation addressed in Section 8.4, we really don't have to take l to infinity, as we know that paths outside of a reasonably large range from the initial and final spatial locations will sum to very close to zero. So we can live with significant, but not infinite, l .

For the second approximation, we only need small enough Δt such that taking a smaller value does not change our answer much.

If we use (38), with judicious choices for Δt and l , we can, in many cases, obtain valid closed form solutions for the amplitude. We can also obtain numerical solutions with a digital computer by using approximations for L between time slices, as we did previously. That is, we can approximate the RHS of (38) in the manner we did for the first line of (37), but extending the approximation of (37) from 2 to n time slices.

9 An Example: Free Particle

We will first determine the amplitude (and thus the detection probability density) of a free particle via the Schroedinger approach and then compare it to that for Feynman's many paths approach.

9.1 Schroedinger Transition Amplitude

Recall, from Section 3.2, that, in the Schroedinger approach, a position eigenstate is a delta function, and as it evolves, the wave function envelope spreads and the peak diminishes. $|U|^2$ for such functions is the probability density at the final point x_f , after time T , where the peak is located. We should then expect $|U|^2$ to decrease as T increases, and to equal infinity when $T=0$.

We start with the Schroedinger transition amplitude relation (9),

$$U(x_i, x_f; T) = \langle x_f | e^{-iHT/\hbar} | x_i \rangle. \quad (41)$$

Since the bra and ket here are Dirac delta functions, with the well known relation

$$\delta(x - x_i) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ik(x-x_i)} dk = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} e^{\frac{iP}{\hbar}(x-x_i)} dp, \quad (42)$$

we can re-write (41) as

$$U(x_i, x_f; T) = \int_{-\infty}^{\infty} (\delta(x - x_f) e^{-iHT/\hbar} \delta(x - x_i)) dx. \quad (43)$$

(For readers unfamiliar with operators in exponents, one can express the exponential quantity as a Taylor series expanded about T , i.e., $f(T) = e^{-iTH/\hbar} = 1 - iTH/\hbar + \frac{1}{2} T^2 H^2 / \hbar^2 + \dots$. Then, operate on the ket/state term by term [getting terms in iET/\hbar to various powers], and finally re-express the resulting Taylor series as an exponential in iET/\hbar . We have taken the ket with time $t_i = 0$ to make things simpler, but even if you like to think of the Hamilton operator as a time derivative, when it acts on that ket, it functions as an energy operator and still yields the energy.)

For the exponential with the H operator acting on the initial state, and $E = p^2/2m$, (43) is

$$\begin{aligned} U(x_i, x_f; T) &= \int_{-\infty}^{\infty} \left(\left(\frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} e^{\frac{i}{\hbar} \frac{p'}{\hbar} (x - x_f)} dp' \right) \left(\frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} e^{-\frac{i}{\hbar} TH} e^{\frac{i}{\hbar} \frac{p}{\hbar} (x - x_i)} dp \right) \right) dx \\ &= \int_{-\infty}^{\infty} \left(\left(\frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} e^{\frac{i}{\hbar} \frac{p'}{\hbar} (x_f - x)} dp' \right) \left(\frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} e^{-iTp^2/2m\hbar} e^{\frac{i}{\hbar} \frac{p}{\hbar} (x - x_i)} dp \right) \right) dx. \end{aligned} \quad (44)$$

We then re-arrange (44) to get

$$\begin{aligned} U(x_i, x_f; T) &= \frac{1}{2\pi\hbar} \iint e^{-iTp^2/2m\hbar} \underbrace{\left(\frac{1}{2\pi\hbar} \int e^{\frac{i}{\hbar} \frac{x}{\hbar} (p' - p)} dx \right) e^{\frac{i}{\hbar} \frac{p' x_f}{\hbar}} e^{-\frac{i}{\hbar} p x_i}}_{\delta(p' - p)} dp' dp \\ &= \frac{1}{2\pi\hbar} \int e^{-iTp^2/2m\hbar} e^{\frac{i}{\hbar} p(x_f - x_i)} dp. \end{aligned} \quad (45)$$

Using the integral formula

$$\int_{-\infty}^{+\infty} e^{-ax^2 + bx} dx = \sqrt{\frac{\pi}{a}} e^{b^2/4a}, \quad (46)$$

we find

$$U(x_i, x_f; T) = \sqrt{\frac{m}{i2\pi\hbar T}} e^{\frac{i}{\hbar} \frac{m}{2T} (x_f - x_i)^2}. \quad (47)$$

The astute reader may question whether (46), with complex a and b , converges. It does because the integrand oscillation rate increases with larger $|p|$ in such a way as to make successive cycles shorter. As $|p|$ gets very large, the cycles become so short that the contribution from each cycle (think area under a sine curve) tends to zero, and it does so in a manner that allows the integral to converge.

From (47), the probability density at event \mathbf{f} is

$$|U(x_i, x_f; T)|^2 = \frac{m}{2\pi\hbar T}, \quad (48)$$

which, as we said it must, decreases with increasing T , and equals infinity for $T = 0$. Note also, that increasing m increases the envelope height, and thus decreases the width (for constant area under the envelope = constant probability.) In other words, the wave packet approaches more classical behavior, i.e., a narrower, more well defined location. Further, if \hbar were to go to zero, the peak would be infinite, i.e., we would have a delta function and an exact particle location, as in classical mechanics.

9.2 Many Paths Transition Amplitude

We now seek to derive (47) using the many paths approach.

A free, non-relativistic, particle has Lagrangian

$$L = \frac{1}{2}mv^2 \approx \frac{1}{2}m\left(\frac{x(t + \Delta t) - x(t)}{\Delta t}\right)^2, \quad (49)$$

where the RHS is an approximation between adjacent time slices. Taking $t_i = 0$, and $l \rightarrow \infty$, (38) becomes

$$\begin{aligned} U(i, f; T) &\approx C \int_{x_n=-\infty}^{x_n=\infty} \dots \int_{x_2=-\infty}^{x_2=\infty} \int_{x_1=-\infty}^{x_1=\infty} e^{i \int_{t_n}^{t_f} \frac{L}{\hbar} dt} e^{i \int_{t_{n-1}}^{t_n} \frac{L}{\hbar} dt} \dots e^{i \int_{t_0}^{t_1} \frac{L}{\hbar} dt} e^{i \int_0^{t_1} \frac{L}{\hbar} dt} dx_1 dx_2 \dots dx_n \\ &\approx C \int_{x_n=-\infty}^{x_n=\infty} \dots \int_{x_2=-\infty}^{x_2=\infty} \int_{x_1=-\infty}^{x_1=\infty} e^{\frac{i}{\hbar} \left[\frac{1}{2} m \left(\frac{x_f - x_n}{\Delta t} \right)^2 \right] \Delta t} \dots e^{\frac{i}{\hbar} \left[\frac{1}{2} m \left(\frac{x_2 - x_1}{\Delta t} \right)^2 \right] \Delta t} e^{\frac{i}{\hbar} \left[\frac{1}{2} m \left(\frac{x_1 - x_0}{\Delta t} \right)^2 \right] \Delta t} dx_1 dx_2 \dots dx_n. \quad (50) \\ &= C \underbrace{\int_{x_n=-\infty}^{x_n=\infty} e^{\frac{im}{2\hbar(\Delta t)}(x_f - x_n)^2} \dots}_{f_\zeta} \underbrace{\int_{x_2=-\infty}^{x_2=\infty} e^{\frac{im}{2\hbar(\Delta t)}(x_2 - x_1)^2} \dots}_{f_\gamma} \underbrace{\int_{x_1=-\infty}^{x_1=\infty} e^{\frac{im}{2\hbar(\Delta t)}(x_1 - x_0)^2} \dots}_{f_\beta} \underbrace{e^{\frac{im}{2\hbar(\Delta t)}(x_1 - x_0)^2}}_{f_\alpha} dx_1 dx_2 \dots dx_n, \end{aligned}$$

where the underbracket notation will help us in subsequent sections.

9.2.1 Background Math

Look, for the moment, at the last two factors (functions f_α and f_β) in the integral. They must be integrated over x_1 , and that result is a function of x_2 . When one of the two functions in such a procedure is a function of $x_2 - x_1$, as it is here, the integral is called a *convolution integral*. (See <http://www-structmed.cimr.cam.ac.uk/Course/Convolution/convolution.html>.)

In mathematical circles (search “Borel’s Theorem”), it is well known that the Fourier (and also, the Laplace) transform of such an integral equals the product of the Fourier (or Laplace) transforms of the two functions. That is, for \mathcal{F} representing Fourier transform,

$$\mathcal{F} \left\{ \int f_\beta(x_2 - x_1) f_\alpha(x_1) dx_1 \right\} = \mathcal{F} \{ f_\beta \} \mathcal{F} \{ f_\alpha \}. \quad (51)$$

Note that although f_α is a function of $x_1 - x_i$, we can write $f_\alpha(x_1)$ because x_i is fixed.

Each factor in the last row of (50), as one moves leftward, plays the part of f_β in the theorem above for the next convolution integral, where the prior convolution integral plays the role of f_α . We get, in essence, a series of nested convolution integrals. Using (51), you should be able to prove to yourself that the transform of (50) equals the product of the transforms of the exponential factors in (50). If you can’t, or don’t want to bother, proving it, then just accept that a corollary to (51) is

$$\begin{aligned} \mathcal{F} \left\{ \int \dots \int \int f_\zeta(x_f - x_n) \dots f_\gamma(x_3 - x_2) f_\beta(x_2 - x_1) f_\alpha(x_1) dx_1 dx_2 \dots dx_n \right\} \\ = \mathcal{F} \{ f_\zeta \} \dots \mathcal{F} \{ f_\gamma \} \mathcal{F} \{ f_\beta \} \mathcal{F} \{ f_\alpha \}. \end{aligned} \quad (52)$$

9.2.2 Evaluating the Amplitude

So, to evaluate (50), using (52), we i) transform each exponential factor f_μ , ii) multiply those transforms together, and iii) take the inverse transform of the result to get U (actually U/C

of (50)). This is made simpler, because each f_μ has the same form, so each transform is the same, i.e.,

$$\mathcal{F}\{f_\alpha\} = \mathcal{F}\{f_\beta\} = \dots = \mathcal{F}\{f_\zeta\}. \quad (53)$$

The Fourier transform of a function f_α is

$$\mathcal{F}\{f_\alpha(x_i)\} = \tilde{f}_\alpha(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} f_\alpha(x_i) e^{-\frac{i}{\hbar} px_i} dx_i. \quad (54)$$

For the f_α of (50), and for convenience, taking the coordinate $x_i = 0$, this is

$$\tilde{f}_\alpha(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{\frac{i}{\hbar 2(\Delta t)} x_i^2} e^{-\frac{i}{\hbar} px_i} dx_i, \quad (55)$$

where here and throughout this section, p is merely a dummy variable allowing us to carry out the math. Using (46), we find (55) becomes

$$\tilde{f}_\alpha(p) = \sqrt{\frac{i(\Delta t)}{m}} e^{-\frac{i}{\hbar 2m} p^2}, \quad (56)$$

and thus, from (50), (52), and (53),

$$\tilde{U}(p) \approx C \tilde{f}_\zeta(p) \dots \tilde{f}_\beta(p) \tilde{f}_\alpha(p) = C \left(\frac{i(\Delta t)}{m} \right)^{N/2} e^{-\frac{i}{\hbar 2m} p^2}. \quad (57)$$

The inverse Fourier transform of (57), is

$$\begin{aligned} U(x_i, x_f; T) &= \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \tilde{U}(p) e^{\frac{i}{\hbar} p(x_f - x_i)} dp \\ &\approx \frac{1}{\sqrt{2\pi\hbar}} \left(\frac{i(\Delta t)}{m} \right)^{N/2} C \int_{-\infty}^{\infty} e^{-\frac{i}{\hbar 2m} p^2} e^{\frac{i}{\hbar} p(x_f - x_i)} dp. \end{aligned} \quad (58)$$

In (58), we could have simply used x_f in the exponent, as we have been taking $x_i = 0$, and our result would have been in terms of x_f . In that case, x_f would have been the distance between x_i and x_f , i.e., $x_f - x_i$. In order to frame our final result in the most general terms, we re-introduced x_i as having any coordinate value in (58).

With (46) again, (58) becomes

$$U(x_i, x_f; T) \approx C \left(\frac{i(\Delta t)}{m} \right)^{N/2} \sqrt{\frac{m}{iT}} e^{\frac{i}{\hbar 2T} (x_f - x_i)^2}. \quad (59)$$

By comparison with (47), we see the phase and dependence on T is the same as in the wave mechanics approach. Using that comparison, we can see that the constant of proportionality is

$$C = \frac{1}{\sqrt{2\pi\hbar}} \left(\frac{m}{i(\Delta t)} \right)^{N/2}. \quad (60)$$

And thus, the probability density at the final event \mathbf{f} is the same as (48), i.e.,

$$\left|U(x_i, x_f; T)\right|^2 = \frac{m}{2\pi\hbar T}, \quad (61)$$

where the equal sign is appropriate for $N \rightarrow \infty$.

Note that for $v = (x_f - x_i)/T$, the amplitude can be expressed in terms of the classical action as

$$U(x_i, x_f; T) = \sqrt{\frac{m}{i2\pi\hbar T}} e^{\frac{i mv^2 T}{\hbar} - \frac{i LT}{\hbar}} = \sqrt{\frac{m}{i2\pi\hbar T}} e^{\frac{i S}{\hbar}}, \quad (62)$$

which agrees with (22) if $A(t)$ there equals the root quantity. In the Appendix, we show it does.

9.3 The Message

It has probably not escaped the reader that the evaluation of a free particle using Feynman's many paths approach is considerably more complicated and lengthy than the Schroedinger approach. This is true for most, if not all, problems in QM.

The disadvantages of the many paths approach in QM are these.

1. It is generally more mathematically cumbersome and time consuming than the wave mechanics approach.
2. The quantity calculated is only proportional to the amplitude, and further analysis is required to determine the precise amplitude.
3. The approach is suitable primarily for position eigenstates and is not readily amenable to more general states, so it is generally not as all encompassing in nature.

The advantages of the many paths approach are these.

1. The approach also applies to quantum field theory (QFT). In a number of instances therein, development of the theory is more direct, and calculation of amplitudes is easier, than with the alternative approach (canonical quantization).
2. Philosophically, we see that there is more than one way to skin a cat. We learn anew that the physical world can be modeled in different, equivalent ways. We learn caution with regard to interpreting a given model as an actual picture of reality.

10 Quantum Field Theory via Path Integrals

So far, we have dealt primarily with non-relativistic quantum mechanics (QM), but the many paths approach is also applicable to relativistic quantum mechanics (RQM), and as noted above, to quantum field theory (QFT). (RQM is often confused with QFT. For a comparison of the similarities and differences between the two, see [Quantum Field Theory: A Pedagogic Intro](#). Further similarities and differences are illustrated in Unifying Chart 2, below.)

We will not go deeply into QFT, and only outline, in a broad overview, how the theory presented herein is applicable therein. This should help those students who continue on to the standard texts for the many paths approach keep the forest in view while examining the trees.

10.1 Particle Theory (QM) vs Field Theory (QFT)

For the many paths approach, we want to make the jump from QM, which is a quantized version of particle theory, to QFT, which is a quantized version of field theory. Unifying Chart 2

below can help us do that. In it, the 2nd and 3rd columns compare particle theory entities/concepts to corresponding field theory entities/concepts. The upper half of the chart, as indicated, summarizes classical theory (non quantum, and implicitly including special relativity). The lower half summarizes quantum theory *via approaches other than many paths*. The chart should be relatively self explanatory, so we will not comment much on it.

We compare the quantum approaches of Unifying Chart 2 to the many paths approach in the next section.

Unifying Chart 2. Comparing Particle Theory to Field Theory: Classical and Quantum

	Particle Theory	Field Theory
	<u>Classical Theory</u>	
Indep variables	<u>1D</u> t	<u>3D</u> t
Depend variables	$x(t)$ position	$\phi(x,y,z,t)$ field
Dynamic variables (functionals)	Particle total value: \mathbf{p}, E, L functions of x, \dot{x}, t (or $\mathbf{r}, \dot{\mathbf{r}}, t$)	Density values (per unit vol): $p, \mathcal{E}, \mathcal{L}$ functions of $\phi, \dot{\phi}, x, y, z, t$ $E = \oint \mathcal{E} d^3x$, etc.
Equations of motion	$\mathbf{F} = m\mathbf{a}$ or equivalently, Euler-Lagrange formulation, $\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} = 0$	$\mathbf{f} = \rho \mathbf{a}$ (force/vol) for media; Maxwell's eqs for e/m, or equivalently, for \mathcal{L} of either, $\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}} \right) - \frac{\partial \mathcal{L}}{\partial \phi} = 0$
Variable correspondences particle \leftrightarrow field	$t \leftrightarrow x, y, z, t$ $x \leftrightarrow \phi$ total values \leftrightarrow density values	
	<u>Quantum Theories</u>	
	QM and RQM via Wave Mechanics	QFT via Wave Mechanics = Canonical Quantization
Quantum character change	x and all dynamic variables \rightarrow operators	ϕ and all dynamic variables \rightarrow operators
New quantum entity	state $ \psi\rangle$ = wave function ψ	state $ \phi\rangle$ different from (operator) field ϕ
Note	Fields create & destroy states. States can be multiparticle ($ \phi_1, \phi_2, \dots\rangle$)	

Operators	functions of x, \dot{x}, t	functions of $\phi, \dot{\phi}, t$
Expectation values of operators	$\bar{E} = \langle \psi H \psi \rangle$ etc. for other opers	$\bar{E} = \langle \phi H \phi \rangle$ or for multiparticle state $\bar{E} = \langle \phi_1, \phi_2 \dots H \phi_1, \phi_2 \dots \rangle$
Equations of motion	For wave function ψ QM: Schrödinger eq RQM: Klein-Gordon, Dirac, Proca eqs or equivalently, Euler-Lagrange formulations	For quantum field ϕ QFT: Klein-Gordon, Dirac, Proca eqs or equivalently, Euler-Lagrange formulations
Macro equations of motion	Deduced from above and expectation values of force, acceleration	Deduced from above and expectation values of relevant quantities
Transition amplitude U	$U(x_i, x_f; T) = \langle x_f e^{-iHT} x_i \rangle$ i & f are eigen states of position	$U(\phi_i, \phi_f; T) = \langle \phi_f e^{-iHT} \phi_i \rangle$ i & f states can be multiparticle
$ U ^2 =$	probability density	probability

10.2 “Derivation” of Many Paths Approach for QFT

From the last row of Unifying Chart 2, we see that the transition amplitude for the QFT canonical approach, which is essentially a wave mechanics approach for relativistic fields, is similar in form to that of the QM wave mechanics approach, given that we note the correspondence $x \rightarrow \phi$ between QM and QFT. An additional fundamental difference between the two is the form of the Hamiltonian H . In QM, H is a non-relativistic function of x , dx/dt , and (rarely) t . In QFT, it is a relativistic function of ϕ , $d\phi/dt$, and (rarely) t .

Since the canonical (wave mechanics) QFT approach mirrors the wave mechanics QM approach, one could postulate (and Feynman probably did) that the many paths approach in QFT would mirror the many paths approach in QM. (See Unifying Chart 1 in Section 7.2 for the corresponding QM transition amplitudes using each approach.) Simply using the same correspondences $x \rightarrow \phi$ and $H_{\text{nonrel}} \rightarrow H_{\text{rel}}$ (and thus, $L_{\text{nonrel}} \rightarrow L_{\text{rel}}$) for the many paths approach yields Unifying Chart 3.

Unifying Chart 3. Comparing QM to QFT for the Many Paths Approach

	<u>Quantum Theories</u>	
	QM and RQM via Many Paths	QFT via Many Paths
Transition amplitude	$U(x_i, x_f; T) \propto \lim_{N \rightarrow \infty} \sum_{j=1}^N e^{iS_j/\hbar}$ $= \int_{x_i}^{x_f} e^{i \int_0^T \frac{L}{\hbar} dt} \mathcal{D}x(t)$	$U(\phi_i, \phi_f; T) \propto \lim_{N \rightarrow \infty} \sum_{j=1}^N e^{iS_j/\hbar}$ $= \int_{\phi_i}^{\phi_f} e^{i \int_0^T \frac{\mathcal{L}}{\hbar} d^4x} \mathcal{D}\phi(x^\mu)$
Note	Above is from Unifying Chart 1 in Section 7.2	Above is a simplified example for a single scalar field.

In the RH column above, all paths, comprising all configurations of the entire field ϕ over all space between its initial and final configurations, are added (integrated). S here is the action for the entire field. \mathcal{L} is the (relativistic) Lagrangian *density* for the field, which, integrated as it is above over all space d^3x , yields L .

Of course, the many paths transition amplitude of Unifying Chart 3 is, at this point, only a guess. However, decades of research, first by Feynman and then by many others, has proven that it is completely valid.

To summarize, briefly

Unifying Chart 4. Super Simple Summary

Correspondences	$x \rightarrow \phi$ $H_{\text{nonrel}} \rightarrow H_{\text{rel}}$	
Wave mechanics amplitude	QM \rightarrow QFT	<u>canonical quantization</u> QFT
Many paths amplitude	QM \rightarrow QFT	<u>functional quantization</u> QFT

10.3 Time Slicing in QFT

Using the same correspondences as in Unifying Chart 4, and the time slicing approximation for QM of (38), we find, for QFT,

$$U(i, f; T = t_f - t_i) \approx C \int \dots \int \int e^{i \int_{t_i}^{t_f} \frac{\mathcal{L}}{\hbar} d^4x} d\phi_1 d\phi_2 \dots d\phi_n , \quad (63)$$

QFT Approximation for Transition Amplitude

where the subscripts refer to different time slices, not to different fields. This example is for only a single field.

The exact form of the transition amplitude, obtained from (39), is given in Unifying Chart 3, and is repeated here,

$$U(i, f; T) = C \int_{\phi_i}^{\phi_f} e^{i \int_0^T \frac{L}{\hbar} d^4x} \mathcal{D}\phi(x^\mu)$$

QFT Exact Expression for Transition Amplitude

(64)

10.4 More Ahead in Path Integral QFT

Of course, we have only scratched the surface of the many paths approach to QFT. There is a great deal more, including some fairly fundamental concepts. However, hopefully, all of the above will provide a solid foundation for that, by explaining more simply, more completely, and in smaller steps of development what traditional introductions to the subject often treat rapidly and in somewhat less than transparent fashion.

If you find errors (typographical or otherwise), or have suggestions on how to make anything herein easier to understand, please help those who come after you by letting me know, so I can make appropriate corrections/modifications. I can be reached via the email address in the home page, [Pedagogic Aides to Quantum Field Theory](#)⁶, for which this material is a sub-section. Thank you.

Distribution of this material to others is encouraged though subject to (fairly liberal) copyright restrictions. These can be found at the above link as well.

Robert D. Klauber
April 30, 2009

Appendix

From (20), with $x=x_{\text{peak}}$,

$$\psi(x_{\text{peak}}, t) = e^{-\frac{i}{\hbar}(\bar{E}t - \bar{p}x)} \underbrace{\frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} g(p) e^{-\frac{i}{2\hbar m}(p - \bar{p})^2} dp}_{A(t)}. \quad (65)$$

We note that (20), and thus (65), are derived from the general wave packet relation (see ref 2)

$$\psi(x, t) = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} g(p) e^{-\frac{i}{\hbar}(Et - px)} dp. \quad (66)$$

At our initial event, take $t=t_i=0$, so the above becomes

$$\psi_i(x, t=0) = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} g(p) e^{\frac{i}{\hbar}px} dp. \quad (67)$$

If (67) is a delta function centered at $x_i=0$, then, from the definition of the delta function,

$$\psi_i(x, t=0) = \delta(x) = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} e^{\frac{i}{\hbar}px} dp. \quad (68)$$

Comparing (68) to (67), we see that for an initial delta function measured at x_i

$$g(p)=1. \quad (69)$$

Using (69) in (65), we obtain

$$A(t) = \frac{1}{2\pi\hbar} \int_{-\infty}^{+\infty} e^{-\frac{i}{2\hbar m}(p-\bar{p})^2} dp . \quad (70)$$

With the integral formula

$$\int_{-\infty}^{+\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}} , \quad (71)$$

we find

$$A(t) = \sqrt{\frac{m}{i2\pi\hbar t}} . \quad (72)$$

¹ <http://www.quantumfieldtheory.info/Introduction-Background.htm>

² Merzbacher, E., *Quantum Mechanics*. 2nd ed. John Wiley & Sons (1970). See Chap 2, Sec 3.

³ Much of the material in this section parallels “Action on Stage: Historical Introduction”, Ogborn, J., Hanc, J. and Taylor, E.F., and “A First Introduction to Quantum Behavior”, Ogborn, J., both from The Girep Conference 2006, Modeling Physics and Physics Education, Universiteit van Amsterdam.

⁴ Feynman, R., *QED: The strange theory of light and matter*. Penguin Books, London (1985).

⁵ <http://www.quantumfieldtheory.info/manypathsgraphicelectronreflection.xls>

⁶ <http://www.quantumfieldtheory.info>

Preface

This book is

1. an attempt to make learning quantum field theory (QFT) as easy as is humanly possible,
2. intended, first and foremost, for new students of QFT, and
3. an introduction to only the most fundamental and central concepts of the theory.

It is not

1. orthodox,
2. an exhaustive treatment of QFT,
3. written for seasoned practitioners in the field, or
4. a presentation of the latest, most modern approach to it.

Students planning a career in field theory will obviously have to move on to more advanced texts, after they digest the more elementary material presented herein. This book is intended to provide a solid foundation in the most essential elements of the theory, nothing more.

In my own teaching experience, and in the course of researching pedagogy, I have come to see that “learning” has at its basis a fundamental three-in-one structure. The wholeness of learning is composed of

- i) the knowledge to be learned,
- ii) the learner, and
- iii) the process of learning itself.

It seems unfortunate that physics and physics textbooks have too often been almost solely concerned with the *knowledge* of physics and only rarely concerned with *those who are learning it* or *how they could best go about learning*. However, there are signs that this situation may be changing somewhat, and I hope that this book will be one stepping stone in that direction.

In writing this book, I have repeatedly tried to visualize the learning process as a new learner would. This viewpoint is one we quickly lose when we, as teachers and researchers, gain familiarity with a given subject, and yet it is a perspective we must maintain if we are to be effective educators. To this end, I have solicited, and continue to solicit, guidance and suggestions from professional educators (those who make learning and education, *per se*, their central focus in life), and more importantly, from those who are studying QFT for the first time. In addition, I have used my own notes, compiled when I was first studying the theory myself, in which I carefully delineated ways the subject could be presented in a more student-friendly manner. In this sense, the text incorporates “peer instruction”, a pedagogic tool of recognized, and considerable, merit, wherein students help teach fellow students who are learning the same subject.

It is my sincere hope that the methodologies I have employed herein have helped me to remain sympathetic to, and in touch with, the perspective of a new learner. Of course, different students find different teaching techniques to have varying degrees of transparency, so there are no hard and fast rules. However, I do believe that most students would consider many of the following principles, which I have employed in the text, to be of pedagogic value.

1) Holistic previews

The entire book, each chapter, and many sections begin with simple, non-mathematical overviews of the material to be covered. These allow the student to gain a qualitative understanding of the “big picture” before he or she plunges into the rigors of the underlying mathematics.

Doing physics is a lot like doing a jig-saw puzzle. We assemble bits and pieces into small wholes and then gradually merge those small wholes into greater ones, until ultimately we end up with the “big picture.” Seeing the picture on the puzzle box before we start has immense value in helping us put the whole thing together. We know the blue goes here, the green there, and the boundary of the two, somewhere in between. Without that picture preview to guide us, the entire job becomes considerably more difficult, more tedious, and less enjoyable. In this book, the holistic previews are much like the pictures on the puzzle boxes. The detail is not there, but the essence of the final goal is. These overviews should

eliminate, or at least minimize, the “lost in a maze of equations” syndrome by providing a “birds-eye road map” of where we have come from, and where we are going. By so doing we not only will keep sight of the forest in spite of the trees, but will also have a feeling, from the beginning, for the relevance of each particular topic to the overriding structure of the wholeness of knowledge in which it is embedded.

2) Schematic diagram summaries (Wholeness Charts)

Enhancing the “birds-eye road map” approach are block diagram summaries, which I call *Wholeness Charts*, so named because they reveal in chart form the underlying connections that unite various aspects of a given theory into a greater whole. Unlike the chapter previews, these are often mathematical and contain considerable theoretical depth.

Learning a computer program line-by-line is immensely harder than learning it with a block diagram of the program, showing major sections and sub-sections, and how they are all interrelated. There is a structure underlying the program, which is its essence and most important aspect, but which is not obvious by looking directly at the program code itself.

The same is true in physics, where line-by-line delineation of concepts and mathematics corresponds to program code, and in this text, Wholeness Charts play the role of block diagrams. In my own learning experiences, in which I constructed such charts myself from my books and lecture notes, I found them to be invaluable aids. They coalesced a lot of different information into one central, compact, easy-to-see, easy-to-understand, and easy-to-reference framework.

The specific advantages of Wholeness Charts are severalfold.

First, in learning any given material we are seeking, most importantly, an understanding of the kernel or conceptual essence, i.e., the main idea(s) underlying all the text. A picture is worth a thousand words, and a Wholeness Chart is a “snapshot” of those thousand words.

Second, although the charts can summarize in-depth mathematics and concepts, they can be used to advantage even when reading through material for the first time. The holistic overview perspective can be more easily maintained by continual reference to the schematic as one learns the details.

Thirdly, comparison with similar diagrams in related areas can reveal parallel underlying threads running through seemingly diverse phenomena. (See, for example, Classical Mechanics Overview Wholeness Chart XXX? and Quantum Mechanics Overview Wholeness Chart XXX? in Chapter XXX? Section ?XXXXX) This not only aids the learning process but also helps to reveal some of the subtle workings and unified structure inherent in Mother Nature.

Further, review of material for qualifying exams or any other future purpose is greatly facilitated. It is much easier to refresh one’s memory, and even deepen understanding, from one or two summary sheets, rather than time consuming ventures through dozens of pages of text. And by copying all of the Wholeness Charts herein and stapling them together, you will have a pretty good summary of the entire book.

Still further, the charts can be used as quick and easy-to-find references to key relations at future times, even years later.

3) Reviews of background material

In situations where development of a given idea depends on material studied in previous courses (e.g., quantum mechanics) short reviews of the relevant background subject matter are provided, usually in chapter introductory sections or later on, in special boxes separate from the main body of the text.

4) Only basic concepts without peripheral subjects

I believe it is of primary importance in the learning process to focus on the fundamental concepts first, to the exclusion of all else. The time to branch out into related (and usually more complex) areas is *after* the core knowledge is assimilated, *not during* the assimilation period.

All too often, students are presented with a great deal of new material, some fundamental, other more peripheral. The peripheral material not only consumes precious study time, but tends to confuse the student with regard to what precisely is essential (what he or she *must* understand), and what is not (what it would be *nice if* he or she also understood).

This book, by careful intention, restricts itself to only the most core principles of QFT. Once those principles are well in hand, the student should then be ready to glean maximum value from standard, more extensive, texts.

5) Optimal “return on investment” exercises

All too often students get tied up, for what seem interminable periods, working through problems from which minimum actual learning is reaped. Study time is valuable and spending it engulfed in great quantities of algebra and trigonometry is probably not its best use.

I have tried, as best I could, to design the exercises in this book so that they consume minimum time but yield maximum return. Emphasis has been placed on gleaning an understanding of concepts without getting mired down.

Later on, when students have become practicing researchers and time pressure is not so great, there will be ample opportunity to work through more involved problems down to every last minute algebraic detail. If they are firmly in command of the *concepts* and *principles* involved, the calculations, though often lengthy, become trivial. If, however, they never got quite grounded in the fundamentals because study time was not efficiently used, then research can go slowly indeed.

6) Many small steps, rather than fewer large ones

Professional educators have known for some time now that learning progresses faster and more profoundly when new material is presented in small bites. The longer, more moderately sloped trail can get one to the mountaintop much more readily than the agonizing climb up the steep vertical face.

Unfortunately, from my personal experience as a student, it often seemed like my textbooks were trying to take me up the steepest grade. I sincerely hope that those using this book do not have this experience. I have made every effort to include each and every relevant step in all derivations and examples.

7) Liberal use of simple concrete examples

Professional educators have also known for quite some time that abstract concepts are best taught by leading into them with simple, physically visualizable examples. Further, their understanding is deepened, broadened, and solidified with even more such concrete examples.

Some may argue that a more formal mathematical approach is preferable because it is important to have a profound, not superficial, understanding. While I completely agree that a profound understanding is essential, it is my experience that the mathematically rigorous introduction, more often than not, has quite the opposite result. (Ask any student about this.) Further, to know any field profoundly we must know it from all angles. We must know the underlying mathematics in detail *plus* we must have a grasp on what it all means in the real world, i.e., how the relevant systems behave, how they parallel other types of systems with which we are already familiar, etc. Since we have to cover the whole range of knowledge from abstract to physical anyway, it seems best to start with the end of the spectrum most readily apprehensible (i.e., the visualizable, concrete, and analogous) and move on from there.

This methodology is employed liberally in this book. It is hoped that so doing will ameliorate the “what is going on” frustration common among students who are introduced to conceptually new ideas almost solely via routes heavily oriented toward abstraction and pure mathematics.

In this context it is relevant that Richard Feynman, in his autobiography, notes,

“I can’t understand anything in general unless I’m carrying along in my mind a specific example and watching it go....(Others think) I’m following the steps mathematically but that’s not what I’m doing. I have the specific, physical example of what (is being analyzed) and I know from instinct and experience the properties of the thing.”

I know from my own experience that I learn in the same way, and I have a suspicion that almost everyone else does as well. Yet few teach that way. This book is an attempt to teach in that way.

8) Margin overview notes

Within a given section of any textbook, one group of paragraphs can refer to one subject, another group to another subject. When reading material for the first time, not knowing exactly where one train of the author’s thought ends and a different one begins can oftentimes prove confusing. In this book, each new idea not set off with its own section heading is highlighted, along with its central message, by notations in the margins. In this way, emphasis is once again placed on the overview, the “big picture” of each topic, even on the subordinate levels within sections and subsections.

Additionally, the extra space in the margins can be used by students, to make their own notes and comments. In my own experience as a student I found this practice to be invaluable. My own remarks written in a book are, almost invariably, more comprehensible to me when reviewing later for exams or other purposes than are those of the author.

9) Definitions and key equations emphasized

As a student, I often found myself encountering a term that had been introduced earlier in the text, but not being clear on its exact meaning, I had to search back through pages clumsily trying to find the first use of the word. In this book, new terminology is underlined when it is introduced or defined, so that it “jumps out” at the reader later when trying to find it again.

In addition, key equations the ones students really need to know have borders around them.

10) Non use of terms like “obvious”, “trivial”, etc.

The text avoids use of emotionally debilitating terms such as “obvious”, “trivial”, “simple”, “easy”, and the like to describe things that may, after years of familiarity, be easy or obvious to the author, but can be anything but that to the new student. (See “A Nontrivial Manifesto” by Matt Landreman, *Physics Today*, March 2005, 52-53.)

The job I have undertaken here has been a challenging one. I have sought to produce a physics textbook which is relatively lucid and transparent to those studying quantum field theory for the first time.

I suspect many physics professors will consider the book too verbose and too simple. I only ask them to try it and let their students be the judges. The proof will be in the pudding - if comprehension comes more quickly and more deeply, then the approach taken here will be vindicated.

Good luck to the new students of quantum field theory! I hope their studies are personally rewarding and professionally fruitful.

Robert D. Klauber

Prerequisites

Quantum field theory takes off where the following subjects end. Those beginning this book should be reasonably well versed in them, at the levels described below.

Quantum Mechanics

An absolute minimum of two undergraduate quarters. Far more preferably, an additional two graduate level quarters. Some exposure to relativistic quantum mechanics would be advantageous, but is not necessary. Optimal level of proficiency: Eugen Merzbacher’s *Quantum Mechanics* (John Wiley) or a similar book.

Classical Mechanics

A semester at the graduate level. Topics covered should include the Lagrangian formulation (for particles, and importantly, also for fields), the Legendre transformation, special relativity, and classical scattering. A familiarity with Poisson brackets would be helpful. Optimal level of proficiency: Herbert Goldstein’s *Classical Mechanics* (Addison-Wesley) or similar.

Electromagnetism

Two quarters at the undergraduate level plus two graduate quarters. Areas studied should comprise Maxwell’s equations, conservation laws, e/m wave propagation, relativistic treatment, Maxwell’s equations in terms of the four potential. Optimal level of proficiency: John David Jackson’s *Classical Electrodynamics* (John Wiley) or similar.

Math/Relativity

Exposure to covariant and contravariant coordinates, and metric tensors, for orthogonal 4D systems, at the level found in Jackson’s chapters on special relativity.