Preface

"All of physics is either impossible or trivial.
It is impossible until you understand it, and then it becomes trivial."

Ernest Rutherford

This book is

1. an attempt to make learning quantum field theory (QFT) as easy, and thus as efficient, 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, particularly as employed in quantum electrodynamics (QED).

It is not

1. orthodox,
2. an exhaustive treatment of QFT,
3. concise (lacking extensive explanation),
4. written for seasoned practitioners in the field, or
5. 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 guidance and suggestions from professional educators (those who make learning and education, *per se*, their central focus in life), and more importantly, from those 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) Brevity Avoided

Conciseness is typically a horror for new students trying to fathom unfamiliar concepts. While it can be advantageous in some arenas, it is almost never so in education. Unfortunately, being succinct, has, in scientific/technical circles, become a goal unto itself, extending even into pedagogy – an area for which it was never suited.

In this book, I have gone to great lengths to avoid conciseness and to present extensive explanations. I often take a paragraph or more for what other authors cover in a single sentence. I do this because I learned a long time ago that the thinnest texts were the hardest. Thicker ones covering the same material actually took less time to get through, and I understood them better, because the authors took time and space to elaborate, rather than leave significant gaps.

Such gaps often contain ambiguities or possibilities for misunderstanding that the author has overlooked and left unresolved. Succinctness may impress peers, but can be terribly misleading and frustrating for students.

2) 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.

3) 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.

Third, comparison with similar diagrams in related areas can reveal parallel underlying threads running through seemingly diverse phenomena. (See, for example, Summary of Classical Mechanics Wholeness Chart 2-2 and Summary of Quantum Mechanics Wholeness Chart 2-5 in Chap. 2, pgs. 20-21 and 30-31.) 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.

4) 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.

5) 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 or advanced. The peripheral/advanced 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 at this point in their development).

As one example, for those familiar with other approaches to QFT, this book does not introduce concepts appropriate to weak interactions, such as ϕ^4 theory, before students have first become grounded in the more elementary theory of quantum electrodynamics.

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 other, more extensive, texts.

6) 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 opportunities to work through more involved problems down to every 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.

7) 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 nearly 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.

In so doing, I have sought to avoid the common practice of letting students work out significant amounts of algebra that typically lies "between the lines". The thinking, as I understand it, is that students are perfectly capable of doing that themselves, so "why take up space with it in a text?"

My answer is simply that including those missing steps makes the learning process more efficient. If it takes the author ten minutes to write out two or three more lines of algebra, then it probably takes the student twenty minutes to do so, provided he/she is not befuddled (which is not rare, and in which case, it can take a great deal longer). That ten minutes spent by the author saves hundreds, or even thousands, of student readers twenty minutes, or more, each. Multiply that by the number of times such things occur per chapter and the number of chapters per book, and we are talking enormous amounts of student time saved.

Students learn very little, if anything, doing algebra. They recapture a lot of otherwise wasted time that can be used for actual learning, if the author types out the missing lines.

8) 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.

9) 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.

10) 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.

11) 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. In so doing, I have employed some decidedly non-traditional tactics, and so anticipated resistance from main stream publishers, who typically have motivations for wanting to do things the way they have been done before. Their respective missions do not seem, at least to me, to be focused primarily on optimizing the process of conveying knowledge.

As an example, a good friend of mine submitted a graduate level physics text manuscript, with student friendly notes in the margins, to one of the world's top academic publishers. He was ordered to remove the margin notes before they would publish the book. Not wanting to fight (and lose) this kind of battle over methodologies I employ, and consider essential in making students' work easier, I have chosen a different route.

I also anticipate resistance from some physics professors who may 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.

If you are a student now, appreciate the pedagogic methodologies used in this book, and end up one day writing a text of your own, I hope you will not forget what advantage you once gained from those methodologies. I hope you will use them in your own book. Above all, I hope your presentation will be profuse with elucidation and not terse.

Good luck to the new students of quantum field theory! May their studies be personally rewarding and professionally fruitful.

Robert D. Klauber
February 2013

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

Advantageous but not essential, as it is covered in the appendix of Chap. 2: Exposure to covariant and contravariant coordinates, and metric tensors, for orthogonal 4D systems, at the level found in Jackson's chapters on special relativity.

Acknowledgements

"You cannot live a perfect day without doing something
for someone who will never be able to repay you."

John Wooden
Hall of Fame UCLA basketball coach.

The people who reviewed, edited, made suggestions for, and corrected draft portions of this book had many candidate perfect days. There is no way I can repay them.

I am most indebted to three, Chris Locke, Christian Maennel and Mike Worsell, who read every word and made innumerable great contributions. Close behind on my gratitude list are Carlo Marino, David Scharf, Jean-Louis Sicaud, and Jon Tyrrell, each of whom read most of the text and provided a substantial number of valuable suggestions and corrections. David, Jon, and Morgan Orcutt deserve further heartfelt thanks for working most of the problems (and finding errors in several of them).

Others making significant, much appreciated contributions include Martin Bäker, Jim Bogan, Ben Brenneman, Brad Carlile, Bill Cohwig, Trevor Daniels, Saurya Das, Lorenzo Del Re, Tony D'Esopo, Paul Drechsel, Michael Gildner, Esteban Herrera, Phil Jones, Ruth Kastner, Lorek Krzysztof, Claude Liechti, Rattan Mann, Lorenzo Massimi, Enda McGlynn, Gopi Rajagopal, Javier Rubio, Girish Sharma, and Dennis Smoot.

Many years before I started writing this text, I fell in debt to my teachers, Robin Ticciati and John Hagelin, who guided me through my earliest sojourns into the quantum theory of fields, and earned both my respect and deep gratitude. Robin, in particular, was generous well beyond the call of duty, in granting me numerous one-on-one sessions to discuss various aspects of the theory.

Non technical, but nonetheless vital support came from my wonderful wife Susan. I cannot thank her enough for her patience, understanding, love, and unswerving devotion throughout the days, weeks, months, and years I spent writing … and re-writing. Last mentioned, yet anything but least, are my amazing and caring parents, without whose support and many, many sacrifices, I would never have gained the education I did, and thus, never have written this book. Thank you, mom and dad.

This book, whatever it is, would be substantially less without these people.

Regardless, any errors or insufficiencies that may still remain are my responsibility, and mine alone.

Chapter 1 *print vers 2/12/13Copyright of Robert D. Klauber*

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 will be going. This allows them to maintain 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).

The first of these is considered by many, and certainly by me, as the easiest way to be introduced to the subject, since it treats particles as objects that one can visualize as evolving along a particular path in spacetime, much as we commonly think of them doing. The path integral approach, on the other hand, treats particles and fields as if they were simultaneously traveling all possible paths, a difficult concept with even more difficult mathematics behind it.

This book is primarily devoted to the canonical quantization approach, though I have provided a simplified, brief introduction to the path integral approach in Chap. 18 near the end. Students wishing to make a career in field theory will eventually need to become well versed in both.

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. $\mu^- \rightarrow e^- + \nu + \bar{\nu}$, where the latter two symbols represent neutrino and antineutrino, respectively).

Limitation of original QM: no transmutation of particles

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.

QFT: transmutation included

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.

Energies <0
RQM yes
QFT no

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. 1-1. 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.

QFT example: Bhabha scattering

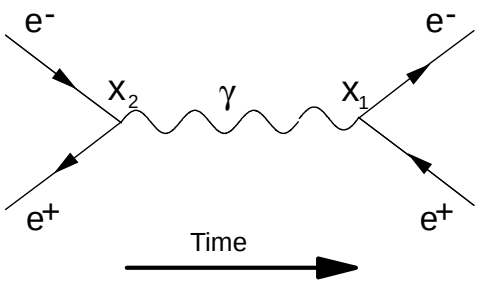


Figure 1-1. Bhabha Scattering

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. 1-1. 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 NRQM 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. 1-1. 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.

$$\text{Transition amplitude} = {}_{x_1}\left\langle e^+ e^- \middle| \underbrace{\left(\bar{\psi}_c \not{A}_d \psi_c\right)_{x_1}} \left(\bar{\psi}_d \not{A}_c \psi_d\right)_{x_2} \right| e^+ e^- \Big\rangle_{x_2} \quad . \tag{1-1}$$

In (1-1), the ket $|e^+e^-\rangle_{x2}$ 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 $\not{A}_c$ is a photon operator that creates (at x_2) a virtual photon, and $\not{A}_d$ 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

Destruction operators leave vacuum ket times a numeric factor

$$\text{transition amplitude} = {}_{x_1}\left\langle e^+ e^- \middle| \underbrace{\left(\bar{\psi}_c \not{A}_d \psi_c\right)_{x_1}} \left(\not{A}_c\right)_{x_2} K_2 \middle| 0 \right\rangle \ , \tag{1-2}$$

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 $\not{A}_c$, creates a virtual photon at x_2 that then propagates from x_2 to x_1, where it is annihilated by $\not{A}_d$. 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.

Propagator action leaves only another numeric factor

$$\text{transition amplitude} = {}_{x_1}\left\langle e^+ e^- \middle| \left(\bar{\psi}_c \psi_c\right)_{x_1} K_\gamma K_2 \middle| 0 \right\rangle \tag{1-3}$$

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 $|e^+e^-\rangle_{x1}$ times a numeric 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

Creation operators leave final state ket plus one more numeric factor

$$\text{transition amplitude} = {}_{x_1}\langle e^+e^-|K_1K_\gamma K_2|e^+e^-\rangle_{x_1} = {}_{x_1}\langle e^+e^-|\underbrace{S_{Bhabha}}_{\substack{\text{just a number}\\ \text{without operators}}}|e^+e^-\rangle_{x_1}$$
$$= S_{Bhabha}\underbrace{{}_{x_1}\langle e^+e^-||e^+e^-\rangle_{x_1}}_{1} = S_{Bhabha}, \quad (1\text{-}4)$$

Calculated amplitude

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*. Note, that if we had ended up with a ket different than the bra (final state), the inner product would be zero, because the two (different) states, represented by the bra and ket, would be orthogonal. Examples are

Bracket of multiparticle state = 1 in QFT

$${}_{x_1}\langle e^+e^-||\underbrace{e^+e^+\rangle}_{\substack{\text{two}\\ \text{positrons}}}{}_{x_1} = 0, \quad {}_{x_1}\langle e^+e^-||\underbrace{\mu^-\mu^+\rangle}_{\substack{\text{muon and}\\ \text{anti-muon}}}{}_{x_1} = 0, \quad \text{and} \quad {}_{x_1}\langle e^+e^-||\underbrace{\gamma\rangle}_{\substack{\text{single}\\ \text{photon}}}{}_{x_1} = 0. \quad (1\text{-}5)$$

The whole process of Fig. 1-1 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

Probability = |Amplitude|²

$$\text{probability of interaction} = S^\dagger_{Bhabha}S_{Bhabha} = |S_{Bhabha}|^2. \quad (1\text{-}6)$$

The quantity $K_1K_\gamma K_2 = S_{Bhabha}$ arising in (1-4) depends on particle momenta, spins, and 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. Further, there are other subtleties, including some integration, that have been suppressed in the above in order to convey the essence of the transition amplitude as simply as possible.

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

1.4 From Whence Creation and Destruction Operators?

In NRQM, the solutions to the relevant wave equation, the Schrödinger 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.

QFT wave equation solutions are operators

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 (dynamical variables like position and momentum) that bear a striking resemblance to commutators in quantum theories. For example, the

Poisson brackets behavior parallels quantum commutators

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, led by Paul Dirac, realized that for each pair of quantum operators that had non-zero (zero) commutators, the corresponding pair of classical dynamical variables also had non-zero (zero) Poisson brackets. They had originally arrived at NRQM by taking classical dynamical 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 dynamical 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, and you should understand it better.)

Quantization: Poisson brackets become commutators

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.

1.5.2 First vs. Second Quantization

Branches of classical mechanics

Classical mechanics has both a non-relativistic and a relativistic side, and each contains a theory of particles (localized entities, typically point-like objects) and a theory of fields (entities extended over space). All of these are represented in the first row of Wholeness Chart 1-1. Properties (dynamical 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 and charge density, or field amplitude at a point, etc. that generally vary from point to point. Poisson brackets in field theories are similar to those for particle theories, except they entail *densities* of the respective dynamical variables, instead of total values.

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

With the success of quantization in NRQM, people soon thought of applying it to relativistic particle theory and found they could deduce RQM in the same way. Shortly thereafter they tried applying it to relativistic field theory, 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.

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.,

$$\{X_i, p_j\} = \delta_{ij} \;\Rightarrow\; [X_i, p_j] = X_i p_j - p_j X_i = i\hbar\delta_{ij} \;. \tag{1-7}$$

In doing (1-7), the classical properties (dynamical 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.

$$\{\phi_r(\mathbf{x},t), \pi_s(\mathbf{y},t)\} = \delta_{rs}\delta(\mathbf{x}-\mathbf{y}) \;\Rightarrow\; [\phi_r(\mathbf{x},t), \pi_s(\mathbf{y},t)] = i\hbar\delta_{rs}\delta(\mathbf{x}-\mathbf{y}), \tag{1-8}$$

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 (1-8), the classical field dynamical 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 momentum* of X. (It is sometimes also called the *conjugate momentum*, as we did above, or the *generalized momentum* of X.)

Our approach here: canonical quantization

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. 18 near 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 left hand column of the chart, we employ the second of these.

Note that because quantum field applications usually involve photons or other relativistic particles, non-relativistic quantum field theory (NRQFT) is not widely applicable and thus rarely taught, at least not at elementary levels. However, in some areas where non-relativistic approximations can suffice, such as condensed matter physics, NRQFT can be useful because calculations are simpler. The term "quantum field theory" (QFT) as used in the physics community generally means "relativistic QFT", and our study in this book is restricted to that.

Wholeness Chart 1-1. The Overall Structure of Physics

	Non-relativistic		**Relativistic**	
	Particle	**Field**	**Particle**	**Field**
Classical mechanics (non-quantum)	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 (Dynamical variables) ⇓ Operators	⇓ 1st quantization ⇓	⇓ 2nd quantization ⇓	⇓ 1st quantization ⇓	⇓ 2nd quantization ⇓
Quantum mechanics	NRQM	NRQFT rarely taught.	RQM	QFT (not gravity)

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.

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) Schrödinger equation, whereas RQM and QFT must employ relativistic counterparts sometimes called relativistic Schrödinger equations. Students of QFT soon learn that each spin type (spin 0, spin ½, and spin 1) has a different relativistic Schrödinger 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

Solutions: RQM→ states QFT→ operators

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 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 in 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.,

$$\bar{E} = \langle \phi | H | \phi \rangle , \tag{1-9}$$

with similar relations for other observables.

These similarities and differences, as well as others, are summarized in Wholeness Chart 1-2. 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 Schrödinger 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 (1-8).

Phenomena in the 3 theories

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), deducing particular experimental results, and vacuum energy.

QFT rarely uses spherical solutions

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.

or wave packets

Wave packets, as well, can seem nowhere 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 generally ignored in introductory (and most advanced) courses.

QFT handles various type interactions

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

The final blocks note the similarities and differences between forces (interactions) in the different theories. As in classical theory, in all three quantum theories, interactions comprise forces that change the momentum and energy of particles. However, in QFT alone, interactions can also involve changes in type of particle, such as shown in Fig. 1-1. At event x_2, the electron and positron are changed into a photon, and in the process energy and momentum is transferred to the photon .

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.

Wholeness Chart 1-2. Comparison of Three Theories

	NRQM	**RQM**	**QFT**
Wave equation	Schrödinger	Klein-Gordon (spin 0) Dirac (spin ½) 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?	No	Yes	No
Particles per state	Single*	Single*	Multi-particle
Expectation values	$\bar{\mathcal{O}} = \langle \phi \vert \mathcal{O} \vert \phi \rangle$	As at left, but relativistic.	As at left in RQM.
Phenomena:			
1. bound states	Yes, non-relativistic	Yes, relativistic	Yes (usually not studied in introductory courses)
2. scattering			
a. elastic	a. Yes	a. Yes	a. Yes
b. inelastic (transmutation)	b. No (though some models can estimate)	b. Yes and no. (i.e., cumbersome and 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 energy	No	No	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.
Interaction types			
1. e/m	No, though can pseudo model	As at left	Yes
2. weak	No	No	Yes
3. strong	No*	No*	Yes
4. gravity	No	No	Not as of this edition date.

Interaction nature			
Transfers energy & momentum?	Yes	Yes	Yes
Can change particle type?	No	No	Yes

*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

The four major parts 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 Schrödinger 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). A trick employed in interaction theory actually lets us use 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. Application to experiment

   The theory of parts 1, 2, and 3 above are put to practical use in determining interaction probabilities and from them, scattering cross sections, decay probabilities (half lives, etc.), and certain other experimental results. 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

Terminology "field" = operator in QFT

When the word "field" is used classically, it refers to an entity, like fluid wave amplitude, **E**, or **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.

Symbol e >0

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 $-e$.

1.9 Big Picture of Our Goal

Our goal: predict scattering and decay seen in Nature

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, certain other experimental results, and elementary particle half lives. To do these things, we need to be able to calculate probabilities for each 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 formulate transition amplitudes, also in Part 2. Necessary refinements will take up Part 3, with experimental application in Part 4.

Steps to our goal

2^{nd} quantization postulates → QFT theory → transition amplitude calculation → probability → scattering, decay, other experimental results → confirmation of QFT

Steps to our goal

In this book our goal is a bit limited, as we will examine a part – an essential part – of the big picture. We will i) develop the fundamental principles of QFT, ii) use those principles to derive quantum electrodynamics (QED), the theory of electromagnetic quantum interactions, and iii) apply the theory of QED to electromagnetic scattering and other experiments. We will not examine herein the more advanced theories of weak and strong interactions, which play essential roles in particle decay, most present day high energy particle accelerator experiments, and composite particle (e.g., proton) structure. Weak and strong interaction theories build upon the foundation laid by QED. First things first.

Our goal in this book: basic QFT principles and QED, theory and experiments

1.10 Summary of the Chapter

Throughout 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) though (1-6), Sect.1.5.2 (1^{st} and 2^{nd} Quantization), Wholeness Chart 1-1 (The Overall Structure of Physics), Wholeness Chart 1-2 (Comparison of Three Theories), and Sect. 1.9 (Big Picture of Our Goal).

1.11 Suggestions?

If you have suggestions to make the material anywhere in this book easier to learn, or if you find any errors, please let me know via the web site address for this book posted on pg. xvi (opposite pg.1). Thank you.

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. Note that your destruction operators must be different than the example in the chapter in that they now destroy a muon and antimuon instead of an electron and positron.

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 dynamical 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 *book vers 3/6/13 copyright of Robert D. Klauber*

Foundations

Tiger got to hunt. Bird got to fly.
Man got to ask himself "why, why, why?".
Tiger got to rest. Bird got to land.
Man got to tell himself he understand.

The Book of Bonkonon 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 that takes the vacuum permittivity ε_o and permeability μ_o values as unity) has been common in NRQM, although standard international units (SI) [essentially, MKS for electromagnetism] are also used. Another is the Heaviside-Lorentz system, which is similar to the Gaussian system except it is structured to eliminate factors of 4π found in the Gaussian form of Maxwell's equations. (See Chap. 5.)

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.

Natural units are "natural" and used in QFT

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.

Any system of units: defined units + laws of nature → additional derived 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 m 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 h. $E = hf$ 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.

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/s.

cgs: cm, g, s defined. Other units derived from laws of nature

From law number two, the dimensions of energy are mass-length^2/time^2 and the units are g-cm^2/s^2. We use shorthand by calling this an erg.

From number three, $\hbar$ has dimensions of energy-time and units of g-cm^2/s, or for short, erg-s. It, like the speed of light, can be measured by experiment and is found to be 1.0545×10^{-27} erg-s.

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**=*m***a** 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.)

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

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 was.) As with everything else, we do this because it will turn out to be advantageous.

Energy in natural units: electron volts (MeV convenient)

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 s^{-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

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

Hybrid units used in experiments: cm, s, MeV

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.

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} $(\text{MeV})^{-1}$. What is its measurable value in seconds? To find out, observe that

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

But the above relation can be expressed in terms of the hybrid MeV-cm-s 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-s. Then,

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

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

The same time interval is described as either 10^{16} $(\text{MeV})^{-1}$ 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-s units) as many times as is necessary to get the units we know that quantity should have in the MeV-cm-s system.

2.1.6 Mass and Energy in the Hybrid and Natural Systems

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

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.

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

Natural Units		**Hybrid Units**		**cgs Units**	
$c = \hbar = 1$		$c = 2.99 \times 10^{10}$ cm/s $\hbar = 6.58 \times 10^{-22}$ MeV-s $\hbar c = 1.973 \times 10^{-11}$ MeV-cm		conversion factor $F = 1.602 \times 10^{-6}$ ergs/MeV	
Quantity, units of (MeV)^M	***M***	**Multiply ← value by ↓ to get →**	**in MeV-cm-s**	**Multiply ← value by ↓ to get →**	**in cgs**
energy	1	1	MeV	*F*	ergs
mass, *m*	1	$1/c^2$	MeV-s^2/cm^2	*F*	erg-s^2/cm^2 = gs
length	-1	$\hbar c$	cm	1	cm
time	-1	$\hbar$	s	1	s
velocity	0	*c*	cm/s	1	cm/s
acceleration, **a**	1	$c/\hbar$	cm/s^2	1	cm/s^2
force	2	*m***a** factors = $1/c\hbar$	MeV/cm	*F*	ergs/cm = dynes
$\hbar$ (= 1)	0	$\hbar$	MeV-s	*F*	erg-s
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-s	*F*	erg-s
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,s,g fundamental, other quantities derived from laws of nature

hybrid: cm,s,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(energy units) = m(energy-s^2/cm^2 units) × ?, 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-s) 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 (which do occur with monotonous regularity.) You may have a point on that. More importantly, the essential mathematical structure of the resulting equations, and the fundamentals of the underlying physics, is more clearly seen without the clutter of relatively unimportant unit scaling factors.

Regardless, 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 defining contravariant components x^μ of the 4D position vector as 3D Cartesian coordinates X_i (see the appendix if you are not comfortable with this), i.e.,

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

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

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

$$(d\tau)^2 = (dt)^2 - dX_1 dX_1 - dX_2 dX_2 - dX_3 dX_3 . \tag{2-2}$$

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

Covariant components have negative 3D Cartesian coordinates

$$x_\mu = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} t \\ -X_1 \\ -X_2 \\ -X_3 \end{bmatrix} = \left[t, -X_i\right]^T , \tag{2-3}$$

then (2-2) becomes

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

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., $dx^{\underline{\mu}} dx_{\underline{\mu}}$ means no summation.

Repeated indices means summation

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

Getting covariant components from contravariant ones

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

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

Contravariant and covariant forms of the metric tensor

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

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

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

Contravariant and covariant derivatives

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

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

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

Raising and lowering indices

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 $w_\mu = g_{\mu\nu}w^\nu$.

Script → density

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.

$p_\mu x^\mu = px$

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

Fig. 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, constant phase angle) propagating through space. Theoretically, the planes extend to infinity in the *y* and *z* directions. The lower parts of Fig. 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.)

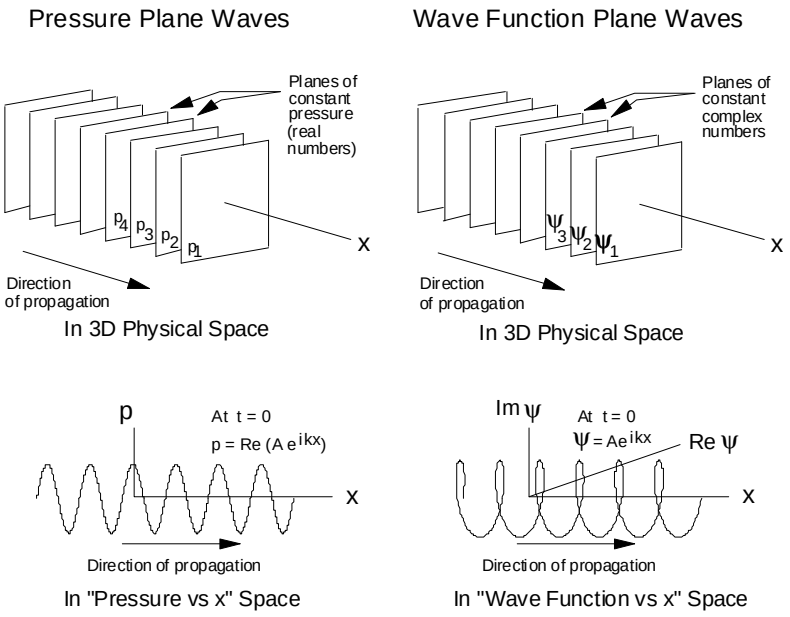


Figure 2-1. Classical vs Quantum Plane Waves

2.4 Review of Variational Methods

2.4.1 Classical Particle Theory

Definition of classical mechanics Lagrangian L

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

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

Governing equation = Euler-Lagrange equation

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

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

This, with (2-9), readily reduces to Newton's 2nd law (with conservative force), $F^i = -\partial V/\partial x^i = m\ddot{x}^i$.

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.

Legendre transformation $H \leftrightarrow L$

Recall also, that given the Lagrangian, we could find the Hamiltonian *H*, via the Legendre transformation (employing a Cartesian system where $x^i = x_i$ and $p^i = p_i$ - see Prob. 8),

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

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.)

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

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

Lagrangian *L* ↔ equations of motion ↔ 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. (Note that Poisson brackets are discussed on pg. 24 and summarized in Wholeness Chart 2-2 on pgs. 20 and 21.)

2.4.2 Pure Mathematics

Variational math can be applied to many diverse fields

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 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 **E** or **B** in electromagnetism).]

Analogous entities in particle and field theories

Particle Theory → Field Theory

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

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. Thus, the Lagrangian density, in terms of kinetic energy density and potential energy densities of the field, is

$$\mathcal{L} = \mathcal{T} - \mathcal{V} \,. \tag{2-12}$$

Intuitive deduction of field relations from particle ones

(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

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

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

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

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

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

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

Formal 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,

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

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

$$\phi\left(x^\mu\right) \rightarrow \phi\left(x^\mu\right) + \delta\phi\left(x^\mu\right), \tag{2-16}$$

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

$$\delta S = 0 \,. \tag{2-17}$$

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

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

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

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

we can express (2-18) as

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

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),

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

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

Classical field real; quantum fields usually complex

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 **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.

2.5 Classical Mechanics: An Overview

Variational classical mechanics overview in Wholeness Chart 2-2

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

Wholeness Chart 2-2.

	Mathematically	**Non-relativistic Particle**
Independent variable(s)	t	t
Coordinates	$q_i = q_i(t)$, $i = 1,.., n$ (generalized)	$x^i = x^i(t)$, $i = 1, 2, 3$ (contravariant)
Lagrangian density	see Fields columns	not applicable for particle
Lagrangian	$L = L\left(q_i, \dot{q}_i, t\right)$	$L = L\left(x^i, \dot{x}^i, t\right) = \sum_i \frac{1}{2} m\left(\dot{x}^i\right)^2 - V\left(x^i, t\right)$
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 ; $p_i = \frac{\partial L}{\partial \dot{x}^i} = m\dot{x}^i$ $\left(= p^i \text{ for Cartesian}\right)$
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) any variable ii) conjugate variables	i) for $v = H$ $\frac{du}{dt} = \{u,H\} + \frac{\partial u}{\partial t}$ ii) for i) plus $u = q_i$ or p_i $\dot{p}_i = \{p_i,H\} = -\frac{\partial H}{\partial q_i}$; $\dot{q}_i = \{q_i,H\} = \frac{\partial H}{\partial p_i}$	i) for $v = H$ $\frac{du}{dt} = \{u,H\} + \frac{\partial u}{\partial t}$ ii) for i) plus $u = x^i$ or p_i $\dot{p}_i = \{p_i,H\} = -\frac{\partial H}{\partial x^i}$; $\dot{x}^i = \{x^i,H\} = \frac{\partial H}{\partial p_i}$
Poisson Brackets for conjugate variables	$\{q_i, p_j\} = \delta_{ij}$ $\{q_i, q_j\} = \{p_i, p_j\} = 0$	$\{x^i, p_j\} = \delta^i{}_j$ $\{x^i, x^j\} = \{p_i, p_j\} = 0$

Summary of Classical (Variational) Mechanics

Non-relativistic Fields	Relativistic Particle	Relativistic Fields
x^i, t $\quad i = 1, 2, 3$	t	$x^\mu \quad \mu = 0, 1, 2, 3$
$\phi^r(x^i, t)$ $\quad r$ = field type = 1, …, n	$x^i = x^i(t)$, $\quad i = 1, 2, 3$	$\phi^r(x^\mu)$ $\quad r$ = field type = 1, …, n
$\mathcal{L} = \mathcal{L}\left(\phi^r, \dot{\phi}^r, \partial_i \phi^r, x^i, t\right)$	not applicable for particle	$\mathcal{L} = \mathcal{L}\left(\phi^r, \partial_\mu \phi^r, x^\mu\right)$
$L = \int \mathcal{L} d^3x$	$L\left(x^i, v^i, t\right) = -m\sqrt{1-v^2} - V$	$L = \int \mathcal{L} d^3x$
$S = \int L dt = \int \mathcal{L} d^3x dt$	$S = \int L dt$	$S = \int L dt = \int \mathcal{L} d^3x dt$
$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\phi}^r}\right) + \frac{d}{dx^i}\left(\frac{\partial \mathcal{L}}{\partial \phi^r{}_{,i}}\right) - \frac{\partial \mathcal{L}}{\partial \phi^r} = 0$	$\frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\right) - \frac{\partial L}{\partial x^i} = 0$	$\frac{\partial}{\partial x^\mu}\left(\frac{\partial \mathcal{L}}{\partial \phi^r{}_{,\mu}}\right) - \frac{\partial \mathcal{L}}{\partial \phi^r} = 0$
$\mathcal{L}$ above in Euler-Lagrange equation	$\frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\right) = -\frac{\partial V}{\partial x^i}$; $V\left(x^i, v^i\right)$	$\mathcal{L}$ 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 ; $p_i = \frac{\partial L}{\partial v^i} = \frac{m v^i}{\sqrt{1-v^2}} - \frac{dV}{dv^i}$	$\pi_r = \frac{\partial \mathcal{L}}{\partial \dot{\phi}^r}$; $\quad \Pi_r = \int \pi_r d^3x$
$\not{\varepsilon}^i = \pi_r \frac{\partial \phi^r}{\partial x^i}$; $\quad p^i = \int \not{\varepsilon}^i d^3x$	n/a ; = conjugate momentum	$\not{\varepsilon}^i = \pi_r \frac{\partial \phi^r}{\partial x^i}$; $\quad p^i = \int \not{\varepsilon}^i d^3x$
$\mathcal{L} = \mathcal{L}\left(\phi^r, \pi_r, \partial_i \phi^r, x^i, t\right)$	$L = L(x^i, p_i, t)$	$\mathcal{L} = \mathcal{L}\left(\phi^r, \pi_r, \partial_i \phi^r, x^i, t\right)$
$\mathcal{H} = \pi_r \dot{\phi}^r - \mathcal{L}$; $\quad H = \int \mathcal{H} d^3x$	n/a ; $\quad H = p_i v^i - L = T + V$	$\mathcal{H} = \pi_r \dot{\phi}^r - \mathcal{L}$; $\quad H = \int \mathcal{H} d^3x$
same form as Relativistic Fields	$\dot{p}_i = -\frac{\partial H}{\partial x^i} = -\frac{\partial V}{\partial x^i} \qquad \dot{x}^i = \frac{\partial H}{\partial p_i}$	$\dot{\pi}_r = -\frac{\delta \mathcal{H}}{\delta \phi^r} \qquad \dot{\phi}^r = \frac{\delta \mathcal{H}}{\delta \pi_r}$ where $\frac{\delta}{\delta \phi^r} = \frac{\partial}{\partial \phi^r} - \frac{\partial}{\partial x^i}\left(\frac{\partial}{\partial \phi^r{}_{,i}}\right)$
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	$\{\phi^r, \pi_s\} = \delta^r{}_s \delta(\mathbf{x} - \mathbf{y}); \{\phi^r, \phi^s\} = \{\pi_r, \pi_s\} = 0$

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.

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

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

1. Generalized coordinates do not have to be independent of each other, and the Lagrangian L can have second and/or higher coordinate derivatives. However, in most cases, including those of Wholeness Chart 2-2, the coordinates are independent and L only contains first derivatives.

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

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.

   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.

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.

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

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.

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

5. For fields,

$$\frac{\partial \phi}{\partial t} = \frac{d\phi}{dt} = \dot{\phi} \tag{2-22}$$

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

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

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

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. 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. 2-2,

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

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

$$\mathcal{H} = \not{p}^i v^i - \mathcal{L} ,$$

where $\not{p}^i$ 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 **E** field, this would be the velocity at which one would travel in order to see a constant **E** value.

Equating the above two equations, we see

$$\not{p}^i v^i = \pi_r \dot{\phi}^r \quad \rightarrow \quad \not{p}^i \frac{\partial x^i}{\partial t} = \pi_r \frac{\partial \phi^r}{\partial t} ,$$

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. (It is different from the same x^i symbol we use in field theory, which is an independent variable that does not depend on time.) 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

$$\not{p}^i = \pi_r \frac{\partial \phi^r}{\partial x^i} .$$

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.

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. (See the Appendix if you do not feel comfortable with the material discussed in this paragraph.) The relativistic particle summary, as outlined in Wholeness Chart 2-2. 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.

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

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 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

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

Poisson bracket definition used in equation of motion

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

$$\frac{du}{dt} = \underbrace{\frac{\partial u}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial u}{\partial p_i}\frac{\partial H}{\partial q_i}}_{\substack{\text{Possion bracket} \\ \text{definition for u and H}}} + \frac{\partial u}{\partial t} = \{u, H\} + \frac{\partial u}{\partial t}, \tag{2-24}$$

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

Forms for 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

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. (See the Appendix Sects. 2.9.3 and 2.9.4, if you do not feel at home with the concepts of this paragraph.) 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* [or simply mass *m* as it is more commonly called in relativity] of a free

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

* 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.

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 mass via $p'^\mu p'_\mu = m^2$.

$\mathcal{L}$ is a Lorentz invariant scalar

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.

L, H, and $\mathcal{H}$ are not Lorentz scalars

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. 9.

End of Key Concepts in Field Theory points

2.6 Schrödinger 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 dynamical variables remain the same.

Different pictures in quantum theory

These different, but equivalent, ways are called different pictures and apply in the same way to all branches of quantum theory (NRQM, RQM, QFT.) Most QM courses more elementary than this one use what is known as the Schrödinger 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 Schrödinger 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.,

Schrödinger Wave Approach	Heisenberg Matrix Approach
1. Schrödinger picture	
2. Heisenberg picture.	

We will review the Schrödinger 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 Schrödinger Picture

Operator expectation value = "expected" or mean measurement

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)

Calculating expectation value

$$\bar{\mathcal{O}} = \int \psi^\dagger \mathcal{O} \psi \, d^3x = \langle \psi | \mathcal{O} | \psi \rangle . \tag{2-25}$$

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 dynamical variable) is

Eq of motion of expectation value

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

In the Schrödinger picture, the solutions to the Schrödinger equation

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

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

are the states ψ_S (or $|\psi\rangle_S$), which are time dependent. The subscript S indicates the Schrödinger 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

An example

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

(2-25) is

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

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

In S.P., NRQM eq of motion of momentum operator (i.e., $p_1{}^S$ *constant in time*)

$$\frac{dp_1^S}{dt} = \frac{\partial p_1^S}{\partial t} = 0\,. \tag{2-30}$$

Equation (2-26) for p_1 is then

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

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 Schrödinger equation (2-27) and its complex conjugate for the ket and bra time derivatives, respectively, in (2-31), we get

Eq of motion of momentum expectation value

$$\frac{d\overline{p}_1}{dt} = {}_S\langle\psi|\, iHp_1^S + \frac{\partial p_1^S}{\partial t} - i\, p_1^S H\,|\psi\rangle_S = {}_S\langle\psi|-i\left[\, p_1^S, H\,\right]|\psi\rangle_S + {}_S\langle\psi|\underbrace{\frac{\partial p_1^S}{\partial t}}_{=0}|\psi\rangle_S\,. \tag{2-32}$$

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 Schrödinger picture. These are generalized to any state and operator in Wholeness Chart 2-4.

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

Note further that the partial time derivative $\partial/\partial t$ in the Schrödinger 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 Schrödinger equation) with either a partial or total time derivative.

2.6.2 The Heisenberg Picture

The Schrödinger 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 (where U is a unitary operator) for this is

Transforming between Schrödinger and Heisenberg pictures

$$U = e^{-iHt/\hbar}\,, \tag{2-33}$$

where states and operators transform as

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

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

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

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,

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

$$\frac{d|\psi\rangle_H}{dt} = 0\,. \tag{2-36}$$

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

$$\frac{d}{dt}\left(U^{\dagger}\mathcal{O}^{S}U\right)=(iH)\underbrace{e^{iHt/\hbar}\mathcal{O}^{S}e^{-iHt/\hbar}}_{\mathcal{O}^{H}}+\underbrace{e^{iHt/\hbar}\left(\frac{\partial\mathcal{O}^{S}}{\partial t}\right)e^{-iHt/\hbar}}_{\text{defined as }\partial\mathcal{O}^{H}/\partial t}+\underbrace{e^{iHt/\hbar}\mathcal{O}^{S}e^{-iHt/\hbar}}_{\mathcal{O}^{H}}(-iH)$$

$$=\frac{d\mathcal{O}^{H}}{dt}=-i\left[\mathcal{O}^{H},H\right]+\underbrace{\frac{\partial\mathcal{O}^{H}}{\partial t}}_{=0\text{ in this book}}\quad . \tag{2-37}$$

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 in (2-37) 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.

Box 2-3. Unitary Transformations in Quantum Theories

A *unitary transformation* is called unitary because its operation on (transformation of) a state vector leaves the magnitude of the state vector unchanged, i.e., the state vector magnitude is multiplied by unity. 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" (magnitude) 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. A unitary transformation multiplies probability by unity.

Recall, from classical mechanics, that an orthogonal transformation represented by a real matrix **A** has an inverse equal to the transpose of that matrix, i.e., $\mathbf{A}^{-1} = \mathbf{A}^{\mathrm{T}}$. 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 so $U^{\dagger}U = 1$. The following example may make this clearer.

Consider $U = e^{-iHt}$, where H is the Hamiltonian operator. By inspection, one knows its magnitude in complex space is unity and so its action on a state vector would not change the length of that state vector (though phase would change by $-Ht$.) Also by inspection, $U^{\dagger}U = 1$. So, U performs a unitary transformation.

Wholeness Chart 2-3. Unitary vs Orthogonal Transformations

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

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.,

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

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 is readily 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.

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),

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

$$H = H^S = H^H \,. \tag{2-38}$$

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

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

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 (dynamical 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. Schrödinger vs. Heisenberg Picture Equations of Motion

	States	**Operators**	**Expectation Values**
Schrödinger Picture	Time dependent $i\frac{d}{dt}\|\psi\rangle_S = H\|\psi\rangle_S$ (Schrödinger eq)	Usually time independent $\frac{d\mathcal{O}^S}{dt} = \frac{\partial\mathcal{O}^S}{\partial t} \underbrace{=0}_{\text{usually}}$	$\frac{d\bar{\mathcal{O}}}{dt} = {}_S\langle\psi|-i\left[\mathcal{O}^S, H\right]+\frac{\partial\mathcal{O}^S}{\partial t}|\psi\rangle_S$ $\|\psi\rangle_S$ 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 $\frac{d\mathcal{O}^H}{dt} = -i\left[\mathcal{O}^H, H\right] + \underbrace{\frac{\partial\mathcal{O}^H}{\partial t}}_{\substack{\text{usually}\\=0}}$	Same as Schrödinger picture above with sub and superscript $S \to H$ $\|\psi\rangle_H$ 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.

Continuation of Wholeness Chart 1-2. Comparison of Three Quantum Theories

	NRQM	**RQM**	**QFT**
Most advantageous picture to use	Schrödinger picture	Schrödinger picture	Heisenberg picture

2.6.3 Visualizing Schrödinger and Heisenberg Pictures

One can think of the S.P. as quantum waves (wave functions, states, 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.

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

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

The philosophical lesson to be learned from this is that we can have different models of reality predicting the same real world phenomena. In this case, in one model the states are waves that move and evolve. In the other model, the states never change. But, both are valid predictors of the laws of nature we observe in the physical universe. Hence, we should be wary of accepting any given model of reality as a "true" picture of what nature is actually doing.

2.7 Quantum Theory: An Overview

Chart 2-5 *summarizes QM*

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.)

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

Note particularly, that in Wholeness Chart 2-5, all relations and quantities are expressed in the Heisenberg picture. If it were expressed in the Schrödinger 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 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 Schrödinger 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

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

Note that everything in the first 12 blocks in the NRQM and RQM columns 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 dynamical variable, represented by u, changes from (2-24) in classical non-relativistic particle theory to

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

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.)

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

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

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 dynamical 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		$[x^i, p_j] = i\hbar\delta^i{}_j$ $[x^i, x^j] = [p_i, p_j] = 0$

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

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

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

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

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 first twelve rows for Wholeness Chart 2-5 as we had in Wholeness Chart 2-2 and the same sort of bracket correspondence for the other rows as in NRQM/RQM, 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.,

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

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

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

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 $\dot{\pi}_r = \frac{-i}{\hbar}[\pi_r, H] = \frac{\delta H}{\delta \phi^r}$; $\dot{\phi}^r = \frac{-i}{\hbar}[\phi^r, H] = \frac{\delta H}{\delta \pi_r}$
	See Non-relativistic Quantum Mechanics section	$\left[\phi^r, \pi_s\right] = i\hbar \delta^r{}_s \delta(\mathbf{x} - \mathbf{y}); \left[\phi^r, \phi^s\right] = [\pi_r, \pi_s] = 0$

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.

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

2.8 Chapter Summary

The bottom right hand block of Wholeness Chart 2-5, Summary of Quantum Mechanics, contains the essence of this chapter (enclosed in box with bold border). 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" (magnitude) 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 dynamical 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 Appendix: Understanding Contravariant and Covariant Components

The concepts of contravariant and covariant components presented in Sect. 2.2 should be somewhat familiar to those who have studied the prerequisite material delineated in the Preface. However, oftentimes, even those who have already been exposed to these concepts still do not feel completely at home with them. For them, and for any newcomers to the subject, I hope the following brief introduction will help.

2.9.1 A Trick for Conveniently Finding 4D Vector Length

Contravariant and covariant components are simply tricks that allow us to represent vectors (and tensors) in a way that helps us carry out certain mathematical procedures, like finding the magnitude of a vector in curved space or the proper time passing on a particle in special relativity. In this book, we will not be dealing with curved space, so all of the applications of contravariant and covariant component theory herein will be for the simpler case of Minkowski space (flat, 4D space with Cartesian space coordinates plus time.) We will, for starters, want to be able to calculate proper time on a particle (decay time of a particle, for instance, depends on proper time, not the lab time we see as the particle whizzes by.)

Consider how we find the length l of a vector in a 3D Cartesian system with one end of the vector at the origin, i.e.,

$$(l)^2 = (X_1)^2 + (X_2)^2 + (X_3)^2 = X_i X_i \quad \left(= \sum_i X_i X_i \text{, repeated indices mean summation. See Sect 2.2} \right)$$

$$= \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} X_1 & X_2 & X_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = X_i \underbrace{\delta_{ij} X_j}_{\sum_j \delta_{ij} X_j = X_i} \tag{2-45}$$

where, with a future purpose in mind, we insert an identity matrix, represented in index notation by the Kronecker delta δ_{ij} (= 0 if row $i \neq$ column j; = 1 if $i = j$), on the RHS.

Now, imagine a spatially 4D Cartesian system, where the length of a 4D vector is

$$(l)^2 = (X_0)^2 + (X_1)^2 + (X_2)^2 + (X_3)^2 = X_\mu X_\mu$$

$$= \begin{bmatrix} X_0 & X_1 & X_2 & X_3 \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} X_0 & X_1 & X_2 & X_3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ X_3 \end{bmatrix} = X_\mu \underbrace{\delta_{\mu\nu} X_\nu}_{X_\mu} \tag{2-46}$$

Now consider the 4D spacetime of special relativity theory (SRT), and the "length" of a 4D vector we have in mind is the proper time τ on an object passing by us. The 0th coordinate is now time instead of a spatial X_0 coordinate. From SRT, we know

$$(c\tau)^2 = (ct)^2 - (X_1)^2 - (X_2)^2 - (X_3)^2 = \text{ (how to write as summed indices?)}$$

$$= \begin{bmatrix} ct & X_1 & X_2 & X_3 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} ct \\ -X_1 \\ -X_2 \\ -X_3 \end{bmatrix} \end{bmatrix} \qquad c = 1 \text{ in natural units} \tag{2-47}$$

Note that because of the minus signs in our "length" (= proper time) calculation in (2-47), we can't use the nice summation symbolism of the first lines of (2-45) and (2-46). That was only good if all of the terms in the summation had the same sign. Fine for purely spatial coordinates of any dimension. Not possible if we have both time and space in the same coordinate system.

But here is a clever idea. Let's define the column matrix of the second line in (2-47) as a different set of vector components, with minus signs in front of the X_i. We could designate it with primes, if we like, so

$$X'_\mu = \begin{bmatrix} ct \\ -X_1 \\ -X_2 \\ -X_3 \end{bmatrix} = \begin{bmatrix} X_0 \\ -X_1 \\ -X_2 \\ -X_3 \end{bmatrix} \quad \text{and} \quad X_\mu = \begin{bmatrix} ct \\ X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ X_3 \end{bmatrix}. \tag{2-48}$$

With this newly defined representation of our 4D vector, and $ct = X_0$, we can represent our vector "length" of (2-47) as

$$(c\tau)^2 = (ct)^2 - (X_1)^2 - (X_2)^2 - (X_3)^2 = (X_0)^2 - (X_1)^2 - (X_2)^2 - (X_3)^2 = X_\mu X'_\mu \,. \tag{2-49}$$

And thus, we have a neat shorthand way to write out a vector length in 4D spacetime.

Unfortunately, the primed notation is used in relativity and elsewhere to indicate a different coordinate system in a different frame. In relativity, this is usually a frame having velocity relative to the unprimed frame. In the present case, we are only working in a single coordinate system. So, a different symbolism has arisen for this case (i.e., for finding vector lengths in the same coordinate system). While it can take a little getting used to, the symbolism entails using no primes, but instead raising the indices for one of the component sets in (2-49), and keeping the indices lowered for the other. We also generally use non-capital letters for 4D position vectors, and capital letters (with subscript indices only) for 3D Cartesian components. Thus, by the convention chosen,

$$x^\mu = \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} = X_\mu = \begin{bmatrix} ct \\ X_1 \\ X_2 \\ X_3 \end{bmatrix} \quad \text{and} \quad x_\mu = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = X'_\mu = \begin{bmatrix} ct \\ -X_1 \\ -X_2 \\ -X_3 \end{bmatrix}. \tag{2-50}$$

With the above convention, our 4D vector length (2-49) becomes

$$(c\tau)^2 = x^0 x_0 + x^1 x_1 + x^2 x_2 + x^3 x_3 = x^\mu x_\mu \,. \tag{2-51}$$

Of course, this can lead to some confusion, as before this, we have always used a superscript solely for raising a quantity to a power. To avoid this confusion, we will have to remember to enclose entities in parentheses when we mean the superscript as a power, as we did on the LHS of (2-51). From now on, superscripts without parentheses will designate components, not powers. Be forewarned, however, that, unfortunately, authors may not always strictly adhere to this practice, and you may have to glean the meaning of a superscript from context. (This isn't so hard *after* you get accustomed to this notation, but it can be difficult *before* you do.)

For reasons beyond the scope of this discussion, $\underline{x^\mu}$ was designated as the contravariant components form, and $\underline{x_\mu}$ as the covariant components form, of the same physical vector. As a mnemonic, just remember that the raised index contravariant components are the 3D Cartesian coordinates plus *ct*. The lowered index covariant components include a minus sign for the 3D part.

Contravariant and covariant components also allow us to readily find the 4D length of any vector, not just the 4D position vector x^μ. For example, the four-velocity of relativity u^μ for an object is

$$u^\mu = \frac{dx^\mu}{d\tau} = \frac{d}{d\tau}\begin{bmatrix} x^0 & x^1 & x^2 & x^3 \end{bmatrix} = \begin{bmatrix} u^0 & u^1 & u^2 & u^3 \end{bmatrix}, \tag{2-52}$$

where

$$u^i = \frac{dx^i}{d\tau} = \frac{dx^i}{\sqrt{1 - v^2/c^2}\,dt} = \frac{v^i}{\sqrt{1 - v^2/c^2}} = \gamma v^i \;; \quad u^0 = \frac{dx^0}{d\tau} = c\frac{dt}{d\tau} = \frac{c}{\sqrt{1 - v^2/c^2}} = \gamma c \,, \tag{2-53}$$

u^i here is the derivative of the spatial coordinate with respect to proper time on the object τ, v^i is that with respect to coordinate time *t*, γ is the usual Lorentz factor common in relativity, and we will henceforth often write vectors as rows, rather than columns, to save space. The 4D length $|u^\mu|$ is found from

$$\left|u^\mu\right|^2 = u^\mu u_\mu = \frac{dx^\mu}{d\tau}\frac{dx_\mu}{d\tau} = \begin{bmatrix} u^0 & u^1 & u^2 & u^3 \end{bmatrix}\begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} u^0 & u^1 & u^2 & u^3 \end{bmatrix}\begin{bmatrix} u^0 \\ -u^1 \\ -u^2 \\ -u^3 \end{bmatrix}$$
$$= \left(u^0\right)^2 - \left(u^1\right)^2 - \left(u^2\right)^2 - \left(u^3\right)^2 = \gamma^2\left(c^2 - \left(v^1\right)^2 - \left(v^2\right)^2 - \left(v^3\right)^2\right) = c^2, \tag{2-54}$$

the last part of which students of relativity may recognize as the correct expression for the square of the magnitude of the four-velocity.

The magnitude of the 4-momentum $p^\mu = mu^\mu$ is then found from

$$\left|p^\mu\right|^2 = p^\mu p_\mu = m^2 u^\mu u_\mu = m^2 c^2 \quad \left(= m^2 \text{ in natural units}\right). \tag{2-55}$$

For any general vector w^μ, with upper case letters representing 3D Cartesian components, we have

$$w^\mu = \begin{bmatrix} w_0 & W_1 & W_2 & W_3 \end{bmatrix} \quad w_\mu = \begin{bmatrix} w_0 & -W_1 & -W_2 & -W_3 \end{bmatrix} \quad \left|w^\mu\right|^2 = w^\mu w_\mu \ . \tag{2-56}$$

In addition, we will often use differential elements of 4 vectors, such as dx^μ, and the relations (2-56) hold for such differential 4 vectors, as well (which should be fairly obvious, as a differential of a vector is also a vector in its own right.)

2.9.2 The Metric

Note that we can use a certain matrix to convert from contraviant to covariant components,

$$x_\mu = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} ct \\ -X_1 \\ -X_2 \\ -X_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} ct \\ X_1 \\ X_2 \\ X_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}}_{g_{\mu\nu}}\underbrace{\begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix}}_{x^\nu} = g_{\mu\nu}x^\nu \ . \tag{2-57}$$

This matrix $g_{\mu\nu}$ represents what is called the metric (of the coordinate space, which in this case is Minkowski coordinate space.) It lowers a raised index. It has an inverse that turns out to have the same form as it does.

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}}_{g_{\mu\nu}}\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}}_{\left(g_{\mu\nu}\right)^{-1} = g^{\nu\alpha}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\delta_\mu{}^\alpha} \ . \tag{2-58}$$

The inverse of the metric can be used to raise indices, i.e.,

$$x^\mu = \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}}_{g^{\mu\nu}}\underbrace{\begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{x_\nu} = g^{\mu\nu}x_\nu \ . \tag{2-59}$$

When indices are repeated, they are summed, and even when they are not, they are only dummy indices symbolizing coordinate axes numbers. So it really doesn't matter what particular Greek letter we take for a summed index. Hence, $g^{\nu\alpha}$ represents the same entity as $g^{\mu\nu}$.

$g_{\mu\nu}$ is sometimes called the covariant metric, and $g^{\mu\nu}$, the contravariant metric. The term *metric* used alone usually means $g_{\mu\nu}$.

Note that with the metric, we can write (2-51) as

$$(c\tau)^2 = x^\mu x_\mu = g_{\mu\nu} x^\mu x^\nu . \tag{2-60}$$

Prove (2-60) to yourself three ways: by substituting the RHS of (2-57) into the middle part of the above, by writing out (2-60) in matrix form, and by doing the summation of terms implied by the repeated indices.

Note that the particular metric form of the metric in (2-57) is specific to Minkowski coordinates, which is all we will use in this book. Other coordinate systems (like 4D having time and a spherical spatial coordinate system) would have other forms for $g_{\mu\nu}$.

The metric in (2-60) plays a role in 4D spacetime similar to the role played by the identity matrix of (2-45) and (2-46) for Cartesian spaces (which are purely spatial, with no time axis.) In fact, for Cartesian systems, the identity matrix *is* the metric. (Do Prob. 8 for more on this.)

The form of the metric tells us a lot, in fact virtually everything, about the coordinate space we are dealing with. It is, in a sense, the *signature* of the coordinate space.

2.9.3 Invariance and Covariance

The quantity $c\tau$ of (2-60) is an example of what is known as a 4D scalar (or world scalar or Lorentz scalar.) It is the length of a vector (timelike here) in spacetime.

In 3D space, a vector length remains the same (invariant) if we change (transform) coordinate systems. The components of the vector are different in a rotated (primed) coordinate system (i.e., $X_i' \neq X_i$), but the length remains the same. $l^2 = X_i X_i = X_i' X_i'$. By definition, a scalar is measured the same by observers using any coordinate system. Scalars are invariant under transformation to a new coordinate system.

The quantity $c\tau$, or simply the proper time τ passed on an object, is the same for all observers, is invariant in 4D spacetime, and hence is a scalar. $(c\tau)^2 = x^\mu x_\mu = x'^\mu x'_\mu$, even though $x'^\mu \neq x^\mu ; x'_\mu \neq x_\mu$. The term Lorentz invariance is commonly used for 4D scalars.

Other such scalars are the magnitudes of the 4-velocity of (2-54) [equal to c] and the 4-momentum of (2-55) [equal to mc.] Change the unprimed coordinate values in those relations to primed coordinates of another observer in another coordinate frame, and the magnitudes remain the same. We will soon encounter yet other such scalars.

As noted, the components of a vector change in different coordinate systems. This is true in 3D if we rotate to new coordinate axes. It is also true in 4D spacetime for coordinate systems in relative motion with respect to one another (unprimed vs primed coordinates). In both cases, the length of the vector remains the same. Objects which behave in this manner (e.g., vectors like x^μ, u^μ, and p_μ) are said to be covariant under transformation to a new coordinate system. For spacetime, the term Lorentz covariance is common.

Note that the same term "covariant", as opposed to "contravariant", is also used with respect to vector components, but the meaning there is different.

2.9.4 Invariance and the QFT Wave Equations

As we will see, beginning in Chap. 3, contravariant/covariant component notation will provide us with a very useful way of writing the relativistic wave equations of RQM and QFT (see first block of Wholeness Chart 1-2 in Chap. 1) and their solutions. Importantly, these forms of the wave equations are *invariant*. By this we mean that the numerical values of the vector components in the equations will change as the coordinate system changes, but the relations between the vector components will remain the same. In other words, the wave equation has the same form (it looks the same mathematically), whether we use unprimed or primed coordinates. The wave equation is invariant. This is the famous principle of relativity known as Lorentz invariance of the laws of nature. Different observers see different vector component values, but they find the same laws of nature governing the behavior of those components. This is a fundamental principle of special relativity theory, and since QFT is grounded in special relativity, it is a fundamental principle of QFT. Any valid relativistic quantum theory must obey Lorentz invariance. Its governing equations must be invariant.

Note that, with respect to equations, the term *Lorentz covariance (of equations)* is used in the literature interchangeably with *Lorentz invariance (of equations*). While the *form* of the equations is invariant, the *vectors* in the equation are covariant. Hence, the practice of using either term.

2.9.5 Other Uses for This Stuff

We have only scratched the surface of the mathematics of metrics, contravariant components, and covariant components, formally called differential geometry (or tensor analysis, or in the old days, Riemannian geometry.) Their enormous power becomes more evident when one studies curved spaces, such as the surface of a sphere or the spacetime around a black hole. However, hopefully, this appendix provides some justification for their use, which is widespread in QFT.

2.10 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 $\partial x^a / \partial x_\beta = g^{\alpha\beta}$ 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 ϕ_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. In a 3D Cartesian coordinate system, the metric $g_{\mu\nu} = \delta_{\mu\nu}$, the Kronecker delta. Show that, in such a system, $x^i = x_i$, velocity $v^i = v_i$, and 3-momentum $p^i = p_i$.

9. 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?

10. (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.)

11. 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 $U|\psi\rangle$?

12. Consider the unitary operator $U = e^{-iH(t-to)}$ and $|\psi_E\rangle = |e^{-i(Eto-\mathbf{p}\cdot\mathbf{x})}\rangle$, 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.)

This page intentionally left blank.

Part One

Free Fields

Like a bird on a wire,
like a drunk in midnight choir,
I have tried in my way to be free.

Sung by Joe Cocker
Lyrics by Leonard Cohen

Chapter 3 print vers 3/6/13 copyright of Robert D. Klauber

Scalars: Spin 0 Fields

..if I look back at my life as a scientist and a teacher, I think the most important and beautiful moments were when I say, "ah-hah, now I see a little better" … this is the joy of insight which pays for all the trouble one has had in this career.

Victor F. Weisskopf
Quarks, Quasars, and Quandaries

3.0 Preliminaries

This chapter presents the most fundamental concepts in the theory of quantum fields, and contains the very essence of the theory. Master this chapter, and you are well on your way to mastering that theory.

3.0.1 Background

Seeking a relativistic quantum theory?

Early efforts to incorporate special relativity into quantum mechanics started with the non-relativistic Schrödinger equation,

$$i\hbar\frac{\partial}{\partial t}\phi = H\phi \qquad \text{where } H = \frac{p^2}{2m} + V = -\frac{\hbar^2}{2m}\nabla^2 + V\,, \tag{3-1}$$

Try relativistic Hamiltonian in Schrödinger equation

and attempted to find a relativistic, rather than non-relativistic, form for the Hamiltonian H.[1] One might guess that approach would lead to a valid relativistic Schrödinger equation. This is, in essence, true but there is one problem, as we will see below.

In special relativity, the 4-momentum vector is Lorentz covariant, meaning its length in 4D space is invariant. For a free particle (i.e., $V = 0$),

$$p^\mu p_\mu = m^2c^2 = g_{\mu\nu}p^\mu p^\nu = \begin{bmatrix} E/c & p^1 & p^2 & p^3 \end{bmatrix}\begin{bmatrix} E/c \\ -p^1 \\ -p^2 \\ -p^3 \end{bmatrix} \;\rightarrow\; \frac{E^2}{c^2} = \mathbf{p}^2 + m^2c^2\,. \tag{3-2}$$

Relativistic energy E

Changing dynamical variables over to operators (as happens in quantization), i.e.,

$$E \rightarrow H \qquad \text{and} \qquad p_i \rightarrow -i\hbar\partial_i\,, \tag{3-3}$$

Relativistic E → relativistic operator H

one finds, from the RHS of (3-2),

---

[1] Actually, Schrödinger first attempted to find a wave equation that was relativistic and came up with what later came to be known as the Klein-Gordon equation, which we will study in this chapter. He discarded it because of problems discussed later on herein, and because it gave wrong answers for the hydrogen atom. Shortly thereafter, he deduced the non-relativistic Schrödinger equation we are familiar with. Some time afterwards, other researchers then tried to "relativize" that equation, as discussed herein.

$$H = \sqrt{-\hbar^2 c^2 \partial_i \partial_i + m^2 c^4}\ , \tag{3-4}$$

Bad news: Relativistic H has square root of a differential operator

seemingly the only form a relativistic Hamiltonian could take. Unfortunately, taking the square root of terms containing a derivative is problematic, and difficult to correlate with the physical world.

The solution to the problem of finding a relativistic Schrödinger equation has been found, however, and as we will see in the next three chapters, turns out to be different for different spin types. This was quite unexpected at first, but has since become a cornerstone of relativistic quantum theory. (See first row of Wholeness Chart 1-2 in Chap. 1, pg. 7.)

But answer has been found, as we will see

Particles with zero spin, such as π-mesons (pions) and the famous Higgs boson, are known as scalars, and are governed by one particular relativistic Schrödinger equation, deduced by (after Schrödinger, actually), and named after, Oscar Klein and Walter Gordon. Particles with ½ spin, such as electrons, neutrinos, and quarks, and known as spinors, by a different relativistic Schrödinger equation, discovered by Paul Dirac. And particles with spin 1, such as photons and the W's and Z's that carry the weak charge, and known as vectors, by yet another relativistic Schrödinger equation, discovered by Alexandru Proça. The Proça equation reduces, in the massless (photon) case, to Maxwell's equations.

Each spin type has its own relativistic wave equation

We will devote a separate chapter to each of these three spin types and the wave equation associated with each. We begin in this chapter with scalars.

3.0.2 Chapter Overview

RQM first,

RQM overview (scalars)

where we will look at

- deducing the Klein-Gordon equation, the first relativistic Schrödinger equation, using the relativistic H^2,
- solutions (which are states = wave functions) to the Klein-Gordon equation,
- probability density and its connection to the funny normalization constant in the solutions, and
- the problem with negative energies in the relativistic solutions.

Then QFT,

QFT overview (scalars)

- using the classical relativistic $\mathcal{L}$ (Lagrangian density) for scalar fields, and the Legendre transformation to get $\mathcal{H}$ (Hamiltonian density),
- from $\mathcal{L}$ and the Euler-Lagrange equation, finding the same Klein-Gordon equation, with the same mathematical form for the solutions, but this time the solutions are fields, not states,
- from 2nd quantization, finding the commutation relations for QFT,
- determining relevant operators in QFT: $H = \int \mathcal{H}\, d^3x$, number, creation/destruction, etc.,
- showing this approach avoids negative energy states,
- seeing how the vacuum is filled with quanta of energy $\frac{1}{2}\hbar\omega$,
- deriving other operators (probability density, 3-momentum, charge) and
- picking up relevant loose ends (scalars = bosons, Fock (multiparticle) space).

And then,

- seeing quantum fields in a different light, as harmonic oscillators.

With finally, and importantly,

- finding the Feynman propagator, the mathematical expression for virtual particles.

Free (no force) Fields

We study free (no interactions) case first

In this chapter, as well as Chaps. 4 (spin ½) and 5 (spin 1), we will deal only with fields/particles that are not interacting, i.e., feel no force = "free". Thus, we will take potential energy $V = 0$. In Chap. 7, which begins Part 2 of the book, we will begin to investigate interactions.

3.1 Relativistic Quantum Mechanics: A History Lesson

3.1.1 Two Possible Routes to RQM

Recall from Chaps. 1 and 2, that 1st quantization, for both non-relativistic and relativistic particle theories, entails i) using the classical form of the Hamiltonian as the quantum form of the

Non-commutating variables must be operators

Hamiltonian, and ii) changing Poisson brackets to commutators. We recall also from Prob. 6 of Chap. 1 that non-commutation of dynamical variables means those variables are operators (because ordinary numbers commute.) For example,

$$\left[p_i , x^j \right] = -i\hbar \delta_i^{\ j} \quad \xleftrightarrow{\text{equivalent}} \quad p_i = -i\hbar \partial_i \tag{3-5}$$

as the RHS above is the only form that satisfies the LHS, and it is an operator.

One might expect that this is the route we would follow to obtain RQM, i.e., 1st quantization of relativistic classical particle theory. However, historically, it was done differently. That is, RQM was first extrapolated from NRQM, not from classical theory. As illustrated in Fig. 3-1, it can be done either way.

In this book, to save space and time, we will only show one of these paths, the historical one represented by the lowest arrow in Fig. 3-1.

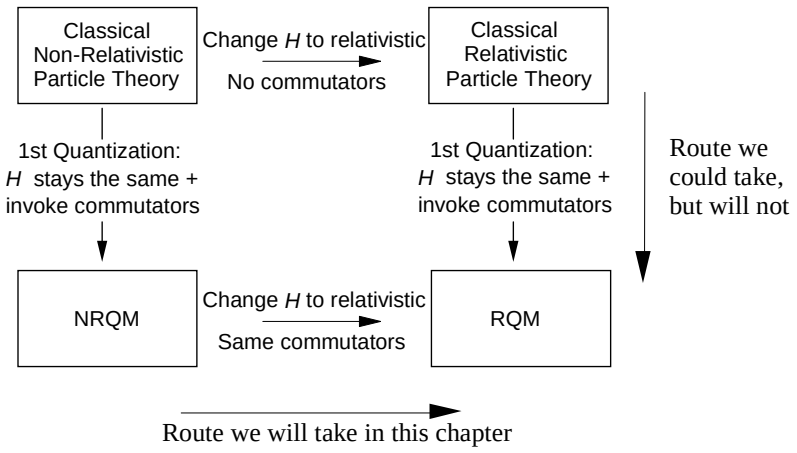


Figure 3-1. Different Routes to Relativistic Quantum Mechanics

3.1.2 Deducing the Klein-Gordon Equation

Let's square operators on both sides of Schrödinger eq

As we saw in Sect. 3.0.1, when we try to use a relativistic Hamiltonian in the Schrödinger equation, we have the problem of the partial derivative operator (see (3-4)) being under a square root sign. So, rather than use H, Klein and Gordon, in 1927, did the next best thing. They used H^2 instead. That is, they squared the operators (operate on each side twice rather than once) in the original Schrödinger equation (3-1) and thus from (3-2), obtained

$$\left(i\hbar \frac{\partial}{\partial t} \right) \left(i\hbar \frac{\partial}{\partial t} \right) \phi = H^2 \phi = \left(\mathbf{p}_{oper}^2 c^2 + m^2 c^4 \right) \phi \,, \tag{3-6}$$

Then use operator form for H^2

which becomes from the square of (3-4)

$$-\frac{\hbar^2}{c^2} \frac{\partial^2}{\partial t^2} \phi = \left(-\hbar^2 \frac{\partial}{\partial X_i} \frac{\partial}{\partial X_i} + m^2 c^2 \right) \phi \;\rightarrow\; -\frac{\partial}{\partial x^0} \frac{\partial}{\partial x_0} \phi = \left(\frac{\partial}{\partial x^i} \frac{\partial}{\partial x_i} + \underbrace{\frac{m^2 c^2}{\hbar^2}}_{\mu^2} \right) \phi \,. \tag{3-7}$$

To get the Klein-Gordon equation

Re-arranging, we have the Klein-Gordon equation (expressed in two equivalent ways with slightly different notation)

$$\left(\frac{\partial}{\partial x^\mu} \frac{\partial}{\partial x_\mu} + \mu^2 \right) \phi = 0 \quad \text{or} \quad \left(\partial_\mu \partial^\mu + \mu^2 \right) \phi = 0 \,, \qquad \mu^2 = \frac{m^2 c^2}{\hbar^2} \left(= m^2 \text{ in nat. units} \right). \tag{3-8}$$

Klein-Gordon equation is specifically for scalars

In 1934, Pauli and Weisskopf [2] showed that the Klein-Gordon equation specifically describes a spin-0 (scalar) particle. This should become evident to us as we study the Dirac and Proça equations, for spin ½ and spin 1, later on, and compare them to the Klein-Gordon equation.

3.1.3 The Solutions to the Klein-Gordon Equation

A solution set to (3-8), readily checked by substitution into (3-8) (which is good practice when using contravariant/covariant notation), is

Solutions to Klein-Gordon equation (discrete)

$$\phi(x) = \sum_{n=1}^{\infty} \frac{1}{\sqrt{2VE_n/\hbar}} \left(A_n e^{-\frac{i}{\hbar}(E_n t - \mathbf{p}_n \bullet \mathbf{x})} \underbrace{+\ B_n^{\dagger} e^{\frac{i}{\hbar}(E_n t - \mathbf{p}_n \bullet \mathbf{x})}}_{\text{absent in NRQM}} \right), \tag{3-9}$$

where we will discuss the funny looking normalization factor in front, containing the volume V and the energy of the nth solution, later. The coefficients $A_{\mathbf{n}}$ and $B_{\mathbf{n}}^{\dagger}$ are constants, and a complex conjugate form for the coefficient of the last term above, i.e., $B_n^{\dagger}$, is used because it will prove advantageous later.

Continuous solutions also exist

This is a discrete set of solutions, typical for cases with waves constrained inside a volume V, though V can be taken as large as one wishes. Each discrete wavelength in the summation of (3-9) fits an integer number of times inside the volume V. Continuous (integral rather than sum) solutions, for waves not constrained inside a specific volume V, exist for (3-8) as well, but we are not concerned with them at this point.

Only plane wave solutions here

This solution set is also specifically for plane waves. We will not consider alternative solution forms for other wave shapes that would exist in problems with cylindrical or spherical geometries.

Solutions in RQM are states (particles)

The solution (3-9), because we are working in RQM, is a state, i.e., $\phi(x)$ above = $|\phi(x)\rangle$, for a single particle. Each individual term in the summation is an eigenstate. $\phi(x)$ is a general state superposition of eigenstates.

Relativistic form has extra set of solutions

Note that in NRQM, we only had terms in the counterpart to (3-9) that had the exponential form of $-i(E_n t - \mathbf{p}_n \cdot \mathbf{x})/\hbar$, because that was the only form that satisfied the non-relativistic Schrödinger equation. Because we are using the square of the relativistic Hamiltonian in RQM, we get additional solutions of exponential form $+i(E_n t - \mathbf{p}_n \cdot \mathbf{x})/\hbar$ that also solve the relativistic Klein-Gordon equation. You should do Prob. 1, at the end of the chapter, to justify the statements in this paragraph to yourself.

Relations for p_μ and k_μ

With an aim towards using natural units, we note the following relations, where wave number $k_i = 2\pi/\lambda_i$,

$$p_\mu = \begin{bmatrix} E/c \\ p_i \end{bmatrix} = \begin{bmatrix} E/c \\ -p^i \end{bmatrix} = \hbar k_\mu = \begin{bmatrix} \hbar\omega/c \\ -\hbar k^i \end{bmatrix} \xrightarrow{\text{nat. units}} p_\mu = \begin{bmatrix} E \\ -p^i \end{bmatrix} = k_\mu = \begin{bmatrix} \omega \\ -k^i \end{bmatrix}, \tag{3-10}$$

and recall the notation introduced in Chap. 2,

Notation review

$$\begin{aligned} px &= p_\mu x^\mu = Et - p^i x^i = Et - \mathbf{p}\bullet\mathbf{x} \qquad \left(= p^\mu x_\mu\right) \\ kx &= k_\mu x^\mu = \omega t - k^i x^i = \frac{Et}{\hbar} - \frac{p^i x^i}{\hbar} = \frac{p_\mu}{\hbar} x^\mu \qquad \left(= k^\mu x_\mu\right) \end{aligned} \tag{3-11}$$

$$\text{in nat. units} \rightarrow \quad E = \omega, \quad p_i = k_i, \quad p_\mu = k_\mu, \quad px = kx.$$

It is then common to re-write (3-9) in natural units with the above notation. In doing so, we also switch the dummy summation variable n, which represents each individual wave in the summation, to the 3D vector quantity $\mathbf{k}$, representing the wave number and direction of each possible wave. For free fields, a given wave with wave number vector $\mathbf{k}$ has a particular energy (see (3-2) with $\mathbf{p} = \mathbf{k}$ in natural units), and we can designate that energy via either $E_{\mathbf{k}}$ or $\omega_{\mathbf{k}}$. It is common practice for scalars to use $\mathbf{k}$ (rather than $\mathbf{p}$) and $\omega_{\mathbf{k}}$ (rather than $E_{\mathbf{p}}$ or $E_{\mathbf{k}}$.)

---

[2] Pauli, W. and Weisskopf, V., Helv. Phys. Acta 7, 709 (1934). Translation in Miller, A. I., *Early Quantum Electrodynamics: A Source Book*, Cambridge U. Press, New York (1994)

The Klein-Gordon equation solutions (3-9) then become, in natural units

$$\phi(x) = \sum_{\mathbf{k}} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}} \left(A_{\mathbf{k}} e^{-ikx} + B_{\mathbf{k}}^{\dagger} e^{ikx} \right) \quad . \tag{3-12}$$

Natural units form of Klein-Gordon solutions

Except for Box 3-1, which reviews NRQM, we will henceforth, in this chapter, use natural units.

Definition of Eigensolutions

As noted previously, in RQM, the solution ϕ of (3-12) is that of a general (sum of eigenstates) single particle state. Each eigenstate has mathematical form (where we are going to omit the $2\omega_{\mathbf{k}}$ part here, because of what is coming)

$$\phi_{\mathbf{k},A} = \frac{e^{-ikx}}{\sqrt{V}} \quad \text{or} \quad \phi_{\mathbf{k},B^{\dagger}} = \frac{e^{ikx}}{\sqrt{V}} \, . \tag{3-13}$$

Eigenstates of Klein-Gordon equation

Each of these forms has what is called unit norm. That is, for $\phi_{\mathbf{k},A}$ (and similarly, for $\phi_{\mathbf{k},B\dagger}$),

$$\int \phi_{\mathbf{k},A}^{\dagger} \phi_{\mathbf{k},A} d^3x = \frac{1}{V} \int_V e^{ikx} e^{-ikx} d^3x = 1 \, , \tag{3-14}$$

Eigenstates have unit norm

or more generally, all such eigenstates are orthonormal, i.e., their inner products are

$$\int \phi_{\mathbf{k},A}^{\dagger} \phi_{\mathbf{k}',A} d^3x = \frac{1}{V} \int_V e^{ikx} e^{-ik'x} d^3x = \delta_{\mathbf{k}\mathbf{k}'} \, . \tag{3-15}$$

and are orthogonal

Similar relations to (3-15) exist for $\phi_{\mathbf{k},B\dagger}$, and every $\phi_{\mathbf{k},A}$ is orthogonal to every $\phi_{\mathbf{k},B\dagger}$. Work this out by doing Prob.2.

Relations (3-13) to (3-15) should look familiar from NRQM. There, (3-14) was the integral of the probability density for a particle in an eigenstate. In RQM, however, things are a little different, as we will see, and we use the term "unit norm" for the property displayed in (3-14).

Unit norm eigenstates were advantageous in NRQM, and they will be in QFT as well. That is the reason we omitted the $2\omega_{\mathbf{k}}$ part of our solutions (3-12) in forming our definitions (3-13). By so doing, the eigenstates then have unit norm, and things just turn out easier later on.

We defined eigenstates to have unit norm because it will be advantageous

3.1.4 Probability Density in RQM

We are going to investigate probability density in RQM, but first look over Box 3-1, and be sure you understand how probability density is derived in NRQM.

Probability Density Using the Klein-Gordon Equation

For RQM, we start with the Klein-Gordon equation rather than Schrödinger equation. First post-multiply it by $\phi^{\dagger}$, then subtract the complex conjugate equation post-multiplied by ϕ, i.e.,

$$\left\{ \frac{\partial^2}{\partial t^2} \phi = \left(\nabla^2 - \mu^2 \right) \phi \right\} \phi^{\dagger} \quad - \quad \left\{ \frac{\partial^2}{\partial t^2} \phi^{\dagger} = \left(\nabla^2 - \mu^2 \right) \phi^{\dagger} \right\} \phi \quad , \tag{3-16}$$

Deduce RQM probability density using relativistic wave equation

and note that $\mu^2 \phi^{\dagger} \phi - \mu^2 \phi \phi^{\dagger} = 0$. The LHS of the result can be replaced with the new LHS in (3-17) below, and the RHS with (3-18).

$$\underbrace{\frac{\partial}{\partial t} \left(\frac{\partial \phi}{\partial t} \phi^{\dagger} - \frac{\partial \phi^{\dagger}}{\partial t} \phi \right)}_{\text{new LHS}} = \underbrace{\frac{\partial^2 \phi}{\partial t^2} \phi^{\dagger} - \frac{\partial^2 \phi^{\dagger}}{\partial t^2} \phi}_{\text{LHS of result above}} + \underbrace{\frac{\partial \phi}{\partial t} \frac{\partial \phi^{\dagger}}{\partial t} - \frac{\partial \phi^{\dagger}}{\partial t} \frac{\partial \phi}{\partial t}}_{=0} \tag{3-17}$$

$$\underbrace{\nabla \cdot \left((\nabla \phi) \phi^{\dagger} - \left(\nabla \phi^{\dagger} \right) \phi \right)}_{\text{new RHS}} = \underbrace{\left(\nabla^2 \phi \right) \phi^{\dagger} - \left(\nabla^2 \phi^{\dagger} \right) \phi}_{\text{RHS of result above}} + \underbrace{\nabla \phi \cdot \nabla \phi^{\dagger} - \nabla \phi^{\dagger} \cdot \nabla \phi}_{=0} \tag{3-18}$$

Box 3-1. Review of Non-Relativistic QM Probability Density

In non-relativistic quantum mechanics (NRQM), we encountered 1) the wave function solution to the Schrödinger equation Ψ, and 2) the particle probability density $\rho = \Psi^\dagger\Psi$ (or equivalently when Ψ is a scalar quantity, $\Psi^*\Psi$.) We review here the derivation of that relation for probability density.

Conserved quantities in field theory:

Recall the continuity equation of continuum mechanics and electromagnetism,

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{j} = 0 \qquad \left(\xrightarrow{\text{implies}} \int_V \rho d^3x = \text{constant in time} \right), \tag{B3-1.1}$$

where ρ is density (mass or charge density), $\mathbf{j}$ is the 3D current density (mass/area-sec or charge/area-sec), and V is all space, or at least large enough so that everywhere outside it, for all time, $\rho = 0$. V is fixed in space and time, whereas ρ can change in space and time inside V. Any conserved quantity (such as total mass M or total charge Q) obeys (B3-1.1).

The general procedure:

Use the governing quantum wave equation to deduce another equation having the form of the continuity equation (B3-1.1), and we will then know that ρ, whatever it turns out to be in that case, must represent a conserved quantity. Its integral over all space is constant in time. If we normalize ρ such that when integrated over all space, the result equals one, we can conjecture that ρ is the particle probability density (which when integrated over all space equals the probability that we will find the particle somewhere in all space, i.e., one.) Then throughout time, as our particle evolves, moves, and rearranges its probability density distribution, the total probability of finding it somewhere in space is always one. It turns out, from experiment, that the conjecture that this quantity ρ in NRQM equals probability density is true.

Probability Density Using the Schrödinger Equation:

First, pre-multiply the Schrödinger equation by the complex conjugate of the wave function, i.e.,

$$\Psi^\dagger\left\{\frac{\partial}{\partial t}\Psi = \frac{1}{i\hbar}\left(-\frac{\hbar^2}{2M}\nabla^2 + V\right)\Psi\right\} \tag{B3-1.2}$$

Then, post-multiply the complex conjugate of the Schrödinger equation by the wave function

$$\left\{\frac{\partial}{\partial t}\Psi^\dagger = \frac{-1}{i\hbar}\left(-\frac{\hbar^2}{2M}\nabla^2 + V^\dagger\right)\Psi^\dagger\right\}\Psi \tag{B3-1.3}$$

where the potential V is real so $V=V^\dagger$. Adding (B3-1.2) to (B3-1.3), we get

$$\Psi^\dagger\frac{\partial\Psi}{\partial t} + \frac{\partial\Psi^\dagger}{\partial t}\Psi = \Psi^\dagger\frac{1}{i\hbar}\left(-\frac{\hbar^2}{2M}\nabla^2 + V\right)\Psi + \left(\frac{-1}{i\hbar}\left(-\frac{\hbar^2}{2M}\nabla^2\Psi^\dagger + V^\dagger\Psi^\dagger\right)\right)\Psi \tag{B3-1.4}$$

or

$$\frac{\partial\left(\Psi^\dagger\Psi\right)}{\partial t} = \frac{-\hbar}{2iM}\underbrace{\left(\Psi^\dagger\left(\nabla^2\Psi\right) - \left(\nabla^2\Psi^\dagger\right)\Psi\right)}_{\nabla\cdot\left[\Psi^\dagger(\nabla\Psi) - (\nabla\Psi^\dagger)\Psi\right]} + \underbrace{\frac{\Psi^\dagger V\Psi}{i\hbar} - \frac{V^\dagger\Psi^\dagger\Psi}{i\hbar}}_{=0 \text{ since } V^\dagger = V} \tag{B3-1.5}$$

This is the same as the continuity equation (B3-1.1) if we take as our probability density

$$\rho = \Psi^\dagger\Psi \ , \tag{B3-1.6}$$

and as our probability current

$$\mathbf{j} = \frac{\hbar}{2iM}\left\{\Psi^\dagger\left(\nabla\Psi\right) - \left(\nabla\Psi^\dagger\right)\Psi\right\} . \tag{B3-1.7}$$

This is how the commonly used relation (B3-1.6) is found.

Manipulations of the wave equation lead to an equation like the continuity equation

Equating the new LHS of (3-17) to the new RHS of (3-18), and to make future work easier, multiplying both sides by the constant *i*, gives the form of the continuity equation

$$i\frac{\partial}{\partial t}\left(\frac{\partial\phi}{\partial t}\phi^\dagger-\frac{\partial\phi^\dagger}{\partial t}\phi\right)=i\nabla\cdot\left((\nabla\phi)\phi^\dagger-(\nabla\phi^\dagger)\phi\right)\ \rightarrow\ \frac{\partial\rho}{\partial t}+\nabla\cdot\mathbf{j}=0\,, \tag{3-19}$$

where probability density and the probability current for a Klein-Gordon particle are

From that, we deduce form of RQM probability density

$$\rho=j^0=i\left(\frac{\partial\phi}{\partial t}\phi^\dagger-\frac{\partial\phi^\dagger}{\partial t}\phi\right),\ \text{and} \tag{3-20}$$

$$\mathbf{j}=-i\left((\nabla\phi)\phi^\dagger-(\nabla\phi^\dagger)\phi\right)\qquad j_i=-i\left(\phi_{,i}\phi^\dagger-\phi^\dagger{}_{,i}\phi\right)\qquad j^i=i\left(\phi^{,i}\phi^\dagger-\phi^{\dagger,i}\phi\right). \tag{3-21}$$

Importantly, and perhaps surprisingly, the relativistic form of the probability density (3-20) is <u>*not*</u> the same as (B3-1.6), the NRQM probability density.

<u>4 Currents</u>

We introduce 4D notation for the scalar and 3D vector of (3-19) and define the <u>scalar 4-current</u>

4-current and 4D *form of continuity equation*

$$j^\mu=\begin{bmatrix}\rho\\ \mathbf{j}\end{bmatrix}=\begin{bmatrix}\rho\\ j^i\end{bmatrix}=\begin{bmatrix}j^0\\ j^i\end{bmatrix}=i\left(\phi^{,\mu}\phi^\dagger-\phi^{\dagger,\mu}\phi\right)\ . \tag{3-22}$$

The <u>4D continuity equation</u> form of (3-19) is then

$$\boxed{\frac{\partial j^\mu}{\partial x^\mu}=\partial_\mu j^\mu=j^\mu{}_{,\mu}=0}\,, \tag{3-23}$$

4-divergence of 4-current of conserved quantity always = 0

where we have shown three common notational ways to designate partial derivative. (3-23) tells us the *important fact* that the <u>4-divergence</u> *of the 4-current of any conserved quantity* (total probability in this case) *is zero.*

<u>Probability for Klein-Gordon Discrete Solutions</u>

Scalar probability density in terms of first Klein-Gordon solution set

For a single particle state in RQM, we are going to <u>assume at first</u>, for simplicity, that the solution (3-12), has <u>only terms with coefficients $A_\mathbf{k}$</u>, i.e., the general state ϕ contains no eigenstates shown with coefficients $B_\mathbf{k}{}^\dagger$. Probability density (3-20) is then (where primes do *not* denote derivatives with respect to spatial coordinates, merely different summation dummy variables)

$$\rho=\left(\sum_\mathbf{k}\frac{\omega_\mathbf{k}A_\mathbf{k}}{\sqrt{2\omega_\mathbf{k}}}\frac{e^{-ikx}}{\sqrt{V}}\right)\left(\sum_{\mathbf{k}'}\frac{A^\dagger_{\mathbf{k}'}}{\sqrt{2\omega_{\mathbf{k}'}}}\frac{e^{ik'x}}{\sqrt{V}}\right)+\left(\sum_{\mathbf{k}'}\frac{\omega_{\mathbf{k}'}A^\dagger_{\mathbf{k}'}}{\sqrt{2\omega_{\mathbf{k}'}}}\frac{e^{ik'x}}{\sqrt{V}}\right)\left(\sum_\mathbf{k}\frac{A_\mathbf{k}}{\sqrt{2\omega_\mathbf{k}}}\frac{e^{-ikx}}{\sqrt{V}}\right), \tag{3-24}$$

where the $\omega_\mathbf{k}$ and $\omega_{\mathbf{k}'}$ came from the time derivatives.

Square of absolute value of coefficient $A_\mathbf{k}$ = probability of finding **k***th eigenstate*

If we integrate ρ over the volume V (which is large enough to encompass the entire state), the result must equal 1. When we do so, all terms with $\mathbf{k}'\neq\mathbf{k}$ go to zero, so the $\omega_{\mathbf{k}'}\rightarrow\omega_\mathbf{k}$ and cancel out. The V term in the denominator cancels in the integration over the volume V, and the two terms result in a factor of 2 that cancels with the 2 in the denominator. The result is

$$\int\rho d^3x=\sum_\mathbf{k}|A_\mathbf{k}|^2=1\,. \tag{3-25}$$

Thus $|A_\mathbf{k}|^2$ is the probability of measuring the **k**th eigenstate, similar to what the coefficients of eigenstates represented in NRQM.

<u>Difference from NRQM</u>

Comparing probability in NRQM and RQM

Note that in RQM

$$\int\underbrace{\phi^\dagger\phi}_{\neq\rho}d^3x=\sum_\mathbf{k}\frac{(A_\mathbf{k})^2}{2\omega_\mathbf{k}}\neq1\quad\text{but}\quad\int\underbrace{i\left(\frac{\partial\phi}{\partial t}\phi^\dagger-\frac{\partial\phi^\dagger}{\partial t}\phi\right)}_{=\rho}d^3x=\sum_\mathbf{k}|A_\mathbf{k}|^2=1\qquad(\text{RQM})\,, \tag{3-26}$$

whereas in NRQM, we had

$$\int \underbrace{\phi^\dagger \phi}_{=\rho}\, d^3x = \sum_{\mathbf{k}} |A_{\mathbf{k}}|^2 = 1 \qquad \text{(NRQM)}\,. \tag{3-27}$$

Normalization Factors

*RQM normalization factors arise from need to have total probability = 1 and $|A_{\mathbf{k}}|^2$ = probability of **k**th state*

Obtaining the RHS of (3-26) is the reason for the normalization factors $1/\sqrt{2\omega_{\mathbf{k}}V}$ used in the solution ϕ of (3-12) and (3-9). Those factors result in a total probability of one for a single particle and $|A_{\mathbf{k}}|^2$ as the probability for measuring the respective eigenstate. That is, the form of the relativistic field equation gave us the form of the probability density in (3-20) (and (3-26)), and the need to have total probability of unity gave us the normalization factors in the solutions.

Relativistic Invariance of Probability

Total probability and $A_{\mathbf{k}}$ are frame independent (relativistically invariant)

This total probability value of unity in (3-25) (and (3-26)) is a relativistic invariant (i.e., a world scalar.) If we change our frame, the energy spectrum (i.e., the $\omega_{\mathbf{k}}$ values) will change (kinetic energy for each energy-momentum eigenstate looks different). But these changes cancel out in the probability calculation, since the $\omega_{\mathbf{k}}$ cancel, and always result in a total probability of one for any frame. Further, the $A_{\mathbf{k}}$ here are constants that do not vary with frame, so the probability of finding any particular state is also independent of what frame the measurements are taken in.

Note that this means the normalization factors chosen provide relativistic invariance of total probability, which we would not have had with any other choice.

3.1.5 Negative Energies in RQM

If we take our traditional operator form for H as $i\partial/\partial t$ and operate on one of our Klein-Gordon solution eigenstates of (3-12) and (3-13), we should get the energy eigenvalue $\omega_{\mathbf{k}}$. When we do this for the eigenstates with exponents in $-ikx$, all looks as expected.

$$H\phi_{\mathbf{k},A} = i\frac{\partial \phi_{\mathbf{k},A}}{\partial t} = E_{\mathbf{k},A}\phi_{\mathbf{k},A} = i\frac{\partial}{\partial t}\frac{e^{-ikx}}{\sqrt{V}} = \omega_{\mathbf{k}}\frac{e^{-ikx}}{\sqrt{V}} = \omega_{\mathbf{k}}\phi_{\mathbf{k},A}\,. \tag{3-28}$$

However, when we do it for the eigenstates with exponents in $+ikx$, we have an "uh-oh", i.e.,

Half of our RQM eigenstates have negative energy

$$H\phi_{\mathbf{k},B^\dagger} = i\frac{\partial \phi_{\mathbf{k},B^\dagger}}{\partial t} = E_{\mathbf{k},B^\dagger}\phi_{\mathbf{k},B^\dagger} = i\frac{\partial}{\partial t}\frac{e^{ikx}}{\sqrt{V}} = -\omega_{\mathbf{k}}\frac{e^{ikx}}{\sqrt{V}} = -\omega_{\mathbf{k}}\phi_{\mathbf{k},B^\dagger}\,. \tag{3-29}$$

Since $\omega_{\mathbf{k}}$ is always a positive number, we have states with negative energies in RQM. We might have expected this, since we used the square of the Hamiltonian as the basis of RQM, and square roots typically have both positive and negative signs.

The bottom line: This is not an attribute of what a good theory has been expected to have, i.e., solely positive energies as we see in our world. As we will shortly see, QFT solved this dilemma (as well as others delineated in Chap. 1.)

3.1.6 Negative Probabilities in RQM

Half of our RQM eigenstates have negative probability density

Do Prob. 3 to prove to yourself that a particle ϕ containing only eigenstates of the exponential form $+i(E_n t - \mathbf{p}_n\cdot\mathbf{x})/\hbar = ikx$ (i.e., those with coefficients $B_{\mathbf{k}}^\dagger$ in (3-12)) has total probability of being measured of –1. The extra states in RQM have physically untenable negative probabilities!

Time to move on to QFT.

3.2 The Klein-Gordon Equation in Quantum Field Theory

3.2.1 States vs Fields

States & fields both spread out in space. But in quantum theories, "field" also means "operator"

It should come as no surprise, to those who have read Chap. 1, that the fundamental scalar wave equation of RQM, the Klein-Gordon equation (3-8), is also the fundamental scalar wave equation of QFT, except that ϕ therein is considered a field, instead of a state. The word "field" in classical theory means an entity that, unlike a particle, is spread out, i.e., is a function of space (it has different values at different spatial locations) and typically also a function of time. The state ϕ of NRQM and RQM certainly fills that bill, but in quantum theory we don't use the word "field" for this, we use the word "state" (or "wave function" or "ket" or "particle".)

The word "field" in quantum theory refers to a quantity that is spread out in space, but also, importantly, as we will soon see, is an operator in QFT. More properly, it is called a quantum field

or an operator field, though the short term field is far more common. Confusingly, we use the same symbol ϕ in QFT for a field as we used for a state in NRQM and RQM.

Notation

In QFT, symbols such as ϕ, which are not part of a ket symbol, do not represent states, but fields. Unless otherwise explicitly noted, in QFT notation,

Notational difference between states and fields. In QFT, ϕ is not a state, but a field

$|\phi\rangle$ symbolizes a state (particle) and ϕ symbolizes a field (operator),

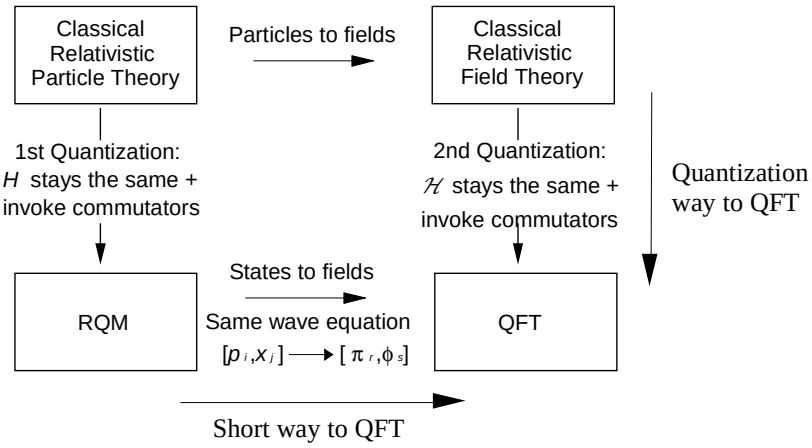
On the other hand, in NRQM and RQM, both symbols above represented the same thing, a state.

We will understand these distinctions a little better later, but for now understand that formally, the Klein-Gordon equation in QFT is called a field equation, because its solution ϕ is a (quantum or operator) field. See the second and third rows of Wholeness Chart 1-2 in Chap. 1, pg. 7.

There are two common ways to derive this equation, which we present in the following two sections, plus a third, which is a good check on the theory and can be found in the Appendix A.

3.2.2 From RQM to QFT

Fig. 3-2 illustrates, schematically, the two basic routes to QFT. The quickest is at the bottom of the figure, for which we simply postulate that the solution ϕ of the Klein-Gordon equation (3-8) describes a field (instead of a particle). This is reasonable, since ϕ is a function of spatial location (and often time), i.e., it is a field in the formal mathematical sense.



Two different routes to QFT

Figure 3-2. Different Routes to Quantum Field Theory

We then must apply the commutation relations for fields (see Chap. 2, pg. 31, Wholeness Chart 2-5, 6th column = 3rd column on right hand page), instead of the commutation relations for particle properties (same chart, 3rd column on left hand page). When we do this, and simply crank the mathematics, we obtain QFT. Because the QFT we then obtain describes the real world so well, it justifies the original postulate.

Short route: RQM → QFT. Similar math as 2nd quantization below

The formal mathematics are much the same as for the alternative route, illustrated on the RHS of Fig. 3-2, and treated in the next section.

3.2.3 From Classical Relativistic Fields to QFT

2nd quantization route: Classical fields → QFT

Classical Scalar Fields

The classical Lagrangian density for a free (no forces), real, relativistic scalar field ϕ has form

Start with classical Lagrangian density for free scalar field

$$\mathcal{L}_0^0 = K\left(\partial_\alpha\phi\partial^\alpha\phi - \mu^2\phi\phi\right) = K\left(\dot\phi\dot\phi + \partial_i\phi\partial^i\phi - \mu^2\phi\phi\right) = K\left(\dot\phi\dot\phi - \underbrace{\partial_i\phi\partial_i\phi}_{\nabla\phi\cdot\nabla\phi} - \mu^2\phi\phi\right), \qquad (3\text{-}30)$$

where ϕ, since it is a classical field, is real (not complex), μ is a constant to be determined by experiment, K is an arbitrary constant, the superscript "0" on $\mathcal{L}$ stands for scalar (with spin 0), and

the subscript "0" means "free". This is not the place to do classical theory, so we will not derive (3-30) here. We do note in passing that (3-30) is a general result derived by insisting that ϕ and $\mathcal{L}$ are Lorentz invariants (i.e., world scalars – see Chap. 2 including appendix) and that the associated Euler-Lagrange equation is also Lorentz invariant in form. (3-30) is the only form that satisfies these conditions and results in a linear field equation (i.e., ϕ appears only to first power.) A non-linear field equation might work, but is far more complicated. For free fields, we will find a linear equation works well.

Using the Legendre transformation, we can readily use (3-30) to find the Hamiltonian density, where ${\pi_0}^0$ is the field conjugate momentum,

Find Hamiltonian density from Legendre transformation

$$\mathcal{H}_0^0 = \pi_0^0\dot{\phi} - \mathcal{L}_0^0 = \underbrace{\frac{\partial \mathcal{L}_0^0}{\partial \dot{\phi}}}_{2K\dot{\phi}}\dot{\phi} - \mathcal{L}_0^0 = K\left(\dot{\phi}\dot{\phi} + \nabla\phi\cdot\nabla\phi + \mu^2\phi\phi\right). \tag{3-31}$$

We may be tempted at this point to proceed with quantization, and simply use the $\mathcal{H}$ and $\mathcal{L}$ above along with the appropriate commutators. However, we know that in quantum mechanics most meaningful things are complex, not real. Quite the reverse of the macroscopic world we live in, and for which real fields of form ϕ generally apply.

Classical field taken as complex

So, we adopt one more postulate, which is that our field ϕ be complex. This means re-expressing our values for $\mathcal{H}$ and $\mathcal{L}$ in terms of a complex field, but such that $\mathcal{H}$ and $\mathcal{L}$ remain real (energy, and energy density $\mathcal{H}$, must be real numbers.) Doing this, where we choose to take K=1, yields the free, complex scalar field Lagrangian and Hamiltonian densities

Re-express Lagrangian and Hamiltonian densities in terms of complex fields

$$\boxed{\mathcal{L}_0^0 = \left(\partial_\alpha\phi^\dagger\,\partial^\alpha\phi - \mu^2\phi^\dagger\phi\right) = \left(\dot{\phi}^\dagger\dot{\phi} - \nabla\phi^\dagger\cdot\nabla\phi - \mu^2\phi^\dagger\phi\right)}, \text{ and} \tag{3-32}$$

$$\boxed{\mathcal{H}_0^0 = \frac{\partial\mathcal{L}_0^0}{\partial\dot{\phi}^r}\dot{\phi}^r - \mathcal{L}_0^0 = \underbrace{\frac{\partial\mathcal{L}_0^0}{\partial\dot{\phi}}}_{\pi_0^0=\dot{\phi}^\dagger}\dot{\phi} + \underbrace{\frac{\partial\mathcal{L}_0^0}{\partial\dot{\phi}^\dagger}}_{\pi_0^{0\dagger}=\dot{\phi}}\dot{\phi}^\dagger - \mathcal{L}_0^0 = \dot{\phi}\dot{\phi}^\dagger + \nabla\phi^\dagger\cdot\nabla\phi + \mu^2\phi^\dagger\phi}\,. \tag{3-33}$$

Take care to realize that ϕ and $\phi^\dagger$ are considered *separate fields in the summation over field types r*, and note the definitions of their respective conjugate momenta. That is, ${\pi_0}^0$ equals the complex conjugate of the time derivative of the field (not the time derivative of the field.) $\pi_0^{0\dagger}$ equals the time derivative of the field, not its complex conjugate.

If, as we progress, we find situations where real, rather than complex, fields are involved, we can simply deal with the special case of a complex field where the imaginary part is zero. Assuming a complex field above means we assumed the most general case.

Deriving the Klein-Gordon Field Equation

Substituting the Lagrangian density (3-32) into the Euler-Lagrange field equation,

Use $\mathcal{L}$ in Euler-Lagrange equation to get Klein-Gordon equation

$$\frac{d}{dx^\mu}\left(\frac{\partial\mathcal{L}}{\partial\phi^r{}_{,\mu}}\right) - \frac{\partial\mathcal{L}}{\partial\phi^r} = 0\ . \tag{3-34}$$

yields the Klein-Gordon equation for fields, where again, the values r=1,2 signify, respectively, the field ϕ and its complex conjugate transpose $\phi^\dagger$ (also called the Hermitian) which, for scalars, is simply the complex conjugate,

$$\boxed{\begin{aligned}\left(\partial_\mu\partial^\mu + \mu^2\right)\phi &= \left(\Box^2 + \mu^2\right)\phi = 0 \qquad \text{(a)}\\ \left(\partial_\mu\partial^\mu + \mu^2\right)\phi^\dagger &= \left(\Box^2 + \mu^2\right)\phi^\dagger = 0. \qquad \text{(b)}\end{aligned}} \tag{3-35}$$

In the above, we have introduced the $\Box^2$ symbol, the 4D equivalent of the 3D Laplacian, $\nabla^2 = \partial^i\partial^i = \partial_i\partial_i = -\partial^i\partial_i$. (Note, some authors use $\Box$ instead of $\Box^2$.) We could, of course, also have obtained (3-35)(b) by taking the complex conjugate transpose of (3-35)(a), since everything inside the parentheses is real.

Recall from Chap. 2, that given any one of $\mathcal{H}$, $\mathcal{L}$, or the field equation, we can deduce any of the others (via the Legendre transformation and the Euler-Lagrange equation). So knowing any one of these is equivalent to knowing any of the others, and our first postulate of 2nd quantization could have stipulated the same $\mathcal{L}$ in classical theory and QFT, or the same field equation, instead of $\mathcal{H}$.

The discrete plane wave solutions to (3-35) are the same as (3-12), and its Hermitian, from RQM, i..e.,[3]

$$\begin{aligned}
\phi(x) &= \underbrace{\sum_{\mathbf{k}} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}} a(\mathbf{k}) e^{-ikx}}_{} + \underbrace{\sum_{\mathbf{k}} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}} b^{\dagger}(\mathbf{k}) e^{ikx}}_{} \qquad \text{(a)} \\
&= \qquad \phi^{+} \qquad + \qquad \phi^{-} \\
\phi^{\dagger}(x) &= \underbrace{\sum_{\mathbf{k}} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}} b(\mathbf{k}) e^{-ikx}}_{} + \underbrace{\sum_{\mathbf{k}} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}} a^{\dagger}(\mathbf{k}) e^{ikx}}_{} \qquad \text{(b)} \\
&= \qquad \phi^{\dagger +} \qquad + \qquad \phi^{\dagger -} \ .
\end{aligned} \tag{3-36}$$

Discrete plane wave solutions to Klein-Gordon field equation

Note the new symbolism for each of the solution forms. We use lower case coefficients in QFT because, as we will see, the coefficients play a much different role in QFT than they did in RQM, and we need to distinguish them.

The continuous plane wave solutions to (3-35) are

$$\begin{aligned}
\phi(x) &= \underbrace{\int \frac{d^3\mathbf{k}}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}}}} a(\mathbf{k}) e^{-ikx}}_{} + \underbrace{\int \frac{d^3\mathbf{k}}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}}}} b^{\dagger}(\mathbf{k}) e^{ikx}}_{} \qquad \text{(a)} \\
&= \qquad \phi^{+} \qquad + \qquad \phi^{-} \\
\phi^{\dagger}(x) &= \underbrace{\int \frac{d^3\mathbf{k}}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}}}} b(\mathbf{k}) e^{-ikx}}_{} + \underbrace{\int \frac{d^3\mathbf{k}}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}}}} a^{\dagger}(\mathbf{k}) e^{ikx}}_{} \ . \qquad \text{(b)} \\
&= \qquad \phi^{\dagger +} \qquad + \qquad \phi^{\dagger -}
\end{aligned} \tag{3-37)}$$

Continuous plane wave solutions to Klein-Gordon field equation

The continuous solutions represent waves that are not constrained to a specific volume. Wavelengths for such solutions do not have to fit an integer number of times inside a particular volume, and thus are not limited to discrete values.

Note also the shorthand notation for each of the four different solution sets underneath the brackets. You will see these symbols again and again, so you might want to consider making a copy of (3-36), pasting it above your desk, and doing memorization tests with yourself every day until they become ingrained in your consciousness. Try to remember that $\phi^{\dagger +}$ is *not* the complex conjugate of ϕ^{+}, for example, contrary to what you might expect. The + sign refers to a term with positive energy in the RQM sense (i.e., – sign before the energy in the exponent.) It might help to think that because † changes the sign of the imaginary part of every complex quantity, it also changes the sign of the symbol ϕ^{-}. So, $(\phi^{-})^{\dagger} = \phi^{\dagger +}$.

Learn the shorthand notation for the four types of solutions

[3] These solutions have the familiar $\pm i(\omega_{\mathbf{k}} t - \mathbf{k}\bullet\mathbf{x})$ form in the exponent. There are actually additional solutions to the Klein-Gordon equation having form $\pm i(\omega_{\mathbf{k}} t + \mathbf{k}\bullet\mathbf{x})$, but these have been widely ignored. These solutions, and their impact on QFT, are discussed in R. D. Klauber, "Mechanism for Vanishing Zero-Point Energy", http://arxiv.org/abs/astro-ph/0309679 (2003).

However you do it, being able to readily recall the definitions of the symbols in (3-36) and (3-37) will help in the future.

Finding μ^2

Do Prob. 4 to prove to yourself that the value of μ^2, which appeared as an unknown constant in the theoretical determination of (3-30), has the same value it did in RQM, i.e.,

$$\boxed{\mu^2 = \frac{m^2c^2}{\hbar^2} \quad \left(= m^2 \text{ in nat. units}\right)}. \tag{3-38}$$

μ^2 has same value in QFT as in RQM

Third Way to Klein-Gordon Equations: A Consistency Check

Recall from Chap 2. and Wholeness Charts 2-2 and 2-5 (pgs. 20 and 30), that we could express the equations of motion for classical fields in terms of Poisson brackets in the former chart, and for Heisenberg picture quantum fields, in terms of commutators in the latter chart. The commutator-based equation of motion for a quantum field in the next to last box in the right hand column of Wholeness Chart 2-5 (reproduced below on the LHS of (3-39)) is in terms of the Hamiltonian and the field. For scalar fields, this equation of motion for ϕ should be essentially the same as the Klein-Gordon equation for ϕ. That is,

There is yet a third way to derive the field equation of motion, this time from a variational math relation

Heisenberg Picture Field Equation of Motion		Klein-Gordon Field Equation
$\dot{\phi} = -i[\phi, H]$	$\xleftrightarrow{\text{should be same thing}}$	$\left(\partial_\mu \partial^\mu + \mu^2\right)\phi = 0$ (3-39)

In the Appendix A of this chapter, we show that this is indeed true, and thus our theory is self consistent. It also proves that the Klein-Gordon field equation of QFT (3-35) (and (3-39)) derived above applies to the Heisenberg, not Schrödinger, picture.

This is a parallel path to do second quantization that is included in the route represented by the vertical arrow on the RHS of Fig. 3-2, but it uses a different, though related, part of the theory.

3.2.4 Summary Chart

All that we have done in this Sect. 3.2, and what we will do in the remainder of this and the next two chapters, is summarized in Wholeness Chart 5-4 at the end of Chap. 5.

Note the summary is at the end of Chap. 5 because each column in it lists the key components in the development of QFT for one of the three spin types (spin 0, ½, and 1), and we won't be doing the latter two until Chaps. 4 and 5.

Be sure to use the summary wholeness chart, as you study this chapter and the next two

You can follow along in the chart, as we develop the theory for scalars, by reading the blocks in the Spin 0 column. You may want to stick a Post-It on that page as a book marker, so you can easily flip to it as you read along in this, and the following two, chapters.

3.3 Commutation Relations: The Crux of QFT

We will soon see how the commutation relations encompassed in the second part of 2nd quantization, found in the last box in the right hand column of Wholeness Chart 2-5 of Chap. 2, pg. 31, and reproduced in (3-40) below, lie at the root of, and structure, all of QFT. For scalars, they are

$$\left[\phi^r(\mathbf{x},t), \pi_s(\mathbf{y},t)\right] = \phi^r\pi_s - \pi_s\phi^r = i\delta^r{}_s\,\delta(\mathbf{x}-\mathbf{y}) \qquad \left[\phi^r, \phi^s\right] = \left[\pi_r, \pi_s\right] = \left[\phi^r, \pi_s{}^\dagger\right] = 0\,. \tag{3-40}$$

2^{nd} *quantization commutation relations determine coefficient commutation*

Note, in passing, that a complex conjugate of a field is considered a different field. In effect, if $\phi^r = \phi$, then $\phi^s = \phi^\dagger$, where $r \neq s$, so that $\left[\phi, \phi^\dagger\right] = 0$.

Of overriding importance in the theory, as we will see, are the following coefficient commutation relations, which we will derive below from the 2nd quantization postulate of (3-40).

$$\boxed{\left[a(\mathbf{k}), a^\dagger(\mathbf{k}')\right] = \left[b(\mathbf{k}), b^\dagger(\mathbf{k}')\right] = \delta_{\mathbf{k}\mathbf{k}'} \text{ (discrete)}; \quad = \delta(\mathbf{k}-\mathbf{k}') \text{ (continuous)}}\,. \tag{3-41}$$

Coefficient com rels 1) play fundamental role in QFT, and 2) imply coefficients are operators

The form of (3-41) should tell us immediately that the Klein-Gordon solution coefficients $a(\mathbf{k})$, $b(\mathbf{k})$, etc. in QFT of (3-36) and (3-37) are far different animals than the $A_\mathbf{k}$, $B_\mathbf{k}$, etc. in RQM of (3-12). The latter are merely numbers, which commute. We must, therefore, suspect that the $a(\mathbf{k})$, $b(\mathbf{k})$, etc. are operators, and as we will see, this suspicion will turn out to be correct.

Proof of coefficient commutation relations

To prove (3-41), start with (3-40) and take different spatial coordinates **x** and **y**, but the same time coordinate *t*, for ϕ and $\pi^0{}_0$. This results in the equal time commutation relations

$$\left[\phi(\mathbf{x},t)\pi_0^0(\mathbf{y},t)-\pi_0^0(\mathbf{y},t)\phi(\mathbf{x},t)\right]=\left[\phi(\mathbf{x},t)\dot{\phi}^\dagger(\mathbf{y},t)-\dot{\phi}^\dagger(\mathbf{y},t)\phi(\mathbf{x},t)\right]=i\delta(\mathbf{x}-\mathbf{y}), \qquad (3\text{-}42)$$

Proving coefficient commutation relations

which are only important at this point as a step in our proof. Then, plugging the discrete solutions (3-36) into the middle part of (3-42), where to save space we use the compressed notation $a_{\mathbf{k}} = a(\mathbf{k})$, etc., we get

$$\left(\sum_{\mathbf{k}}\frac{1}{\sqrt{2V\omega_{\mathbf{k}}}}a_{\mathbf{k}}e^{-i(\omega_{\mathbf{k}}t-\mathbf{k}\bullet\mathbf{x})}+\sum_{\mathbf{k}}\frac{1}{\sqrt{2V\omega_{\mathbf{k}}}}b_{\mathbf{k}}^\dagger e^{i(\omega_{\mathbf{k}}t-\mathbf{k}\bullet\mathbf{x})}\right)\left(\sum_{\mathbf{k}'}\frac{-i\omega_{\mathbf{k}'}}{\sqrt{2V\omega_{\mathbf{k}'}}}b_{\mathbf{k}'}e^{-i(\omega_{\mathbf{k}'}t-\mathbf{k}'\bullet\mathbf{y})}+\sum_{\mathbf{k}'}\frac{i\omega_{\mathbf{k}'}}{\sqrt{2V\omega_{\mathbf{k}'}}}a_{\mathbf{k}'}^\dagger e^{i(\omega_{\mathbf{k}'}t-\mathbf{k}'\bullet\mathbf{y})}\right)$$

$$-\left(\sum_{\mathbf{k}'}\frac{-i\omega_{\mathbf{k}'}}{\sqrt{2V\omega_{\mathbf{k}'}}}b_{\mathbf{k}'}e^{-i(\omega_{\mathbf{k}'}t-\mathbf{k}'\bullet\mathbf{y})}+\sum_{\mathbf{k}'}\frac{i\omega_{\mathbf{k}'}}{\sqrt{2V\omega_{\mathbf{k}'}}}a_{\mathbf{k}'}^\dagger e^{i(\omega_{\mathbf{k}'}t-\mathbf{k}'\bullet\mathbf{y})}\right)\left(\sum_{\mathbf{k}}\frac{1}{\sqrt{2V\omega_{\mathbf{k}}}}a_{\mathbf{k}}e^{-i(\omega_{\mathbf{k}}t-\mathbf{k}\bullet\mathbf{x})}+\sum_{\mathbf{k}}\frac{1}{\sqrt{2V\omega_{\mathbf{k}}}}b_{\mathbf{k}}^\dagger e^{i(\omega_{\mathbf{k}}t-\mathbf{k}\bullet\mathbf{x})}\right)$$

$$=\sum_{\mathbf{k}}\sum_{\mathbf{k}'}\frac{i\omega_{\mathbf{k}'}}{2V\sqrt{\omega_{\mathbf{k}}\omega_{\mathbf{k}'}}}\begin{pmatrix}-a_{\mathbf{k}}b_{\mathbf{k}'}e^{-i(\omega_{\mathbf{k}}+\omega_{\mathbf{k}'})t}e^{i(\mathbf{k}\bullet\mathbf{x}+\mathbf{k}'\bullet\mathbf{y})}+a_{\mathbf{k}}a_{\mathbf{k}'}^\dagger e^{-i(\omega_{\mathbf{k}}-\omega_{\mathbf{k}'})t}e^{i(\mathbf{k}\bullet\mathbf{x}-\mathbf{k}'\bullet\mathbf{y})}\\-b_{\mathbf{k}}^\dagger b_{\mathbf{k}'}e^{i(\omega_{\mathbf{k}}-\omega_{\mathbf{k}'})t}e^{-i(\mathbf{k}\bullet\mathbf{x}-\mathbf{k}'\bullet\mathbf{y})}+b_{\mathbf{k}}^\dagger a_{\mathbf{k}'}^\dagger e^{i(\omega_{\mathbf{k}}+\omega_{\mathbf{k}'})t}e^{-i(\mathbf{k}\bullet\mathbf{x}+\mathbf{k}'\bullet\mathbf{y})}\\+b_{\mathbf{k}'}a_{\mathbf{k}}e^{-i(\omega_{\mathbf{k}}+\omega_{\mathbf{k}'})t}e^{i(\mathbf{k}\bullet\mathbf{x}+\mathbf{k}'\bullet\mathbf{y})}+b_{\mathbf{k}'}b_{\mathbf{k}}^\dagger e^{i(\omega_{\mathbf{k}}-\omega_{\mathbf{k}'})t}e^{-i(\mathbf{k}\bullet\mathbf{x}-\mathbf{k}'\bullet\mathbf{y})}\\-a_{\mathbf{k}'}^\dagger a_{\mathbf{k}}e^{-i(\omega_{\mathbf{k}}-\omega_{\mathbf{k}'})t}e^{i(\mathbf{k}\bullet\mathbf{x}-\mathbf{k}'\bullet\mathbf{y})}-a_{\mathbf{k}'}^\dagger b_{\mathbf{k}}^\dagger e^{i(\omega_{\mathbf{k}}+\omega_{\mathbf{k}'})t}e^{-i(\mathbf{k}\bullet\mathbf{x}+\mathbf{k}'\bullet\mathbf{y})}\end{pmatrix}=i\delta(\mathbf{x}-\mathbf{y}). \qquad (3\text{-}43)$$

K-G solutions into equal time commutator

Using the math identity for the 3D Dirac delta function

$$\delta(\mathbf{x}-\mathbf{y})=\frac{1}{V}\sum_{n=-\infty}^{+\infty}e^{-i\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{y})}\quad\left(=\frac{1}{V}\sum_{\mathbf{k}}e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{y})}\text{ in our notation}\right) \qquad (3\text{-}44)$$

Re-express Dirac delta function

after the last equal sign in (3-43), and matching terms, we see that all terms where $|\mathbf{k}| \neq |\mathbf{k}'|$ must equal zero. That is,

$$\frac{i\omega_{\mathbf{k}'}}{2V\sqrt{\omega_{\mathbf{k}}\omega_{\mathbf{k}'}}}\underbrace{\left(b_{\mathbf{k}'}a_{\mathbf{k}}-a_{\mathbf{k}}b_{\mathbf{k}'}\right)}_{\text{must}=0}e^{-i(\omega_{\mathbf{k}}+\omega_{\mathbf{k}'})t}e^{i(\mathbf{k}\bullet\mathbf{x}+\mathbf{k}'\bullet\mathbf{y})}+$$

$$\frac{i\omega_{\mathbf{k}'}}{2V\sqrt{\omega_{\mathbf{k}}\omega_{\mathbf{k}'}}}\underbrace{\left(a_{\mathbf{k}}a_{\mathbf{k}'}^\dagger-a_{\mathbf{k}'}^\dagger a_{\mathbf{k}}\right)}_{\text{must}=0}e^{-i(\omega_{\mathbf{k}}-\omega_{\mathbf{k}'})t}e^{i(\mathbf{k}\bullet\mathbf{x}-\mathbf{k}'\bullet\mathbf{y})}+ \qquad (3\text{-}45)$$

... etc. for all terms in **k** and **k′** with $\mathbf{k}\neq\pm\mathbf{k}'$ (no terms on RHS in **k** and **k′**)

All terms where $|\mathbf{k}| \neq |\mathbf{k}'|$

These must vanish, so their commutators must = 0

So, all possible coefficient commutators with $\mathbf{k} \neq \mathbf{k}'$ or $-\mathbf{k}'$ vanish. The remaining terms all have $|\mathbf{k}| = |\mathbf{k}'|$, which means $\omega_{\mathbf{k}} = \omega_{\mathbf{k}'}$. Some of these have an exponential form $i(\omega_{\mathbf{k}} + \omega_{\mathbf{k}'})t$, and thus

$$\frac{i}{2V}\frac{\omega_{\mathbf{k}}}{\omega_{\mathbf{k}}}\underbrace{\left(b_{\mathbf{k}}a_{\mathbf{k}}-a_{\mathbf{k}}b_{\mathbf{k}}\right)}_{\text{must}=0}\underbrace{e^{-i2\omega_{\mathbf{k}}t}}_{\neq 0}e^{i\mathbf{k}\bullet(\mathbf{x}+\mathbf{y})}+$$

$$\frac{i}{2V}\underbrace{\left(b_{\mathbf{k}}^\dagger a_{\mathbf{k}}^\dagger-a_{\mathbf{k}}^\dagger b_{\mathbf{k}}^\dagger\right)}_{\text{must}=0}e^{i2\omega_{\mathbf{k}}t}e^{-i\mathbf{k}\bullet(\mathbf{x}+\mathbf{y})}+ \qquad (3\text{-}46)$$

... etc. for all similar terms = 0 (no time dependence on RHS).

Some of terms where $|\mathbf{k}| = |\mathbf{k}'|$*, i.e., those of form* $exp(i(\omega_{\mathbf{k}}+\omega_{\mathbf{k}'})t)$

For these terms, the coefficient commutators must vanish because the exponential in $\omega_{\mathbf{k}}$ varies in time, whereas there is no such variation on the RHS.

Commutators must = 0

The remaining terms have exponential form $i(\omega_{\mathbf{k}} - \omega_{\mathbf{k}'})t$ and $|\mathbf{k}| = |\mathbf{k}'|$. Adding the terms for given **k** with the terms for $-\mathbf{k}$ yields

$$\frac{i}{2V}\underbrace{\left(a_{\mathbf{k}}a_{\mathbf{k}}^{\dagger}-a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}}\right)}_{\text{must=1}}\underbrace{e^{-i(\omega_{\mathbf{k}}-\omega_{\mathbf{k}})t}}_{1}e^{i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})}+\frac{i}{2V}\underbrace{\left(a_{\mathbf{k}}a_{\mathbf{k}}^{\dagger}-a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}}\right)}_{\text{must=1}}e^{-i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})}+$$
$$\frac{i}{2V}\underbrace{\left(b_{\mathbf{k}}b_{\mathbf{k}}^{\dagger}-b_{\mathbf{k}}^{\dagger}b_{\mathbf{k}}\right)}_{\text{must=1}}e^{-i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})}+\frac{i}{2V}\underbrace{\left(b_{\mathbf{k}}b_{\mathbf{k}}^{\dagger}-b_{\mathbf{k}}^{\dagger}b_{\mathbf{k}}\right)}_{\text{must=1}}e^{i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})}=\frac{i}{V}e^{-i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})}+\frac{i}{V}e^{i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})}\ , \tag{3-47}$$

Remaining terms where $|\mathbf{k}| = |\mathbf{k}'|$, i.e., those of form $exp(i(\omega_{\mathbf{k}} - \omega_{\mathbf{k}'})t)$

Commutators must = 1

where we see that the coefficient commutators need to equal unity in order to match corresponding terms on the RHS. (A subtle point for the purists: we also made the reasonable assumption that the commutators in $a_{\mathbf{k}}\,a_{\mathbf{k}}^{\dagger}$ and the commutators in $b_{\mathbf{k}}\,b_{\mathbf{k}}^{\dagger}$ have the same value.)

The commutation relations shown above in (3-45) to (3-47) are the same as (3-41). QED.

If you are ambitious, have extra time, and/or simply have to prove everything to yourself, do Prob. 7 to derive the continuous solution commutators of (3-41).

End of coefficient commutation relations proof

With the coefficient commutator relations in hand, we are finally ready to dive into the real core of QFT.

3.4 The Hamiltonian in QFT

We find the Hamiltonian by integrating the Hamiltonian density $\mathcal{H}$ over all space (a volume V containing the discrete solutions, which we can make as large as we like.) In QFT, we express $\mathcal{H}$ in terms of a complex field and substitute our field equation solutions.

$H = \int \mathcal{H}\,dV$

3.4.1 The Free Scalar Hamiltonian in Terms of the Coefficients

For a free scalar field $\mathcal{H} = \mathcal{H}_0{}^0$, as in (3-33), where we employ our discrete, plane wave solutions (3-36) we get

$H = \int \mathcal{H}\,dV$ in terms of the fields

$$H_0^0=\int\mathcal{H}_{\text{o}}^{\text{o}}d^3x=\int\left(\dot{\phi}\dot{\phi}^{\dagger}+\nabla\phi^{\dagger}\cdot\nabla\phi+\mu^2\phi^{\dagger}\phi\right)d^3x=$$
$$\int\left(\sum_{\mathbf{k}}\frac{\partial}{\partial t}\frac{1}{\sqrt{2V\omega_{\mathbf{k}}}}\left(a(\mathbf{k})e^{-ikx}+b^{\dagger}(\mathbf{k})e^{ikx}\right)\right)\left(\sum_{\mathbf{k}'}\frac{\partial}{\partial t}\frac{1}{\sqrt{2V\omega_{\mathbf{k}'}}}\left(b(\mathbf{k}')e^{-ik'x}+a^{\dagger}(\mathbf{k}')e^{ik'x}\right)\right)d^3x \tag{3-48}$$
$$+\int\left(-\partial_i\phi^{\dagger}\partial^i\phi+\mu^2\phi^{\dagger}\phi\right)d^3x.$$

Deriving H in terms of the coefficients ↓

The middle line of (3-48), i.e., the $\int\dot{\phi}\dot{\phi}^{\dagger}d^3x$ part, becomes

$$\int\left(\sum_{\mathbf{k}}\frac{i\omega_{\mathbf{k}}}{\sqrt{2V\omega_{\mathbf{k}}}}\left(-a(\mathbf{k})e^{-ikx}+b^{\dagger}(\mathbf{k})e^{ikx}\right)\right)\left(\sum_{\mathbf{k}'}\frac{i\omega_{\mathbf{k}'}}{\sqrt{2V\omega_{\mathbf{k}'}}}\left(-b(\mathbf{k}')e^{-ik'x}+a^{\dagger}(\mathbf{k}')e^{ik'x}\right)\right)d^3x. \tag{3-49}$$

or

$$\sum_{\mathbf{k}}\sum_{\mathbf{k}'}\frac{-\sqrt{\omega_{\mathbf{k}}}\sqrt{\omega_{\mathbf{k}'}}}{2V}\int\begin{pmatrix}a(\mathbf{k})b(\mathbf{k}')e^{-ikx}e^{-ik'x}-a(\mathbf{k})a^{\dagger}(\mathbf{k}')e^{-ikx}e^{ik'x}\\-b^{\dagger}(\mathbf{k})b(\mathbf{k}')e^{ikx}e^{-ik'x}+b^{\dagger}(\mathbf{k})a^{\dagger}(\mathbf{k}')e^{ikx}e^{ik'x}\end{pmatrix}d^3x\,. \tag{3-50}$$

The sum over $\mathbf{k}$ and $\mathbf{k}'$ is from negative infinity to positive infinity in the x, y, and z directions.

All terms in the integration in (3-50) result in zero except when $\mathbf{k}' = \mathbf{k}$ or $\mathbf{k}' = -\mathbf{k}$, because we are integrating orthogonal functions between their boundaries. (This is similar to $sin(2X)sin(4X)$ integrated with respect to X along a complete number of wavelengths, where here $\mathbf{k} = 2$ and $\mathbf{k}' = 4$.) Since the volume of integration in (3-50) equals V, we end up with

$$\int\dot{\phi}\dot{\phi}^{\dagger}d^3x=\sum_{\mathbf{k}}\frac{\omega_{\mathbf{k}}}{2}\left(-a(\mathbf{k})b(-\mathbf{k})e^{-2i\omega t}+a(\mathbf{k})a^{\dagger}(\mathbf{k})+b^{\dagger}(\mathbf{k})b(\mathbf{k})-b^{\dagger}(\mathbf{k})a^{\dagger}(-\mathbf{k})e^{2i\omega t}\right)$$
$$=\sum_{\mathbf{k}}\frac{(\omega_{\mathbf{k}})^2}{2\omega_{\mathbf{k}}}\left(-a(-\mathbf{k})b(\mathbf{k})e^{-2i\omega t}+a(\mathbf{k})a^{\dagger}(\mathbf{k})+b^{\dagger}(\mathbf{k})b(\mathbf{k})-b^{\dagger}(-\mathbf{k})a^{\dagger}(\mathbf{k})e^{2i\omega t}\right). \tag{3-51}$$

In the second row of (3-51), the sign change on the **k** in the first and last terms is justified since we are summing over all **k**, so for every term with **k** in it, there is another with – **k**. This modification will make things easier a bit later.

Following similar steps for the next term in (3-48) we get

$$\begin{aligned}
&-\int \partial_i\phi^\dagger\partial^i\phi d^3x = \int \partial_i\phi^\dagger\partial_i\phi d^3x \\
&= \int\left(\sum_{\mathbf{k}}\frac{ik_i}{\sqrt{2V\omega_{\mathbf{k}}}}\left(b(\mathbf{k})e^{-ikx}-a^\dagger(\mathbf{k})e^{ikx}\right)\right)\left(\sum_{\mathbf{k}'}\frac{ik_i'}{\sqrt{2V\omega_{\mathbf{k}'}}}\left(a(\mathbf{k}')e^{-ik'x}-b^\dagger(\mathbf{k}')e^{ik'x}\right)\right)d^3x \qquad (3\text{-}52)\\
&= \sum_{\mathbf{k}}\frac{\mathbf{k}^2}{2\omega_{\mathbf{k}}}\left(b(\mathbf{k})a(-\mathbf{k})e^{-2i\omega t}+a^\dagger(\mathbf{k})a(\mathbf{k})+b(\mathbf{k})b^\dagger(\mathbf{k})+a^\dagger(\mathbf{k})b^\dagger(-\mathbf{k})e^{2i\omega t}\right)
\end{aligned}$$

where we note that terms in the summation with both **k** and – **k** have an extra sign change since $k_i = -k_i'$ in the multiplication in the second line of (3-52).

Similarly, for the mass term in (3-48), we get (do Prob. 8 at the end of the chapter to prove it)

$$\int \mu^2\phi^\dagger\phi d^3x = \sum_{\mathbf{k}}\frac{\mu^2}{2\omega_{\mathbf{k}}}\left(b(\mathbf{k})a(-\mathbf{k})e^{-2i\omega t}+b(\mathbf{k})b^\dagger(\mathbf{k})+a^\dagger(\mathbf{k})a(\mathbf{k})+a^\dagger(\mathbf{k})b^\dagger(-\mathbf{k})e^{2i\omega t}\right). \quad (3\text{-}53)$$

Adding the final parts of (3-51), (3-52), and (3-53), and using $\mathbf{k}^2+\mu^2=(\omega_{\mathbf{k}})^2$ along with the coefficient commutation relations (3-41), we end up with

$$\begin{aligned}
H_0^0 &= \sum_{\mathbf{k}}\frac{\omega_{\mathbf{k}}}{2}\left(\underbrace{a(\mathbf{k})a^\dagger(\mathbf{k})}_{\text{use commutator}}+a^\dagger(\mathbf{k})a(\mathbf{k})+b^\dagger(\mathbf{k})b(\mathbf{k})+\underbrace{b(\mathbf{k})b^\dagger(\mathbf{k})}_{\text{use commutator}}\right) \\
&= \sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(a^\dagger(\mathbf{k})a(\mathbf{k})+\tfrac{1}{2}+b^\dagger(\mathbf{k})b(\mathbf{k})+\tfrac{1}{2}\right).
\end{aligned} \qquad (3\text{-}54)$$

or simply

$$\boxed{H_0^0=\sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(N_a(\mathbf{k})+\tfrac{1}{2}+N_b(\mathbf{k})+\tfrac{1}{2}\right)}, \qquad (3\text{-}55)$$

$H=\int\mathcal{H}dV$ in terms of the coefficients

where

$$\boxed{N_a(\mathbf{k})=a^\dagger(\mathbf{k})a(\mathbf{k}) \qquad N_b(\mathbf{k})=b^\dagger(\mathbf{k})b(\mathbf{k})}. \qquad (3\text{-}56)$$

Expressions (3-55) and (3-56) lie at the heart of QFT, as we are about to see.

3.4.2 Number Operators

Consider what we must get if the Hamiltonian of (3-55) operates on a state (a ket) comprised of two free scalar particles, each in the same eigenstate of energy $\omega_{\mathbf{k}_1}$. We would expect that (multiparticle) state would have an energy eigenvalue equal to its total energy $2\omega_{\mathbf{k}_1}$ i.e.,

$$H_0^0\left|2\phi_{\mathbf{k}_1}\right\rangle = 2\omega_{\mathbf{k}_1}\left|2\phi_{\mathbf{k}_1}\right\rangle . \qquad (3\text{-}57)$$

But from (3-55), that means

$$\sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(N_a(\mathbf{k})+\tfrac{1}{2}+N_b(\mathbf{k})+\tfrac{1}{2}\right)\left|2\phi_{\mathbf{k}_1}\right\rangle = 2\omega_{\mathbf{k}_1}\left|2\phi_{\mathbf{k}_1}\right\rangle . \qquad (3\text{-}58)$$

How can we make sense of (3-58)? The answer is that it is not quite true, and that we can make sense of it all if, instead of (3-57) and (3-58), we consider

$$H_0^0\left|2\phi_{\mathbf{k}_1}\right\rangle = \sum_{\mathbf{k}}\omega_{\mathbf{k}}\Big(\underbrace{N_a(\mathbf{k})+\tfrac{1}{2}}_{\substack{a\text{ type}\\ \text{particles}}}+\underbrace{N_b(\mathbf{k})+\tfrac{1}{2}}_{\substack{b\text{ type}\\ \text{particles}}}\Big)\underbrace{\left|2\phi_{\mathbf{k}_1}\right\rangle}_{\substack{a\text{ type}\\ \text{particles}}} = \left(2\omega_{\mathbf{k}_1}+\sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(\tfrac{1}{2}+\tfrac{1}{2}\right)\right)\left|2\phi_{\mathbf{k}_1}\right\rangle, \qquad (3\text{-}59)$$

$H=\int\mathcal{H}dV$ in terms of number operators

with the following interpretation.

$N_a(\mathbf{k})$ = number operator with eigenvalue $n_a(\mathbf{k})$ = number of *a* particles with 3-mom **k** in the ket,
$N_b(\mathbf{k})$ = number operator with eigenvalue $n_b(\mathbf{k})$ = number of *b* particles with 3-mom **k** in the ket,
and, the vacuum has ½ quantum of energy for each **k** for *a* particles, and also for *b* particles.

This might, at first, be considered a separate postulate, but if the $\mathcal{H}_0{}^0$ derived by 2nd quantization for quantum scalar fields is correct, this is the only possible interpretation of (3-59) that works. The part about the vacuum would be surprising to anyone who had not already heard that the vacuum is a seething caldron of virtual quanta. More on this shortly.

For $H = \int \mathcal{H}\, dV$ to work, the vacuum must be filled with ½ quanta

We also anticipate that the *b* type particles will be antiparticles, and the *a* types, normal particles. More on this later, as well.

Examples of number operators and kets

In light of the above, the following examples should be relatively straightforward. Note we designate *b* type particles with an overbar.

Example #1: 10 particle state

$$H_0^0\,|5\phi_{\mathbf{k}_1},2\phi_{\mathbf{k}_2},\underbrace{3\bar{\phi}_{\mathbf{k}_3}}_{\substack{b\text{ type}\\ \text{particles}}}\rangle = \sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(N_a(\mathbf{k})+\tfrac{1}{2}+N_b(\mathbf{k})+\tfrac{1}{2}\right)\left|5\phi_{\mathbf{k}_1},2\phi_{\mathbf{k}_2},3\bar{\phi}_{\mathbf{k}_3}\right\rangle$$
$$=\left(5\omega_{\mathbf{k}_1}+2\omega_{\mathbf{k}_2}+3\omega_{\mathbf{k}_3}+\sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(\tfrac{1}{2}+\tfrac{1}{2}\right)\right)\left|5\phi_{\mathbf{k}_1},2\phi_{\mathbf{k}_2},3\bar{\phi}_{\mathbf{k}_3}\right\rangle. \qquad (3\text{-}60)$$

Every state has vacuum as part of it, so every state has huge vacuum energy

and so we see that *b* type particles in our theory have positive energy. This resulted from our interpretation of $H_0{}^0$ in (3-59), i.e., from our postulate for the properties of $N_b(\mathbf{k})$ in (3-59) and the box following it. We know from experiment that antiparticles have positive energy, and this interpretation will lead to *b* particles filling the role of antiparticles. We will see later that *b* particles have opposite charge from the *a* particles, and so they fit nicely into our theory as antiparticles.

Example #2: Vacuum state

$$H_0^0\ \underbrace{|0\rangle}_{\substack{\text{vacuum}\\ \text{state}}} = \sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(N_a(\mathbf{k})+\tfrac{1}{2}+N_b(\mathbf{k})+\tfrac{1}{2}\right)|0\rangle = \underbrace{\sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(\tfrac{1}{2}+\tfrac{1}{2}\right)}_{\text{infinite energy}}|0\rangle. \qquad (3\text{-}61)$$

Note that every state is superimposed on the vacuum, so every state actually has infinite energy. We saw this in Example #1.

3.4.3 Zero Point (Vacuum) Energy

Zero point (vacuum) energy results from 2nd quantization postulate of non-commutation

The infinite sum of ½ quanta in (3-61) represents the now famous perspective on the vacuum as being almost inconceivably crammed with energy, known as zero point energy (ZPE.) In actuality, the sum, while enormous, is usually not considered infinite, but for reasons beyond the scope of our current discussion, to terminate at a very high level, known as the Planck scale.

It is important to recognize that this *vacuum energy arose from our postulate of 2nd quantization*, that a field and its conjugate momentum don't commute (see (3-40). Because of this we got the coefficient commutation relations (3-41), and those were used in our derivation of the form of the Hamiltonian (see (3-54), which resulted in the appearance of $\frac{1}{2}\omega_{\mathbf{k}}$ terms.

Non-commutation leads to uncertainty. Source of statements that uncertainty principle results in ZPE

The field commutation relations of QFT are siblings to the particle commutation relations for NRQM and RQM. (See Wholeness Chart 2-5 in Chap. 2, pg. 30.) In the latter, particle position and momentum do not commute and this results in the renowned Heisenberg uncertainty relation between position and momentum. In QFT, the field and its conjugate momentum do not commute, implying a parallel uncertainty relationship between them. This is the source of the statements, often heard, that the uncertainty principle is the cause of the zero point energy.

Are the ½ Quanta Effervescent?

Ruminations on vacuum effervescence

One often also hears that the ½ quanta "pop" in and out of the vacuum effervescently in particle/antiparticle pairs. I submit that this is a heuristic, at best, representation for the popular media. According to (3-55), there is no "popping", no evanescent physical reality alternated with nothingness, no pairing of particles at a creation event followed by a common mutual destruction event as one might see in a Feynman diagram (different from, but similar in nature to, what we saw in Fig. 1-1 of Chap. 1, pg. 2) Via our fundamental QFT relation for $H_0{}^0$, the ½ quanta are simply

sitting in the vacuum. They may be virtual in some sense, and not real, but the QFT *H* operator does not suggest any intermittent sort of existence.

Further, our derivation of $H_0{}^0$ has been exclusively for free fields, where no interactions are included (more on interactions in later chapters). The particles (and antiparticles), for which $H_0{}^0$ determines the energy, do not interact with other particles or antiparticles. This means they can't create or destroy in pairs, since that is, above all, an interaction between the associated particles and antiparticles. So $H_0{}^0$ specifically does not measure the energy of such pairs and the ½ energy terms therein must be for free fields that are not "popping" in and out of the vacuum in pairs.

Casimir plates do not prove ZPE existence

Some argue that experimental measurements of Casimir plate forces, the Lamb shift, and the anomalous magnetic moment of the electron demonstrate the existence of vacuum fluctuations. However, Casimir forces are generally computed by considering the ½ quanta such as those in (3-61) to be standing waves between the plates. That is, they do *not need*, in those analyses, to be "popping" in and out of the vacuum, but merely be sitting there continually. Further, the Casimir force can be computed without reference to zero-point energies at all, and thus may not be the conclusive proof for their existence it is widely taken to be. (See R. L. Jaffe, "Casimir Effect and the Quantum Vacuum", Phys. Rev. D72 021301(R) (2005) http://arxiv.org/abs/hep-th/0503158 .)

Neither are Lamb shift nor anomalous magnetic moment solution due to free field ½ quanta (*ZPE*)

The Lamb shift calculation involves so-called "vacuum fluctuations", but they are actually higher order corrections to propagators for interacting fields (which we study in later chapters), not the ½ quanta vacuum energy of free fields (which we study in this chapter.) The same is true of the QFT correction to the magnetic moment of the electron (the famed "anomalous magnetic moment".)

As a caveat, I note that the remarks in this sub-section entitled "Are the ½ Quanta Effervescent" reflect my personal position on vacuum energy. The majority of physicists believe quanta are continually "bubbling" in and out of existence in the vacuum. I simply have not seen a sound derivation of this in the literature, and don't believe it is supported by the formal derivation of ZPE.

We will discuss this issue further when we get to interaction theory. (See Chap. 10.)

3.4.4 Positive Energy in QFT

$H = \int \mathcal{H}\, dV$ leads to positive energy for both QFT particle types

Note that unlike RQM, all particles in QFT have positive energy. The QFT energy operator $H_0{}^0$ operating on states yields positive eigenvalues for both *a* and *b* types of particles. The RQM energy operator operating on states did not do that, as we saw in Sect. 3.1.5 (pg. 47).

Continuation of Wholeness Chart 1-2. Comparison of Three Quantum Theories

	NRQM	**RQM**	**QFT**
Hamiltonian	$i\frac{\partial}{\partial t}$	$i\frac{\partial}{\partial t}$	$H = \int \mathcal{H}\, d^3x$
Sign of Energy *E*	positive	positive & negative	positive

3.4.5 Unit Norms and Orthogonality for Multiparticle States

Recall from NRQM, that it was advantageous to normalize states, i.e., change the constant in front of the ket such that the inner product of the state and its complex conjugate transpose (the bracket of the bra and ket) equaled unity. That is, we defined

$$\langle \phi_{\mathbf{k}} || \phi_{\mathbf{k}} \rangle = \int_V \underbrace{\phi_{\mathbf{k}}{}^{\dagger}(\mathbf{x},t)\phi_{\mathbf{k}}(\mathbf{x},t)}_{\text{states}} d^3x = 1 \quad \text{and} \quad \langle \phi_{\mathbf{k}} || \phi_{\mathbf{k}'} \rangle = \int_V \underbrace{\phi_{\mathbf{k}}{}^{\dagger}(\mathbf{x},t)\phi_{\mathbf{k}'}(\mathbf{x},t)}_{\text{states}} d^3x = 0, \ \mathbf{k} \neq \mathbf{k}' . \quad (3\text{-}62)$$

As aside, note that in (3-62), it is assumed that the ket is expressed in the position basis. For example, a plane wave momentum eigenstate in that basis would have form[4]

---

[4] The ket symbol $|\phi\rangle$ in general represents a particle state, but the form of the ket when we write it out mathematically, such as we did in the RHS of (3-63), changes with the basis we care to use. For example we could express the ket in the momentum basis (in momentum space) instead of the position basis *x*; or in a number of other ways. Mathematically, (3-63) is really $|\phi\rangle_{x\,\text{basis}} = \langle x|\phi\rangle = Ae^{-i(Et-\mathbf{k}\cdot\mathbf{x})}$.

$$\left|\phi_{\mathbf{k}}\right\rangle = Ae^{-i(Et-\mathbf{k}\cdot\mathbf{x})} \qquad \left(\left|\phi_{\mathbf{k}}\right\rangle \text{ here is expressed in the position basis}\right). \tag{3-63}$$

In this book, unless otherwise stated, when we express $|\phi\rangle$ mathematically, we will assume the position basis as in (3-63). At such times, $|\phi\rangle$ will = $|\phi\rangle_{\text{x basis}}$. Now, back to the main point.

In NRQM and RQM, states are single particle states. In QFT, they are typically multiparticle, but we will also find it advantageous therein to normalize. So, we define our symbols for multiparticle states so that every such state is normalized (i.e, has unit norm.) For example, for a state comprising two *a* particles of 3-momentum **k**, one of 3-momentum **k**′, and five *b* particles of 3-momentum **k**″, we would have (where the middle part is just a reminder of what we mean by the bracket notation[5])

Multiparticle states in QFT have unit norm

$$\left\langle 2\phi_{\mathbf{k}},\phi_{\mathbf{k}'},5\bar{\phi}_{\mathbf{k}''}\right\|\left.2\phi_{\mathbf{k}},\phi_{\mathbf{k}'},5\bar{\phi}_{\mathbf{k}''}\right\rangle = \int_V \left(\underbrace{2\phi_{\mathbf{k}}\phi_{\mathbf{k}'}5\bar{\phi}_{\mathbf{k}''}}_{\text{states}}\right)^{\dagger}\left(\underbrace{2\phi_{\mathbf{k}}\phi_{\mathbf{k}'}5\bar{\phi}_{\mathbf{k}''}}_{\text{states}}\right)d^3x = 1\ . \tag{3-64}$$

Note that any (multiparticle) state is orthogonal to every other state that is not identical to it in particle types, particle numbers, and **k** values for each. For examples,

Different multiparticle states in QFT are orthogonal

$$\left\langle 2\phi_{\mathbf{k}},\phi_{\mathbf{k}'},5\bar{\phi}_{\mathbf{k}''}\right\|\left.\phi_{\mathbf{k}'},5\bar{\phi}_{\mathbf{k}''}\right\rangle = 0 \qquad \left\langle 2\phi_{\mathbf{k}}\right\|\left.\phi_{\mathbf{k}}\right\rangle = 0 \qquad \left\langle 5\phi_{\mathbf{k}''}\right\|\left.5\bar{\phi}_{\mathbf{k}''}\right\rangle = 0 \qquad \left\langle \phi_{\mathbf{k}}\right\|\left.\phi_{\mathbf{k}'}\right\rangle = 0\ . \tag{3-65}$$

Note on Notation

It is common practice in QFT to employ the bracket notation of LHS of (3-64), and virtually never, the integral form shown between the equal signs. As noted earlier, in QFT, symbols such as $\phi_{\mathbf{k}}$, which are not part of a ket symbol, normally do not represent states, but operators/fields. When they play the role of states, as in (3-62), we must label them specifically, as there and in (3-64).

QFT virtually never uses $\phi_{\mathbf{k}}$ to represent a state, nor an integral to represent an inner product of states

3.5 Expectation Values and the Hamiltonian

Note that the expectation value relation for an operator $\mathcal{O}$ in QFT for a single particle state is the same as that in the rest of quantum mechanics, i.e., provided $|\phi\rangle$ has unit norm,

Expectation value for single particle in QFT like those in NRQM and RQM

$$\bar{\mathcal{O}} = \langle\phi|\mathcal{O}|\phi\rangle\ . \tag{3-66}$$

where we don't confuse the overbar (used here outside a bra or ket) for expectation (or average) value with its use inside a bra or ket, where it signifies *b* type particles. The expectation value is the average value we would measure over a large number of measurements of the state. If the particle is in an eigenstate of an observable (an operator), then every measurement of that observable for that state would be the same (the eigenvalue), and thus equal to the expectation value. An eigenstate of energy would measure the same value for energy upon every measurement. This is true for a single particle state, such as in (3-66).

It is also true for a multiparticle state, such as those we run into in QFT. For example, the multiparticle state in (3-60) is in an eigenstate of energy (the sum of the energies from each of the ten particles in the state.) Each particle therein has fixed mass plus a fixed momentum **k**, and hence fixed total energy. Thus, the energy expectation value for the state in (3-60) is

$$\begin{aligned}
\bar{H}_0^0 &= \left\langle 5\phi_{\mathbf{k}_1},2\phi_{\mathbf{k}_2},3\bar{\phi}_{\mathbf{k}_3}\right|H_0^0\left|5\phi_{\mathbf{k}_1},2\phi_{\mathbf{k}_2},3\bar{\phi}_{\mathbf{k}_3}\right\rangle = \\
&= \left\langle 5\phi_{\mathbf{k}_1},2\phi_{\mathbf{k}_2},3\bar{\phi}_{\mathbf{k}_3}\right|\underbrace{\left(5\omega_{\mathbf{k}_1}+2\omega_{\mathbf{k}_2}+3\omega_{\mathbf{k}_3}+\sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(\tfrac{1}{2}+\tfrac{1}{2}\right)\right)}_{\text{a number, not an operator, so can move outside}}\left|5\phi_{\mathbf{k}_1},2\phi_{\mathbf{k}_2},3\bar{\phi}_{\mathbf{k}_3}\right\rangle \\
&= \left(5\omega_{\mathbf{k}_1}+2\omega_{\mathbf{k}_2}+3\omega_{\mathbf{k}_3}+\sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(\tfrac{1}{2}+\tfrac{1}{2}\right)\right)\underbrace{\left\langle 5\phi_{\mathbf{k}_1},2\phi_{\mathbf{k}_2},3\bar{\phi}_{\mathbf{k}_3}\right|\left.5\phi_{\mathbf{k}_1},2\phi_{\mathbf{k}_2},3\bar{\phi}_{\mathbf{k}_3}\right\rangle}_{=1} \\
&= 5\omega_{\mathbf{k}_1}+2\omega_{\mathbf{k}_2}+3\omega_{\mathbf{k}_3}+\sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(\tfrac{1}{2}+\tfrac{1}{2}\right).
\end{aligned} \tag{3-67}$$

[5] For the purists, we note that (3-64) can be taken as an invariant relationship if we define our states properly. That is, we can define our multiparticle eigenstate with a factor of $1/\sqrt{V}$ in front, so the integrand in (3-64) yields a factor 1/*V*. The integral over 3D space then yields a factor of *V*. *V* is non-invariant, but one in the numerator and one in the denominator cancel, leaving an invariant final result.

Expectation value for multiparticle state has same form as single particle state

In general, for any operator $\mathcal{O}$, the expectation value for any multiparticle state is

$$\boxed{\bar{\mathcal{O}} = \langle \phi_1, \phi_2, \phi_3, ... | \mathcal{O} | \phi_1, \phi_2, \phi_3, ... \rangle} , \tag{3-68}$$

where we will typically find the operator expressed in terms of number operators.

Expectation value for the vacuum, VEV

A concept that becomes important later on in QFT is that of the vacuum expectation value, or simply the VEV, whose symbol and mathematical expression are

$$\langle \mathcal{O} \rangle = \langle 0 | \mathcal{O} | 0 \rangle . \tag{3-69}$$

If you don't see it right away, do Prob. 9 to prove to yourself that the VEV of the free field scalar Hamiltonian is

Hamiltonian VEV

$$\left\langle H_0^0 \right\rangle = \sum_{\mathbf{k}} \omega_{\mathbf{k}} . \tag{3-70}$$

We expect to measure infinite (or enormous, if nature has a maximum |**k**|) energy in the vacuum.

3.6 Creation and Destruction Operators

The coefficients are creation and destruction operators

In this section, we will prove what is perhaps the most fundamental aspect of QFT, which we foreshadowed in Chap. 1, that the Klein-Gordon solution coefficients $a(\mathbf{k})$, $a^\dagger(\mathbf{k})$, $b(\mathbf{k})$, and $b^\dagger(\mathbf{k})$ are not numbers, but operators that create and destroy particles. Since certain combinations of them do not commute, we should expect them to be operators of some kind.

3.6.1 Proving It

Proving it.

Proof that $a(\mathbf{k})$ is a Particle Destruction Operator

With the notation $|n_{\mathbf{k}}\rangle$ denoting a multiparticle state of $n_{\mathbf{k}}$ a type particles (no b types for now), all with the same 4-momentum k^μ, what can we say about the state

$$a(\mathbf{k}) | n_{\mathbf{k}} \rangle = | m_{\mathbf{k}} \rangle \, ? \tag{3-71}$$

To see, first operate on this state with our number operator $N_a(\mathbf{k}) = a^\dagger(\mathbf{k})\, a(\mathbf{k})$

$$N_a(\mathbf{k}) | m_{\mathbf{k}} \rangle = N_a(\mathbf{k}) a(\mathbf{k}) | n_{\mathbf{k}} \rangle = \underbrace{a^\dagger(\mathbf{k}) a(\mathbf{k})}_{\text{use commutator}} a(\mathbf{k}) | n_{\mathbf{k}} \rangle . \tag{3-72}$$

Then use the commutation relations from (3-41), to find (3-72) equals

$$\begin{aligned} &\left(a(\mathbf{k}) a^\dagger(\mathbf{k}) - 1\right) a(\mathbf{k}) | n_{\mathbf{k}} \rangle = a(\mathbf{k}) a^\dagger(\mathbf{k}) a(\mathbf{k}) | n_{\mathbf{k}} \rangle - a(\mathbf{k}) | n_{\mathbf{k}} \rangle = a(\mathbf{k}) N_a(\mathbf{k}) | n_{\mathbf{k}} \rangle - a(\mathbf{k}) | n_{\mathbf{k}} \rangle \\ &= a(\mathbf{k}) n_{\mathbf{k}} | n_{\mathbf{k}} \rangle - a(\mathbf{k}) | n_{\mathbf{k}} \rangle = n_{\mathbf{k}} a(\mathbf{k}) | n_{\mathbf{k}} \rangle - a(\mathbf{k}) | n_{\mathbf{k}} \rangle = (n_{\mathbf{k}} - 1) a(\mathbf{k}) | n_{\mathbf{k}} \rangle = (n_{\mathbf{k}} - 1) | m_{\mathbf{k}} \rangle . \end{aligned} \tag{3-73}$$

So

$$N_a(\mathbf{k}) | m_{\mathbf{k}} \rangle = (n_{\mathbf{k}} - 1) | m_{\mathbf{k}} \rangle = m_{\mathbf{k}} | m_{\mathbf{k}} \rangle \qquad m_{\mathbf{k}} = n_{\mathbf{k}} - 1 . \tag{3-74}$$

Since the number operator operating on $|m_{\mathbf{k}}\rangle$ gives a number of particles one less than it did when operating on $|n_{\mathbf{k}}\rangle$, the operation of $a(\mathbf{k})$ on $|n_{\mathbf{k}}\rangle$ in (3-71) reduces the number of particles in the state by one. We conclude that $a(\mathbf{k})$ is a particle destruction operator.

End of Proof

Do Prob. 10, or at least part of it, to prove one or more of the last three relations below to yourself.

Summary of operator functions of coefficients

$$\begin{aligned} N_a(\mathbf{k}) \left(a(\mathbf{k}) | n_{\mathbf{k}} \rangle \right) &= (n_{\mathbf{k}} - 1) \left(a(\mathbf{k}) | n_{\mathbf{k}} \rangle \right) \quad a(\mathbf{k}) \text{ destroys an } a \text{ particle with momentum } \mathbf{k} \\ N_a(\mathbf{k}) \left(a^\dagger(\mathbf{k}) | n_{\mathbf{k}} \rangle \right) &= (n_{\mathbf{k}} + 1) \left(a^\dagger(\mathbf{k}) | n_{\mathbf{k}} \rangle \right) \quad a^\dagger(\mathbf{k}) \text{ creates an } a \text{ particle with momentum } \mathbf{k} \\ N_b(\mathbf{k}) \left(b(\mathbf{k}) | \bar{n}_{\mathbf{k}} \rangle \right) &= (\bar{n}_{\mathbf{k}} - 1) \left(b(\mathbf{k}) | \bar{n}_{\mathbf{k}} \rangle \right) \quad b(\mathbf{k}) \text{ destroys a } b \text{ particle with momentum } \mathbf{k} \\ N_b(\mathbf{k}) \left(b^\dagger(\mathbf{k}) | \bar{n}_{\mathbf{k}} \rangle \right) &= (\bar{n}_{\mathbf{k}} + 1) \left(b^\dagger(\mathbf{k}) | \bar{n}_{\mathbf{k}} \rangle \right) \quad b^\dagger(\mathbf{k}) \text{ creates a } b \text{ particle with momentum } \mathbf{k} . \end{aligned} \tag{3-75}$$

Creation operators $a^\dagger(\mathbf{k})$ and $b^\dagger(\mathbf{k})$ are sometimes called raising operators, because they raise the number of particles in a state. Destruction operators $a(\mathbf{k})$ and $b(\mathbf{k})$ are sometimes called lowering

operators, for what should be obvious reasons. States that have been operated on by a raising operator are sometimes called raised states; those by a lowering operator, lowered states.

3.6.2 Normalization Factors for Raised and Lowered States

Raising and lowering particle number results in a factor in front of the new ket

When a raising operator operates on a ket, the resulting raised ket does not generally have unit norm (is not normalized.) Consider

$$a^\dagger(\mathbf{k})\left|n_\mathbf{k}\right\rangle = A\left|n_\mathbf{k}+1\right\rangle \ , \tag{3-76}$$

where A is some constant (which is a number, not an operator, and could be complex). The original ket $|n_\mathbf{k}\rangle$ and the ket $|n_\mathbf{k}+1\rangle$ (without the constant A) in (3-76) have unit norm. (See (3-64) for one example.) Also, by taking the complex conjugate transpose of (3-76), we see the $a(\mathbf{k})$ acting leftward on the bra has the same raising effect as the $a^\dagger(\mathbf{k})$ acting on the ket,

$$\left(A\left|n_\mathbf{k}+1\right\rangle\right)^\dagger = \left(a^\dagger(\mathbf{k})\left|n_\mathbf{k}\right\rangle\right)^\dagger = \left\langle n_\mathbf{k}\right|a(\mathbf{k}) = \left\langle n_\mathbf{k}+1\right|A^\dagger \ . \tag{3-77}$$

Note that

$$\underbrace{\left\langle n_\mathbf{k}\right|a(\mathbf{k})}_{\left\langle n_\mathbf{k}+1\right|A^\dagger}a^\dagger(\mathbf{k})\left|n_\mathbf{k}\right\rangle = \left\langle n_\mathbf{k}+1\right|A^\dagger A\left|n_\mathbf{k}+1\right\rangle = A^\dagger A\underbrace{\left\langle n_\mathbf{k}+1\right|\left|n_\mathbf{k}+1\right\rangle}_{1} = A^\dagger A \ . \tag{3-78}$$

(3-78) also equals

$$\left\langle n_\mathbf{k}\right|\underbrace{a(\mathbf{k})a^\dagger(\mathbf{k})}_{\text{use commutator}}\left|n_\mathbf{k}\right\rangle = \left\langle n_\mathbf{k}\right|\underbrace{a^\dagger(\mathbf{k})a(\mathbf{k})}_{N_a(\mathbf{k})}+1\left|n_\mathbf{k}\right\rangle = \left\langle n_\mathbf{k}\right|n_\mathbf{k}+1\left|n_\mathbf{k}\right\rangle = n_\mathbf{k}+1 \ . \tag{3-79}$$

Equating the RHS's of (3-78) and (3-79), and for simplicity taking A as real (complex would also work, but be more complicated) yields

$$A = \sqrt{n_\mathbf{k}+1} \ . \tag{3-80}$$

From (3-76), we then have the first line in (3-81) below. Identical logic leads to the third line. Do Prob. 11 if you can't just accept the second and fourth lines without seeing for yourself how they are obtained.

Factors arising from action of creation and destruction operators

$$\boxed{\begin{aligned} a^\dagger(\mathbf{k})\left|n_\mathbf{k}\right\rangle &= \sqrt{n_\mathbf{k}+1}\left|n_\mathbf{k}+1\right\rangle \\ a(\mathbf{k})\left|n_\mathbf{k}\right\rangle &= \sqrt{n_\mathbf{k}}\left|n_\mathbf{k}-1\right\rangle \\ b^\dagger(\mathbf{k})\left|\bar{n}_\mathbf{k}\right\rangle &= \sqrt{\bar{n}_\mathbf{k}+1}\left|\bar{n}_\mathbf{k}+1\right\rangle \\ b(\mathbf{k})\left|\bar{n}_\mathbf{k}\right\rangle &= \sqrt{\bar{n}_\mathbf{k}}\left|\bar{n}_\mathbf{k}-1\right\rangle \end{aligned}} \tag{3-81}$$

2^{nd} quantization responsible for creation/ destruction operators

Note that the above results are ultimately due to 2nd quantization. The non-commutation of fields and their conjugate momenta resulted in the coefficient commutation relations, which was a crucial part in the proof that the coefficients create and destroy states, as well as the derivation of the normalization constants shown above. *Second quantization turned the solution coefficients in RQM, which were merely constants, into creation and destruction operators in QFT.*

3.6.3 Annihilating the Vacuum

The vacuum state has unit norm

Note that the vacuum $|0\rangle$ has unit norm, like any other state, i.e.,

$$\left\langle 0\right|\left|0\right\rangle = 1 \ . \tag{3-82}$$

Note also that 0 (zero) is a number representing nothing, and is different from $|0\rangle$, the vacuum state, which actually is something. From (3-81), the action of a destruction operator on the vacuum results in zero. That is,

A destruction operator annihilates the vacuum, i.e., leaves 0

$$a(\mathbf{k})\left|0\right\rangle = \sqrt{0}\left|-1\right\rangle = 0 \ . \tag{3-83}$$

Don't worry about the funny looking ket in the middle (which is actually meaningless and not something you will ever see in the literature). The root of zero controls the final result.

In QFT lingo, one says "a lowering operator destroys (or annihilates) the vacuum".

3.6.4 Total Particle Number

For future use, we define the total particle number as the number of particles (i.e. *a* types) minus the number of antiparticles (*b* types). For scalars, the total particle number operator is

Total particle number is number of particles minus number of antiparticles

$$N(\phi) = \sum_{\mathbf{k}} \left(N_a(\mathbf{k}) - N_b(\mathbf{k}) \right). \tag{3-84}$$

Note the subtle difference in phraseology in that we commonly use the term "number of particles" as being equal to the number of particles *plus* the number of antiparticles. "Total particle number" on the other hand refers to a negative value for the number of antiparticles.

We will soon see that *b* particles have opposite charge from *a* particles, and thus, in many senses, represent their negatives. So designating them as having negative total particle number seems reasonable.

3.6.5 $\phi(\mathbf{x})$ and $\phi^\dagger(\mathbf{x})$ as Operator Fields

Since the field solutions $\phi(x)$ and $\phi^\dagger(x)$ contain the operator coefficients, they are then also operators, or more properly, operator fields or quantum fields. As noted in Chap. 1 and earlier in the present chapter, this is often shortened in QFT to simply fields. At long last, we have proven our earlier statements about the solutions to the wave equation in QFT being operators (fields).

ϕ and $\phi^\dagger$ are operator fields since they contain coefficient operators

Note that for our field solutions of (3-36), ϕ acts as a total particle number lowering operator, because it destroys particles (via $a(\mathbf{k})$) and creates antiparticles (via $b^\dagger(\mathbf{k})$). The former decreases a positive total particle number, whereas the latter increases the magnitude of a negative total particle number. For $\phi^\dagger$, the situation is reversed: $a^\dagger(\mathbf{k})$ creates particles and $b(\mathbf{k})$ destroys antiparticles, both actions increasing the total particle number.

Thus, the total particle lowering operator field is (see (3-36) for full expression)

ϕ is a total particle number lowering operator field

$$\phi = \underbrace{\phi^+}_{\substack{\text{destroys} \\ \text{particles}}} + \underbrace{\phi^-}_{\substack{\text{creates} \\ \text{anti-particles}}}, \tag{3-85}$$

and the total particle raising operator field is

$\phi^\dagger$ is a total particle number raising operator field

$$\phi^\dagger = \underbrace{\phi^{\dagger +}}_{\substack{\text{destroys} \\ \text{anti-particles}}} + \underbrace{\phi^{\dagger -}}_{\substack{\text{creates} \\ \text{particles}}}. \tag{3-86}$$

When we originally saw the field solutions (3-36), it was suggested, as a mnemonic, that you make a copy of them and stick it over your desk. It would be good now to insert (3-85) and (3-86) into that copy and make them part of it, as we will be using those symbols and what they represent, over and over.

3.6.6 Normal Ordering

When the infinite sum of ½ quanta energy in (3-59) was first found, physicists wanted desperately to make it go away. The amount of energy involved should, via general relativity, curve the universe to such an enormous degree that the light emanating from your finger would be bent so much that it would never reach your eyes. But that isn't what happens in our world, so something isn't correct. In fact, the difference in mass-energy level of the vacuum, between what is predicted by theory and what is observed, is on the order of a factor of 10^{120}, the biggest discrepancy between theory and experiment in the history of science.

One approach to solving ("hiding" may be a better word) this problem is something called normal ordering. Normal ordering, in any term, consists of moving all destruction operators to the right hand side of that term. This has little impact for operators that commute, such as $a(\mathbf{k})$ and $b^\dagger(\mathbf{k})$, for example. That is, changing the term $a(\mathbf{k})b^\dagger(\mathbf{k})$ to $b^\dagger(\mathbf{k})a(\mathbf{k})$ is not an issue, since the factors commute, i.e., $a(\mathbf{k})b^\dagger(\mathbf{k}) = b^\dagger(\mathbf{k})a(\mathbf{k})$. However, the term $a(\mathbf{k})a^\dagger(\mathbf{k})$ is a different story, since $a(\mathbf{k})a^\dagger(\mathbf{k}) \neq a^\dagger(\mathbf{k})a(\mathbf{k})$.

Normal ordering puts all destruction operators in any term on the RHS

In particular, note the effect of normal ordering in our derivation of the number operator form of the Hamiltonian in (3-54). Instead of employing the commutator relations for the $a(\mathbf{k})a^\dagger(\mathbf{k})$ and $b(\mathbf{k})b^\dagger(\mathbf{k})$ terms as done in (3-54), we simply move all the destruction operators to the RHS, so those

terms become $a^\dagger(\mathbf{k})a(\mathbf{k})$ and $b^\dagger(\mathbf{k})b(\mathbf{k})$. Thus, we never end up with the $\frac{1}{2}\omega_{\mathbf{k}}$ terms, the Hamiltonian is finite, just what we would originally have expected it to be, and the vacuum has zero energy, i.e.,.

Normal ordering makes vacuum energy go away

$$H_0^0 = \sum_{\mathbf{k}} \omega_{\mathbf{k}} \left(a^\dagger(\mathbf{k}) a(\mathbf{k}) + b^\dagger(\mathbf{k}) b(\mathbf{k}) \right) \quad (\text{normal ordered} = \text{what is observed}). \qquad (3\text{-}87)$$

The Hamiltonian only has number operators yielding $n_{\mathbf{k}}\omega_{\mathbf{k}}$ energy for $n_{\mathbf{k}}$ particles, each having 3-momentum $\mathbf{k}$.

Although use of normal ordering became quite widespread, it suffers from a pretty fundamental problem. It violates the foundational postulate of non-commutation of certain operators, upon which all of QFT stands. Invoking normal ordering means assuming, in this one area of QFT, that $a(\mathbf{k})$ and $a^\dagger(\mathbf{k})$ (as well as $b(\mathbf{k})$ and $b^\dagger(\mathbf{k})$) commute! But they don't. And the fact that they don't is fundamental to every other part of QFT[6]. In normal ordering, we simply suspend commutation long enough to get a zero energy for the vacuum, then bring it back for the rest of the theory. It is not unreasonable to conclude that use of normal ordering for this purpose is questionable, at best.

Some argue that using normal ordering in this way is internally inconsistent

Caveat: The position expressed in the above paragraph is not widespread, and normal ordering to remove the ½ quanta terms is often invoked with little comment or qualification. I, and a handful of others[7] I am aware of, contend it should simply be jettisoned.

In any case, in spite of common use of normal ordering to "clean up" the Hamiltonian, the huge vacuum energy issue has not gone away in most physicists eyes, and it remains a widely discussed, unsolved problem as of 2012, the year of this book (though I offer a possible solution in the article cited in the footnote on page 50.

Paradoxically both 1) *normal ordering is widely used to eliminate vacuum energy, and* 2) *vacuum energy is generally accepted as a fact*

It may seem to you the reader that the entire issue is fraught with ambiguity, and that is probably a reasonable assessment. In spite of that, virtually everyone considers vacuum energy to be a reality.

3.6.7 The Observable Hamiltonian

One can distinguish observables, which in quantum theories are represented by operators, from the theoretically obtained expressions for the corresponding operators. That is, the $\frac{1}{2}\omega_{\mathbf{k}}$ terms in H_0^0 are not observed, so for reasons of practicality, we can consider the normal ordered Hamiltonian of of (3-87) to be what we will call the observable Hamiltonian.

The normal ordered Hamiltonian is the observable Hamiltonian.

The relation (3-87), even though derived via the *ad hoc* and mathematically questionable normal ordering process, results in what is actually observed. This is why normal ordering persists.

Since we will not do a lot with the vacuum, we will find it more convenient and streamlined to use (3-87) for the Hamiltonian, except when we are specifically interested in vacuum energy. Just don't forget, as we do that, what the complete Hamiltonian, as derived via our theory, looks like.

3.7 Probability, Four Currents, and Charge Density

Probability in QFT is found in essentially the same way as we did for NRQM (see Box 3-1) and RQM (see Sect. 3.1.4, pg. 44.) That is, we use the governing wave equation and manipulate it to obtain a relationship like the continuity equation (3-19) (or (3-23) in 4D notation). The integral over all space of the quantity ρ in that relationship is conserved, and since total probability (of finding one or more particles) is also conserved, ρ has a good chance of being probability density. Experiment can confirm, or deny, that.

Probability in QFT found in similar way as in NRQM and RQM, i.e., from wave equation

3.7.1 Four Currents Operators and Probability Density in QFT

In the present case, the solutions to the governing equation are operator fields, not states, so we would expect the resulting density ρ to be an operator density, rather than a numeric density. Our expectation will turn out to be true, as we see below.

---

[6] As an example of one such part, see Sect. 3.6.1, where commutation relations result in $a(\mathbf{k})$ and $b(\mathbf{k})$ being destruction operators, and $a^\dagger(\mathbf{k})$ and $b^\dagger(\mathbf{k})$ being creation operators.

[7] P. Teller, *An Interpretive Introduction to Quantum Field Theory*, Princeton University Press (1995). Teller submits that a complete and true theory would not have such an artificial, and arbitrarily imposed, feature as normal ordering. On page 130, with reference to normal ordering he states, "If, as appears to be the case, at this point one must use mathematically illegitimate tricks, concern is an appropriate response."

Since our governing equation is the Klein-Gordon equation, and that is the same as in RQM, similar steps (3-16) to (3-20) can be followed. The result is the same 4-current relations as (3-22) and (3-23), except that ϕ and $\phi^\dagger$ are now operator fields (or simply, in QFT lingo "fields"),

Same wave equation as RQM → same form for probability 4-current

$$\boxed{j^\mu{}_{,\mu} = 0 \quad \text{with} \quad j^\mu = i\left(\phi^{,\mu}\phi^\dagger - \phi^{\dagger,\mu}\phi\right) \qquad j^\mu \text{ is an operator}}, \tag{3-88}$$

so

But now 4-current is an operator

$$\rho = j^0 = i\left(\frac{\partial\phi}{\partial t}\phi^\dagger - \frac{\partial\phi^\dagger}{\partial t}\phi\right). \tag{3-89}$$

Since (3-89) is an operator, we need its expectation value to find measurable probability density,

So, need to find expectation value of probability operator

$$\bar{\rho} = \langle\phi_1,\phi_2,\phi_3,....|\,\rho\,|\phi_1,\phi_2,\phi_3,....\rangle. \tag{3-90}$$

To evaluate (3-90), we need first to substitute our free field solutions (3-36) into (3-89). Do Prob. 12 to prove to yourself that if we restrict ourselves to particles in **k** eigenstates (which is typically the case in QFT), then this results in an effective density operator

Probability operator expressed in terms of number operators

$$\rho = \frac{1}{V}\sum_{\mathbf{k}}\left(a^\dagger(\mathbf{k})a(\mathbf{k}) - b^\dagger(\mathbf{k})b(\mathbf{k})\right) = \frac{1}{V}\sum_{\mathbf{k}}\left(N_a(\mathbf{k}) - N_b(\mathbf{k})\right). \tag{3-91}$$

3.7.2 Single Particle State

Let's now find the expectation value of ρ for a single *a* type particle state $|\phi_{\mathbf{k}'}\rangle$. We find all the number operators except the one for an *a* type particle with momentum **k**′ yield zero, so

Probability expectation value for single type a particle = 1, no surprise

$$\bar{\rho} = \langle\phi_{\mathbf{k}'}|\,\rho\,|\phi_{\mathbf{k}'}\rangle = \langle\phi_{\mathbf{k}'}|\frac{1}{V}\sum_{\mathbf{k}}\left(N_{\mathbf{a}}(\mathbf{k}) - N_{\mathbf{b}}(\mathbf{k})\right)|\phi_{\mathbf{k}'}\rangle = \langle\phi_{\mathbf{k}'}|\frac{1}{V}|\phi_{\mathbf{k}'}\rangle = \frac{1}{V}. \tag{3-92}$$

For a plane wave, this is exactly our probability density, a flat distribution over the volume, whose integral over the volume equals one. So far, ρ looks like it could well be a probability distribution.

3.7.3 Multiparticle State

But now let's look at a multiparticle state.

Probability expectation value for four type a particles = 4 > 1 !

$$\begin{aligned}\bar{\rho} &= \langle 3\phi_{\mathbf{k}_1},\phi_{\mathbf{k}_2}|\,\rho\,|3\phi_{\mathbf{k}_1},\phi_{\mathbf{k}_2}\rangle = \langle 3\phi_{\mathbf{k}_1},\phi_{\mathbf{k}_2}|\frac{1}{V}\sum_{\mathbf{k}}\left(N_{\mathbf{a}}(\mathbf{k}) - N_{\mathbf{b}}(\mathbf{k})\right)|3\phi_{\mathbf{k}_1},\phi_{\mathbf{k}_2}\rangle \\ &= \langle 3\phi_{\mathbf{k}_1},\phi_{\mathbf{k}_2}|\frac{4}{V}|3\phi_{\mathbf{k}_1},\phi_{\mathbf{k}_2}\rangle = \frac{4}{V}.\end{aligned} \tag{3-93}$$

When (3-93) is integrated over *V*, we get 4, the number of particles in the state! Since total probability is never greater than 1, our interpretation of ρ as a probability density seems to be in trouble for multiparticle states.

So, QFT deals with number of particles instead of probability

Partially for this reason, QFT rarely deals with probability densities for states. It concerns itself, instead, with *numbers* of particles (and antiparticles) in a state. Thus, the number operators play a major role. As we will see, this works well, and allows us to solve the kinds of problems in QFT we need to solve.

3.7.4 Antiparticles (Type b Particles)

Now consider the expectation value of ρ on a *b* type single particle state.

Probability expectation value for single type b particle = – 1 < 0 !!

$$\bar{\rho} = \langle\bar{\phi}_{\mathbf{k}'}|\,\rho\,|\bar{\phi}_{\mathbf{k}'}\rangle = \langle\bar{\phi}_{\mathbf{k}'}|\frac{1}{V}\sum_{\mathbf{k}}\left(N_{\mathbf{a}}(\mathbf{k}) - N_{\mathbf{b}}(\mathbf{k})\right)|\bar{\phi}_{\mathbf{k}'}\rangle = \langle\bar{\phi}_{\mathbf{k}'}|\frac{-1}{V}|\bar{\phi}_{\mathbf{k}'}\rangle = \frac{-1}{V}. \tag{3-94}$$

So total probability of a *b* type particle would be negative! And for 3 such particles, it would be -3.

This was another tip to early researchers that the density they were dealing with was more readily related to charge density, and the *b* particles were antiparticles, with opposite charge (and charge density) from particles.

Thus, we have started to see that the concept of probability density, and the mathematics associated with it, seem to lose some applicability in QFT. Particle number, however, takes on significance, as now antiparticles would simply be designated as having negative total particle numbers.

Led to conclusion that $\rho \propto$ charge probability density, and its spatial integral is total charge

3.7.5 Charge Density Not Probability Density

If we multiply our four current operator (3-88) by the charge of a scalar particle q it behaves like a charge density operator, which we will designate by s^{μ}.

Multiply 4-*current operator by particle charge to get charge* 4-*current operator*

$$s^{\mu}{}_{,\mu} = 0 \quad \text{with} \quad s^{\mu} = qj^{\mu} = iq\left(\phi^{,\mu}\phi^{\dagger} - \phi^{\dagger,\mu}\phi\right) \ , \tag{3-95}$$

so

$$\rho_{charge} = qj^{0} = iq\left(\frac{\partial\phi}{\partial t}\phi^{\dagger} - \frac{\partial\phi^{\dagger}}{\partial t}\phi\right). \tag{3-96}$$

This makes sense, as charge would be distributed in parallel fashion to probability density, i.e., denser charge where the particle is more concentrated. Further, total charge using (3-93) multiplied by q would yield $4q$, the charge on the state. Similarly, the total charge on the state in (3-94) would be $-q$.

Thus, re-interpreting the operator ρ, as charge density, and the 3D part of the four current as charge current density is consistent. In actuality, it is demanded in order for our theory to agree with experiment. That empirical reality also forces us to accept *b* type particles as antiparticles.

Take care that in the future, we may use the symbol ρ as simply charge density, without a subscript. Since we will rarely, if ever, deal again with probability density, hopefully, there will be little confusion.

3.7.6 Caution in Evaluating Expectation Values of Density Operators

Some care must be taken in the evaluation of expectation values similar to that of (3-90). The bracket, expressed in the position basis, is an integration over space. But for operators with a spatial dependence such as ρ often has (and which is typical of charge, mass or any type of density), the spatial dependence in the operator is <u>not</u> included in the integration. That is, writing out the expectation value as an integral, we integrate over the $\mathbf{x}'$ of the state, but not the $\mathbf{x}$ of the operator.

Integration implied in expectation value is over ket $\mathbf{x}'$*, not over operator* $\mathbf{x}$

$$\left\langle \rho(\mathbf{x},t)\right\rangle = \left\langle \phi_{\mathbf{k}}(\mathbf{x}',t)\,|\,\rho(\mathbf{x},t)\,|\phi_{\mathbf{k}}(\mathbf{x}',t)\right\rangle = \int \phi^{\dagger}_{\mathbf{k},state}(\mathbf{x}',t)\rho(\mathbf{x},t)\phi_{\mathbf{k},state}(\mathbf{x}',t)\,d^3x' \ . \tag{3-97}$$

This was not evident in (3-92) and similar relations above, because there (for plane waves, specifically) the operator ρ was not a function of space.

The point in (3-97) generalizes to other types of operator functions that would be sandwiched inside a bra and a ket. We will run into these in the future.

3.7.7 The ϕ and $\phi^{\dagger}$ Normalization Constants Again

We just assumed, in all of our discussion so far, that the normalization constants in our solutions, $1/\sqrt{2\omega_{\mathbf{k}}V}$, that we derived in RQM, are also valid in QFT. Since our field solutions ϕ and $\phi^{\dagger}$ in QFT had the same form as the state solutions in RQM, and our 4 current $j^{\mu} = (\rho, \mathbf{j})$ in each case had the same form as well, this seems like a reasonable assumption. The assumption can be considered justified by our results in the above few sections. For example, (3-91) worked out as a correct form for density (probability or charge) only because of the form chosen for our constants. The square root of $2\omega_{\mathbf{k}}$ dropped out in getting (3-91) because of the two terms, each with two field factors multiplied, and the time derivatives in (3-89).

ϕ and $\phi^{\dagger}$ normalization constants from RQM work for QFT

We can therefore consider the results of the sections above as justification for the choice of normalization constants in the field solutions to the Klein-Gordon equation. All of so many other results, yet to be seen in our studies, will be further justification.

3.8 More on Observables

QFT, like the quantum theories studied before it, is interested in observable quantities, such as energy, 3-momentum, charge, and spin, which are represented in each of those theories by

operators. The eigenvalues of those operators are what we measure. Expectation values of those operators are the averages of what we measure over many trials.

Finding operators other than H in terms of number operators

Regarding energy, we have already discussed the observable operator corresponding to it. See (3-87). Regarding spin, scalar particles have none, so we will put off discussion of particles that do have it to later chapters.

3.8.1 Charge Operator

Regarding charge, we need merely to integrate our charge density operator qj^0 of (3-96) and (3-91) over the entire volume, to get the charge operator

$$Q = \int s^0 d^3x = q\int j^0 d^3x = q\sum_{\mathbf{k}} \left(N_a(\mathbf{k}) - N_b(\mathbf{k}) \right). \tag{3-98}$$

Total charge operator yields total charge of a (multiparticle) state

A typical multiparticle state is in a charge eigenstate with an eigenvalue of charge equal to the sum of the charges of all particles in the state. Hence, the eigenvalue equals the charge expectation value, since we will measure the same charge with each measurement. For a sample state,

$$\bar{Q} = \left\langle 7\phi_{\mathbf{k}_1}, \phi_{\mathbf{k}_2}, 5\bar{\phi}_{\mathbf{k}_3} \left| \underbrace{q\sum_{\mathbf{k}} \left(N_a(\mathbf{k}) - N_b(\mathbf{k}) \right)}_{Q} \right| 7\phi_{\mathbf{k}_1}, \phi_{\mathbf{k}_2}, 5\bar{\phi}_{\mathbf{k}_3} \right\rangle = 7q + q - 5q = +3q\,. \tag{3-99}$$

Vacuum has zero charge

Note, we derived (3-98) using (3-91). If you did Prob. 12, you saw that in deriving (3-91) we summed terms in ½ and – ½ that cancelled to net zero. In other words, (3-98) actually has a $+q/2$ and a $-q/2$ term for each **k**. Thus, Q acting on the vacuum would sum up an infinite number of half charge quanta for both particles (positive charge) and anti-particles (negative charge), leaving zero total vacuum charge. (Thankfully, it does. If it didn't, our theory would be bound for the trash heap.)

3.8.2 Three Momentum Operator

The three momentum operator can be found using the relationship for physical momentum density at the bottom of Box 2-2 in Chap. 2, pg. 23, and integrating over the volume. (Also shown in the 9th block under the title in the RH column of Wholeness Chart 2-2, pg. 21.) That is,

$$p^i = \int \mathscr{P}^i d^3x = \int \pi_r \frac{\partial \phi^r}{\partial x^i} d^3x = \int \left(\frac{\partial \mathscr{L}}{\partial \dot{\phi}} \frac{\partial \phi}{\partial x^i} + \frac{\partial \mathscr{L}}{\partial \dot{\phi}^\dagger} \frac{\partial \phi^\dagger}{\partial x^i} \right) d^3x \ . \tag{3-100}$$

Substituting the Klein-Gordon solutions (3-36) and their conjugate momenta into(3-100), one obtains the 3-momentum operator (do Prob. 13 to prove it)

Total 3-momentum operator yields total 3-momentum of a (multiparticle) state

$$\mathbf{P} = \sum_{\mathbf{k}} \mathbf{k} \left(N_a(\mathbf{k}) + N_b(\mathbf{k}) \right), \tag{3-101}$$

which is pretty much what we would have expected. **P** operating on a multiparticle ket, with all particles in **k** eigenstates, would yield an eigenvalue equal to the number of particles and antiparticles with 3-momentum **k** multiplied by **k**. If this is less than obvious to you, do Prob. 14.

Vacuum has zero 3-momentum

It is interesting that, similar to what happened to charge, we have ½ quanta in the vacuum with 3-momentum, but the total for the vacuum sums to zero. That is, in deriving (3-101), we get terms in the summation of ½ **k** + ½ **k** = **k** (one ½ quanta for each particle and one for each antiparticle), similar to what we had for energy. But unlike energy, this is a vector summation, and for every 3-momentum **k** in the sum, there is a 3-momentum – **k**, as well. The net is nil 3-momentum for the vacuum, which again, is a welcome result.

So far in our theory, only energy has proved problematic in having a non-zero vacuum expectation value (VEV.)

3.8.3 The Four Momentum Operator

As discussed in the Appendix of Chap. 2, and elsewhere, the four momentum has energy in the 0th component (E/c in non-natural units) and 3-momentum for the other three components. Given (3-87) for H_0^0, and (3-101) for the free scalar field p^i, the four momentum operator is

$$\underbrace{P^\mu = K^\mu}_{\text{operators here}} = \underbrace{\begin{pmatrix} H \\ \mathbf{P} \end{pmatrix}}_{\text{for free scalars}} = \sum_{\mathbf{k}} \underbrace{\begin{pmatrix} \omega_{\mathbf{k}} \\ \mathbf{k} \end{pmatrix}}_{\text{usually what we mean by symbol } k^\mu} \left(N_a(\mathbf{k}) + N_b(\mathbf{k}) \right), \tag{3-102}$$

4-momentum operator includes energy and 3-momentum

where we note that k^μ usually refers to the numeric (not operator) 4 vector $(\omega_{\mathbf{k}},\mathbf{k})$.

3.9 Real Fields

So far, we have only dealt with complex fields. It is possible to have real fields, and in fact, we will see they play a key role in the theory. They turn out to be associated with neutral particles, which are, in fact, their own antiparticles.

To see this, look at a special case for our general field equation solutions (3-36) where $\phi = \phi^\dagger$, i.e., ϕ is real. In order for this to be true, we must have $a(\mathbf{k}) = b(\mathbf{k})$, and of course, $a^\dagger(\mathbf{k}) = b^\dagger(\mathbf{k})$. Thus, for this case,

$$\phi = \phi^\dagger = \sum_{\mathbf{k}} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}} a(\mathbf{k}) e^{-ikx} + \sum_{\mathbf{k}} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}} a^\dagger(\mathbf{k}) e^{ikx} \quad (\text{for } \phi \text{ real}), \tag{3-103}$$

which makes sense, since adding a complex number and its complex conjugate yields a real number.

In the charge operator (3-98), we would then have $N_a(\mathbf{k}) = N_b(\mathbf{k})$ (i.e., $a^\dagger(\mathbf{k})a(\mathbf{k}) = b^\dagger(\mathbf{k})b(\mathbf{k})$), so charge would be zero for any such particle(s) state. Each *b* type particle operator (creation, destruction, charge, energy, etc.) will be the same operator as that for *a* particles. There is only one type of particle (for a real field), so that particle must be its own antiparticle.

Real (not complex) fields create and destroy neutral charge particles

Conclusion: Real fields are associated with charge-neutral particles, which are their own anti-particles.

Note on nomenclature: The term "real field" refers to the (operator) field solution to the field equation (Klein-Gordon for scalars) that are not complex. The term "real particle" refers to a particle that is not virtual, but manifest and detectable.

Terminology difference between real fields and real particles

3.10 Characteristics of Klein-Gordon States

3.10.1 Bosons vs Fermions

Although at this point in your career, you should be familiar with bosons and fermions, and their behavior, Wholeness Chart 3-1 can serve as a refresher course. It should need no further comment.

Wholeness Chart 3-1. Bosons vs Fermions

	Bosons	**Fermions**
What role	typically forces	typically matter
Some examples	elementary: photons, Higgs composite: mesons	elementary: electrons, neutrinos, quarks composite: baryons (e.g., proton, neutron)
Behavior	can occupy same state	can't occupy the same state
Spin	integer spin Scalars: spin 0 Vectors (e.g. photons): spin 1 Graviton: spin 2	half integer spin Spinors: spin ½ Gravitinos: spin 3/2

3.10.2 Klein-Gordon States are Bosons

The same scalar creation operator that operates repeatedly on a state results in a raised state containing a number of the same particle with the same **k** (and thus the same energy and identical in all regards.) For example, using the creation operator of the first line of (3-81) acting first on the vacuum, then repeatedly on the newly created states, we have

$$a(\mathbf{k})^\dagger|0\rangle = |\phi_\mathbf{k}\rangle \;\rightarrow\; a(\mathbf{k})^\dagger|\phi_\mathbf{k}\rangle = \sqrt{2}\,|2\phi_\mathbf{k}\rangle \;\rightarrow\; a(\mathbf{k})^\dagger|2\phi_\mathbf{k}\rangle = \sqrt{3}\,|3\phi_\mathbf{k}\rangle \rightarrow ..... \qquad (3\text{-}104)$$

Scalar kets can have > 1 particle in same state, so scalars are bosons, not fermions

We are not concerned, for this discussion, with the square root numeric coefficients, but with the fact that we can have multiparticle states with more than one individual particle in the same individual state.

This means Klein-Gordon states must be bosons. We sort of knew this because we were told that they have zero spin. But here we prove it.

As we will see in the next chapter, spinors do not have the characteristic displayed by (3-104).

3.10.3 Commutators with Scalars, Not Anti-Commutators

Let's see what happens if anti-commutators were used instead of commutators with the Klein-Gordon field equation solutions. That is, in the derivation of our number operator form of the Hamiltonian, in equation (3-54), try the anti-commutators

$$\begin{aligned}\left[a(\mathbf{k}),a^\dagger(\mathbf{k}')\right]_+ &= a(\mathbf{k})a^\dagger(\mathbf{k}') + a^\dagger(\mathbf{k}')a(\mathbf{k}) = \delta_{\mathbf{k}\mathbf{k}'} \\ \left[b(\mathbf{k}),b^\dagger(\mathbf{k}')\right]_+ &= b(\mathbf{k})b^\dagger(\mathbf{k}') + b^\dagger(\mathbf{k}')b(\mathbf{k}) = \delta_{\mathbf{k}\mathbf{k}'}.\end{aligned} \qquad (3\text{-}105)$$

A minus sign is then introduced, such that instead of (3-55), we would get

$$H_0^0 = \sum_\mathbf{k} \omega_\mathbf{k}\left(\tfrac{1}{2} + \tfrac{1}{2}\right) , \qquad (3\text{-}106)$$

Using anti-commutator instead of commutator → theory with faulty energy prediction

or for an observable Hamiltonian (ignoring vacuum energy)

$$H_0^0 = 0 \quad \text{and thus,} \quad H_0^0|\phi_{\mathbf{k}_1}\rangle = 0\,. \qquad (3\text{-}107)$$

Every real state would have zero energy, which is certainly not physically true. Therefore, we can only use commutators with spin 0 boson fields, and not anti-commutators.

Klein-Gordon scalars:
1) *must be bosons, and*
2) *can only use commutation*

We will find in the next chapter, that fermions in QFT are governed by anti-commutators in parallel fashion to the way in which bosons (scalars, at least, to this point in our studies) are governed by commutators. Just as anti-commutators can't work for bosons (scalars), as we saw above, we will also see later that commutators can't work for fermions.

3.11 Odds and Ends

3.11.1 Usefulness of 3-Momentum Discrete Eigenstates

Discrete form of solutions (*not wave packets*) *suffice for extremely accurate real particle predictions*

As you will see in time, QFT can find real world experimental values, for things like scattering cross sections and decay half lives, using only discrete **k** eigenstate forms for real states. These eigenstates, unlike wave packets, typically extend indefinitely in the **k** direction. But for experimental predictions, the particle states, which are actually wave packets, can be approximated to extremely high precision by such discrete **k** eigenstates.

One exception is the propagator, the mathematical representation of virtual particles, which is best derived, as shown in Sect. 3.13, and most useful practically, via incorporation of the continuous (integral) form of the field equation solutions. This is, at least in part, because virtual particles are not constrained by boundary conditions to discrete **k** values and in certain cases must be integrated over all possible, continuous values of **k**.

Except, we will need continuous solutions to find Feynman propagator (*for virtual particles*)

3.11.2 Nevertheless, What about Non-eigen States?

In NRQM (and RQM) we commonly dealt with general states, i.e., non-eigen states, which were superpositions of two or more eigenstates. Granted, as noted above, that we can solve almost all QFT problems using **k** eigenstates, we might still ask "how does QFT compare in this regard to what we learned in NRQM?" It is a good question, troubling many students, no doubt, and not treated in any other text I am aware of.

A closely related question is "what is created or destroyed by the general solutions $\phi(x)$ (or $\phi^\dagger(x)$), which for discrete eigenstates, is a summation of terms, each containing a single particle eigenstate creation/destruction operator?" Does operation of $\phi^\dagger(x)$ on the vacuum, for instance, create an infinite number of single particles, or a single particle comprising an infinite number of

momentum eigenstates? If the latter, what amplitudes (whose absolute values squared are probabilities) are assigned to each such eigenstate?

Creating a General Single Particle State (Discrete Solution Form)

To create a general single particle state, , we would need a creation operator of form

Creation operator for single particle general (non-eigen) state, discrete case.

$$C = \sum_{\mathbf{k}} A_{\mathbf{k}} \mathbf{a}_{\mathbf{k}}^{\dagger} , \tag{3-108}$$

so that operation of C on the vacuum results in a sum of eigenstates,

$$C|0\rangle = \sum_{\mathbf{k}} A_{\mathbf{k}} \mathbf{a}_{\mathbf{k}}^{\dagger}|0\rangle = A_1|\phi_1\rangle + A_2|\phi_2\rangle + A_3|\phi_3\rangle + ... = |\phi\rangle . \tag{3-109}$$

In (3-108) and (3-109) $A_{\mathbf{k}}$ is a numerical coefficient, the square of the absolute value of which (for proper normalization of the ket eigenstates) equals the probability of finding the **k** eigenstate.

If only one term in C is used, then only one eigenstate with $|A_{\mathbf{k}}| = 1$ is created. If a more general state, comprising a sum of eigenstates, is created, then we are free to select the $A_{\mathbf{k}}$ as we please in order to create the particular general state we like, provided (for conservation of probability and correct normalization so total probability is unity)

$$\sum_{\mathbf{k}} |A_{\mathbf{k}}|^2 = 1 . \tag{3-110}$$

Destroying a General Single Particle State (Discrete)

Destruction operator for single particle general (non-eigen) state, discrete case.

Note that the general single particle destruction operator

$$D = \sum_{\mathbf{k}} \mathbf{a}_{\mathbf{k}} \tag{3-111}$$

acting on any single particle general state will lower that state to the vacuum. That is,

$$\begin{aligned}
&\left(\sum_{\mathbf{k}} \mathbf{a}_{\mathbf{k}}\right)|\phi\rangle = \left(\sum_{\mathbf{k}} \mathbf{a}_{\mathbf{k}}\right)\left(A_1|\phi_1\rangle + A_2|\phi_2\rangle + A_3|\phi_3\rangle + ...\right) \\
&= \left(\left(\sum_{\mathbf{k}} \mathbf{a}_{\mathbf{k}}\right) A_1|\phi_1\rangle + \left(\sum_{\mathbf{k}} \mathbf{a}_{\mathbf{k}}\right) A_2|\phi_2\rangle + \left(\sum_{\mathbf{k}} \mathbf{a}_{\mathbf{k}}\right) A_3|\phi_3\rangle + \dots\right) \\
&= A_1 \underbrace{\mathbf{a}_1|\phi_1\rangle}_{|0\rangle} + A_1 \underbrace{\mathbf{a}_2|\phi_1\rangle}_{0} + A_1 \underbrace{\mathbf{a}_3|\phi_1\rangle}_{0} + ... + A_2 \underbrace{\mathbf{a}_1|\phi_2\rangle}_{0} + A_2 \underbrace{\mathbf{a}_2|\phi_2\rangle}_{|0\rangle} + 0 + ... = \underbrace{(A_1 + A_2 + ...)}_{\text{can normalize = 1}}|0\rangle .
\end{aligned} \tag{3-112}$$

Creating and Destroying Multi-particle State (Discrete)

For multiparticle general states, repeatedly apply C and D operators

Applying operators similar in form to (3-108) (with typically different values for $A_{\mathbf{k}}$ in each operator) twice in succession creates a two particle state where each particle is a single particle general state (i.e., each is a summation of momentum eigenstates.) Any number of such operators may be applied to create a state of any number of particles, each in a general (not eigen) state.

Applying (3-111) repeatedly will destroy one general state single particle upon each application.

What $\phi(x)$ and $\phi^{\dagger}(x)$ Create When Acting on the Vacuum

$\phi(x)$ acting on the vacuum will create a single general antiparticle state comprising a superposition of an infinite number of eigenstates, each with a constant coefficient in front of it, i.e.,

$$\phi(x)|0\rangle = \sum_{\mathbf{k}} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}} e^{-ikx} \underbrace{a(\mathbf{k})|0\rangle}_{0} + \sum_{\mathbf{k}} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}} e^{ikx} \underbrace{b^{\dagger}(\mathbf{k})|0\rangle}_{|\bar{\phi}_{\mathbf{k}}\rangle} = \sum_{\mathbf{k}} \underbrace{\frac{1}{\sqrt{2V\omega_{\mathbf{k}}}}}_{\text{constant}} \underbrace{e^{ikx}}_{\text{phase factor}} |\bar{\phi}_{\mathbf{k}}\rangle . \tag{3-113}$$

Similarly, $\phi^{\dagger}(x)$ acting on the vacuum will create a single particle general state comprising a superposition of particle eigenstates. For discrete solutions, these operations have little use in QFT.

Creating and Destroying Continuous Solution Forms of States

In analogous fashion, the continuous form of $\phi(x)$ (or $\phi^{\dagger}(x)$) acting on the vacuum yields a single antiparticle (particle) wave packet, i.e., an integral over all **k**, rather than a sum. You don't need to know more now, but for the future, details are in this book's web site (address on pg. xvi, opposite pg. 1) at the link titled Non Eigen States, Wave Packets, and the Hamiltonian in QFT. A summary

of that can be found in Appendix B of Chap. 10 herein. We will work more with these operations when we derive the Feynman propagator in Sect. 3.13.

3.11.3 c-numbers vs q-Numbers

q-numbers are quantum numbers (operators or their eigenvalues);

c-numbers are classical numbers

The terms <u>*c*-number</u> and <u>*q*-number</u> were introduced by Paul Dirac to distinguish between *classical numbers* (real or complex), which commute, and *quantum operators*, which do not always commute. The term *q*-number can equally apply to the *eigenvalue* of a given quantum operator.

Thus, the 3-momentum of a classical particle is a *c*-number. The 3-momentum of a quantum state, in a 3-momentum eigenstate, is a *q*-number.

Eigenstates are often labeled by their *q*-numbers (eigenvalues). For example, the *n*, *l*, and *m* numbers for electron levels in the hydrogen atom are quantum, or *q*-, numbers. *n* represents the energy level number (which is simpler than specifying the energy itself); *l*, the angular momentum magnitude; and *m*, the *z* component of angular momentum. By specifying *n*, *l*, and *m*, one specifies the eigenstate of the electron in the atom.

3.11.4 Fock Space and Hilbert Space

As you should (hopefully) remember, a quantum state in NRQM is an abstract vector in an abstract vector space, analogous to a physical vector in 3D physical space. The same thing is true in RQM and QFT. This is summarized in Wholeness Chart 3-2.

In all quantum theories, basis vectors (which are typically eigenstates) are abstractions of the unit basis vectors along the 3D axes. A general state is a vector sum of certain amounts of each basis vector state. Operators in each kind of space act on the states in that space.

In NRQM and RQM, the states are single particle states and the abstract space they inhabit is called <u>Hilbert space</u>, which has a different single particle eigenstate as the basis vector of each "axis". The dimension of the Hilbert space for a given system is simply the number of eigenstates in that system. This can, for many systems, be infinite.

Fock space is Hilbert space generalized to multiparticle states

In QFT, states are multiparticle, so the basis eigenstate of each "axis" is a multiparticle state. One "axis" basis vector might be an electron and a photon, each with particular 3-momentum. Another might be 2 photons and a positron with particular 3-momenta. Yet another might be an electron and photon like the first, except that at least one of them has different 3-momentum from the first. The multiparticle abstract state space of QFT is called <u>Fock space</u>, which is simply an extension of Hilbert space to multiparticle states.

Wholeness Chart 3-2. Physical, Hilbert, and Fock Spaces

	<u>3D Physical Space</u>	**<u>Hilbert Space</u>**	**<u>Fock Space</u>**
Character of a vector	Position vector in 3D	State vector $\|\Psi\rangle$ in NRQM, RQM Single particle	State vector $\|\phi_1, \phi_2, \ldots\rangle$ in QFT Multi particle
Orthonormal basis vectors along "axes"	$\hat{\mathbf{i}}_1, \hat{\mathbf{i}}_2, \hat{\mathbf{i}}_3$	Normalized eigenvectors $\|\Psi_1\rangle, \|\Psi_2\rangle, \|\Psi_3\rangle, \|\Psi_4\rangle, \ldots.$	Normalized eigenvectors $\|0\rangle, \|\phi_1\rangle, \|\phi_2\rangle, .. \mid \phi_1, \phi_2\rangle, ..\|\phi_1, \phi_2, \phi_3\rangle, \ldots$
Inner product	$\hat{\mathbf{i}}_i \bullet \hat{\mathbf{i}}_j = \delta_{ij}$	$\langle\Psi_r\|\Psi_s\rangle = \delta_{rs}$	$\langle\phi_1\| \phi_1, \phi_2\rangle = 0$; $\langle\phi_1, \phi_2\| \phi_1, \phi_2\rangle = 1$; etc.
General state vector	$\mathbf{r} = x^1\hat{\mathbf{i}}_1 + x^2\hat{\mathbf{i}}_2 + x^3\hat{\mathbf{i}}_3$	$\|\Psi\rangle = C_1\|\Psi_1\rangle + C_2\|\Psi_2\rangle + C_3\|\Psi_3\rangle + \ldots.$	$\|\Phi\rangle = C_1\|\phi_1\rangle + C_2\|\phi_2\rangle + \ldots + C_{12}\|\phi_1, \phi_2\rangle + C_{13}\|\phi_1, \phi_3\rangle + \ldots + C_{123}\|\phi_1, \phi_2, \phi_3\rangle + \ldots$
State vector & its components	point in 3D space = amount along each basis vector	"point" in Hilbert Space = amount of each single particle basis vector	"point" in Fock Space = amount of each multi particle basis vector
Operators	Matrices operate on vectors	Hamiltonian H, 3-momentum $\mathbf{P}$, etc operate on states	Hamiltonian H, $\mathbf{P}$, creation $a^\dagger(\mathbf{k})$, $b^\dagger(\mathbf{k})$, destruction $a(\mathbf{k}), b(\mathbf{k})$, charge Q operate on states

3.11.5 a(**k**) *Destroys Any State without Single a Type Particle in* **k** *Eigenstate*

Destruction operators of given kind and **k** *destroy any state not having like kind of particle in a* **k** *eigenstate*

Keep in mind (as we actually already did in the last row of (3-112)) that, for example,

$$\mathbf{a}_2 \left| \phi_1, 4\phi_3, 7\bar{\phi}_2 \right\rangle = 0 , \tag{3-114}$$

which is zero, not the vacuum state. In general, a particular type particle destruction operator of given **k** acting on a state that has no particles of that type of the same 3-momentum **k** results in zero.

3.12 Harmonic Oscillators and QFT

One sometimes hears that particles in QFT can be considered to be harmonic oscillators. The reason for this can be seen with the aid of Wholeness Chart 3-3, which summarizes the states of the NRQM harmonic oscillator (relativistic form is similar, but more complicated, so we won't bother with it) and particles in QFT.

QFT particle states have similarities to harmonic oscillator energy states

One sees immediately that the energy levels of the two look very similar. Each level is $\hbar\omega$ above the one below it. (We keep the symbol $\hbar$ for this discussion, since it makes the NRQM summary look more familiar.) And strikingly, each also has a lowest level of energy, when $n = 0$ ($n_{\mathbf{k}} = 0$ in QFT, to be precise), of ½ quantum (½$\hbar\omega$ or ½$\hbar\omega_{\mathbf{k}}$.) More striking still, each has raising and lowering operators that raise and lower energy levels by $\hbar\omega$ (or $\hbar\omega_{\mathbf{k}}$ for each extra particle in QFT.)

These similarities led people to think in terms of QFT particles as harmonic oscillators. The vacuum was the lowest excitation of the quantum field. (Really, one should say the "state", not the "field", but people commonly express it this way. Confusing? Yes.) Each state above (in QFT, each additional particle) was simply a more excited state of the lowest state (in QFT, the vacuum state.) Operators acting on states raise or lower the number of particles, and thus the energy level, and so excite, or de-excite, the vacuum. Particles are just excitations of an underlying vacuum field.

Wholeness Chart 3-3. Quantum Harmonic Oscillator Compared to QFT Free States

	NRQM Harmonic Oscillator	**QFT Free States**
Energy Levels	$(n + ½)\,\hbar\omega$	$(n_{\mathbf{k}} + ½)\,\hbar\omega_{\mathbf{k}}$
Interpretation of n and $n_{\mathbf{k}}$	single particle energy level 0,1,2,…	number of particles at $\hbar\omega_{\mathbf{k}}$ energy
Interpretation of ω and $\omega_{\mathbf{k}}$	natural frequency of classical oscillator	frequency of particle of energy $\hbar\omega_{\mathbf{k}}$
Lowest energy level	½ $\hbar\omega$	½ $\hbar\omega_{\mathbf{k}}$
Interpretation of ↑	real particle in lowest state	vacuum, virtual particle
Raising operator	raises single particle energy one level	raises number of particles by one and thus, also raises energy one level
Lowering operator	lowers single particle energy one level	lowers number of particles by one and thus, also lowers energy one level
Wave form	Hermite polynomial	$e^{\pm ikx}$
Nature of wave form	real, non-sinusoid	complex, sinusoid
Motion	oscillates in one place	wave that moves
Spatial constraints	bound state, local region	unbound state, unlimited volume
Free or interaction	harmonic oscillator potential = force	free, no force

3.12.1 "Derivation" of QFT via Harmonic Oscillators

Some "derive" QFT from harmonic oscillator assumption

Some treatments actually introduce QFT via assuming states therein are harmonic oscillators. I submit this assumption can only be made after one already knows the form of the Hamiltonian (3-55) and the raising/lowering operators (3-75) as we derived them in Sect. 3.4 (pg. 53.) Otherwise,

how could anyone understand they should simply assume the QFT states have energy levels similar to those of the harmonic oscillator?

I contend that assuming harmonic oscillator behavior in QFT is an unreasonable, and unfounded, assumption, but that starting with 2nd quantization (a parallel track to what was known to work in NRQM) is a reasonable assumption. However, the former approach is common.

3.12.2 Harmonic Oscillators Have Different Behavior than States

But there are differences between QFT particle states and harmonic oscillator states

Note that the wave form for the harmonic oscillator is a Hermite polynomial, far different from the complex sinusoid of $e^{\pm ikx}$ that fields (and states) have in QFT. And, a harmonic oscillator doesn't move in space (other than up and down, or side to side, in one location), whereas waves (particles = states) do, i.e., they travel from place to place. Further, the free fields (and particles) we have been dealing with in QFT are unrestricted in space (for discrete solutions, volume V can be as large as the universe; for continuous solutions, there is no volume constraint), whereas harmonic oscillators are confined to a local region. Still further, harmonic oscillators are not free states like those we have treated, but feel force/interaction (due to the harmonic oscillator potential).[8]

For discussion of a counter argument that might be made here for the vacuum, see Appendix B.

Note the caveat: This section and the one above it are my personal position on this matter, and not, to my knowledge, shared by many others. You, the reader, should make your own call

3.12.3 Vacuum Excitations = Real Particles

States can be thought of as excitations of the vacuum (the least excited, or ground, state)

In spite of the foregoing, one can still think of real states as stable, excited states of the vacuum, since our raising operators can create a particle state from that vacuum, i.e.,

$$a^\dagger(\mathbf{k})|0\rangle = |\phi_\mathbf{k}\rangle . \tag{3-115}$$

The RHS above can be considered as the next highest state above the ground state (above the vacuum), and thus, an excited state of the vacuum. Considering such excited states specifically as *harmonic oscillator* excited states is a different matter.

3.13 The Scalar Feynman Propagator

Feynman propagator not simple to understand

The Feynman propagator, the mathematical formulation representing a virtual particle, such as the one represented by the wavy line in Fig. 1-1 of Chap. 1, pg. 2, is one of the toughest thing, in my opinion, to learn and feel comfortable with in QFT. If you don't feel comfortable with it right away, don't worry about it. That is how virtually everyone feels. Over time, it will become more familiar, and if you are lucky and work hard, maybe even easy.

Use wholeness chart as you study the derivation

I have tried to take the derivation of the propagator one step at a time, and emphasize what each step entails. Wholeness Chart 5-4 (at the end of Chap. 5) breaks these steps out clearly, and should be used as an aid when studying the propagator derivation.

Feynman propagator for QFT virtual particles is different from propagator for real particles of NRQM & RQM

Propagators: NRQM vs QFT and Real vs Virtual Particles

Note that the propagator for real particles, which you may have studied in NRQM, is *not* the same as the Feynman propagator, which is explicitly for virtual particles in QFT. It may be confusing, but the Feynman propagator is often simply called, "the propagator". You will have to get used to discerning the difference from context.

In QFT, as we will see when we study interactions, a propagator for real particles is not generally needed, and we will not derive one here.

3.13.1 The Approach

We'll use the Feynman propagator when we get to interaction theory

The first part of QFT is a free particle theory (no interactions, as in this chapter and the next three). After this, interactions are introduced. In the course of deriving the interaction theory, a mathematical relationship arises that is called the Feynman propagator. Physically, it can be visualized as representing a virtual particle that exists fleetingly and carries energy, momentum, and

[8] All of this harmonic oscillator business confused me greatly as a student. I simply could not understand how QFT states could possibly be essentially identical to harmonic oscillators. I was not confident enough to bring up the counter points mentioned herein, and they were never addressed. So if you have seen and been confused by the harmonic oscillator approach, you are not alone.

in some cases, charge from one real particle to another. Thus, it is the carrier, or mediator, of force (interaction.) See the virtual photon of Fig. 1-1 in Chap. 1, pg. 2.

It will help us pedagogically to derive the Feynman propagator now, rather than when we get to interactions. The derivation of interaction theory is fairly complicated and it will be easier, as we develop it, if we already know the mathematical relation for the Feynman propagator, rather than diverting our attention for several pages to derive it then.

But it's easier in the long run if we derive it here

Heuristically, it may help to consider the virtual particle as created at a particular spacetime point and destroyed at a later spacetime point, and this is how Feynman diagrams portray it. From this (heuristic) perspective the operator field $\phi^{\dagger}(y)$ can be considered to create a virtual scalar particle at event y (we used the symbol x_2 in Fig. 1-1), and the field operator $\phi(x)$ destroys that virtual particle at event x (x_1 in Fig. 1-1.) The scalar propagator incorporates these two field operators in a sort of "short-hand" way.

Note that the above "creation/destruction at a point" perspective can help initially in understanding the derivation of the propagator, but we caution that it will have to be modified and refined. We will save that to the end when, after digesting the derivation to follow, this modification will be easier to understand.

We will now derive a relationship for the propagator using the field operators acting on the vacuum, and will later see (Chap. 7) that this derived relationship arises naturally in the full mathematical development of the interaction theory.

3.13.2 Milestones in the Derivation

Start with a physical visualization of the propagator and follow 5 *distinct steps*

We develop the Feynman propagator in five distinct steps, starting with a physical interpretation. We represent that interpretation mathematically and then "massage" it in subsequent steps with more mathematics, until we obtain the form of the propagator that is most useful (in QFT interaction analysis).

The entire derivation is for continuous (not discrete) eigenstate solutions of the field equation (Klein-Gordon here), since the propagator represents a virtual particle in the vacuum and the vacuum is not confined to a volume V. We represent the scalar Feynman propagator with the symbol $i\Delta_F(x-y)$. (Including the imaginary factor i is common practice.)

Step 1*: Math interpretation of the physical propagator*

Step 1: Express the Feynman propagator $i\Delta_F$ as a mathematical representation of a particle or antiparticle created at one point in space and time in the vacuum and destroyed at another place and time.

Step 2: Express $i\Delta_F$ in terms of two commutators (one for particles and one for anti-particles).

Step 3: Express those two commutators as real integrals.

Step 4: Re-express those two real integrals as two contour (complex plane) integrals.

Step 5: Re-express the two contour integrals as a single integral over real, not complex, space, the form most suitable for analysis.

Step 5*: Combining two integrals in complex space into one over real space*

Step 1: The Feynman Propagator as the Vacuum Expectation Value of a Time Ordering Operator

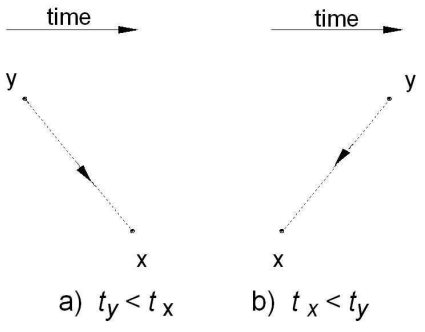


Figure 3-3. Creation & Destruction of Virtual Particle/Antiparticle

Fig 3-3a represents creation of a particle, which will be virtual, at y and destruction of it at x. Fig. 3-3b represents creation of an antiparticle at x and destruction of it at y. Virtual particles are never detected when real particles interact, so the same effect on the real particles could be realized by either of the processes in Fig. 3-3. For example, a virtual particle carrying charge from y to x would represent the same charge exchanges as an antiparticle carrying opposite charge from x to y. Thus, we need a relationship for the propagator that includes both scenarios as possibilities.

Step 1*, first part, defining the time ordering operator T and seeing how it represents creation of either a virtual particle or antiparticle followed by its destruction*

That is, we need an operator that will create a particle first if $t_y < t_x$, but create an antiparticle first if $t_x < t_y$. Our Klein-Gordon solutions (3-37), (3-85), and (3-86) provide the means for the desired creation and

destruction operations. But these have to be arranged to provide us with the time ordering dependence of Fig. 3-3. To this end, consider the time ordering operator *T*, defined as follows.

If $t_y < t_x$, $\phi^\dagger(y)$ operates first (creates a particle) and is placed on the right, with $\phi(x)$ operating second (destroys the particle) and placed on the left.

$$\text{for } t_y < t_x \quad \text{(particle)} \qquad T\{\phi(x)\phi^\dagger(y)\} = \phi(x)\phi^\dagger(y). \tag{3-116}$$

Of course, from (3-85), and (3-86), in (3-116), $\phi(x)$ also creates an antiparticle and $\phi^\dagger(x)$ also destroys an anti-particle, but we will see this effect ultimately drops out and does not play a role in the Feynman propagator.

If $t_x < t_y$, $\phi(x)$ operates first (creates an antiparticle) and is placed on the right with $\phi^\dagger(y)$ operating second (destroying the antiparticle) and placed on the left (where the effects of these operators for particles will drop out, as we will see.)

Step 1, *second part,defining the transition amplitude as equal to the VEV of time ordering operator T*

$$\text{for } t_x < t_y \quad \text{(anti-particle)} \qquad T\{\phi(x)\phi^\dagger(y)\} = \phi^\dagger(y)\phi(x). \tag{3-117}$$

We now define what is called the transition amplitude, which equals the vacuum expectation value (VEV) of the above time ordering operator. It is an amplitude, similar to the amplitude of a wave function in NRQM, because, as we will shortly see, the square of its magnitude equals the probability density of it being observed. (As the square of the magnitude of the amplitude for a component of the wave function equals the probability of it being observed.)

This transition amplitude is

$$\langle 0|T\{\phi(x)\phi^\dagger(y)\}|0\rangle, \tag{3-118}$$

We use the VEV because we will be interested in the expectation of finding a virtual particle traveling in the vacuum

which is the vacuum expectation value (VEV) of *T*, and this, as we will see below, represents both possible scenarios of Fig. 3-3. In wave mechanics for the position basis, the bracket above is an integration over all space. This is still true, but note carefully that the integration variable is over the space variable of the bra and ket (think $\mathbf{x}'$), but not the time ordering variables **x** and **y**. (See Sect. 3.7.6, pg 63.) In QFT notation, we tend to merely think of a bracket as equaling zero unless the bra and ket represent the same state.

To gain insight into (3-118), consider the transition amplitude operating on the vacuum when a virtual particle is propagated. Then, from (3-85), and (3-86), where an overbar in a state represents an antiparticle,

Gaining insight into the time ordering operator T acting on the vacuum

$$\begin{aligned} T\{\phi(x)\phi^\dagger(y)\}|0\rangle &= \phi(x)\phi^\dagger(y)|0\rangle \qquad \text{for } t_y < t_x \quad \text{(particle)} \\ &= \left(\underbrace{\phi^+(x)}_{\substack{\text{destroys}\\ \text{particle}}} + \underbrace{\phi^-(x)}_{\substack{\text{creates}\\ \text{antiparticle}}}\right)\left(\underbrace{\cancel{\phi^{\dagger+}(y)|0\rangle}}_{\substack{\text{destroys antiparticle,}\\ \text{annihilates vacuum}}} + \underbrace{\phi^{\dagger-}(y)|0\rangle}_{\substack{\text{creates}\\ \text{particle}}}\right) \\ &= \left(\phi^+(x) + \phi^-(x)\right)F(y)|\phi\rangle \\ &= G(x)F(y)|0\rangle + H(x)F(y)|\bar{\phi}\phi\rangle. \end{aligned} \tag{3-119}$$

G, *F*, and *H* are numeric factors that result from the creation and destruction operations (such as the normalization coefficients that are part of the field operators), which we will not express explicitly here. Thus, we have a general ket left, which in this case is part vacuum state, with the amplitude of the vacuum state part being *GF*, and the amplitude of the multiparticle state (scalar plus anti-scalar) part being *HF*. As we (hopefully) remember from NRQM, and which is true for all quantum theories, for appropriate normalization, the square of the magnitude of the amplitude of a state equals the probability of finding that state. Thus $(GF)^\dagger(GF)$ represents the probability of observing the vacuum state (no particles left after the transition.) To find the amplitude *GF*, we need only form an inner product of the last line of (3-119) with $\langle 0|$, i.e.,

Taking the inner product of the above $T|0\rangle$ with $\langle 0|$ to get the transition amplitude

$$\begin{aligned} \langle 0|T\{\phi(x)\phi^\dagger(y)\}|0\rangle &= \langle 0|G(x)F(y)|0\rangle + \langle 0|H(x)F(y)|\bar{\phi}\phi\rangle \\ &= G(x)F(y)\underbrace{\langle 0|0\rangle}_{=1} + H(x)F(y)\underbrace{\langle 0|\bar{\phi}\phi\rangle}_{=0} = G(x)F(y). \end{aligned} \tag{3-120}$$

Note in (3-119) that the $\phi^{\dagger-}(y)$ part of $\phi^{\dagger}(y)$, which created a particle, left the $F(y)$ factor, but the $\phi^{\dagger+}(y)$ part, which destroys an anti-particle or, in this case, annihilates the vacuum, resulted in zero. Also, the $\phi^{+}(x)$ part of $\phi(x)$, which destroyed a particle, left us with the $G(x)$ factor, but the $\phi^{-}(x)$ part created an anti-particle in the ket, and thus left zero (because the particle + antiparticle ket was orthogonal to the vacuum, the bra, leaving a bracket = 0.)

So, as we said above with reference to our original definition (3-116) of the time ordering operator, only the part of $\phi(x)$ that destroys a particle and the part of $\phi^{\dagger}(y)$ that creates a particle will be relevant.

In a similar way, the same time ordering operator can be used for antiparticle propagation (with time for x and y reversed) as in Fig 3-3b and (3-117). You can prove this by doing Prob. 17.

The square of the absolute value of the transition amplitude is a probability density (for the transition to occur)

So, the VEV of the time ordering operator is an amplitude, the square of whose magnitude is the probability of the transition from the vacuum initially (represented by $|0\rangle$) to the vacuum finally (represented by $\langle 0|$). Actually, $|G(x)F(y)|^2$ is a probability density (to be precise, a double density), because it is a function of **x** and **y**. That is, the location **y** where the virtual particle is created could be anywhere, and so could the location **x** where it is destroyed. We would need to integrate the probability density over all possible **x** and all possible **y** to get the actual probability, and this is what one does in interaction theory to calculate probabilities and cross sections.

Redefine the transition amplitude as the Feynman propagator

Given all of this, we can define our mathematical relationship for the processes shown in Fig. 3-3 as the VEV of the time ordering operator T. This is called, in honor of its discoverer, the Feynman propagator $i\Delta_F$ (where we insert a factor of i because it makes things easier later on),

$$\boxed{i\Delta_F(x-y)=\langle 0|T\{\phi(x)\phi^{\dagger}(y)\}|0\rangle}. \tag{3-121}$$

Step 2, expressing Feynman propagator in terms of commutators

Step 2: Expressing $i\Delta_F$ in Terms of Commutators

Note for $t_y < t_x$, the case for a virtual particle (not antiparticle), the Feynman propagator equals

$$\begin{aligned} i\Delta_F(x-y) &= \langle 0|\phi(x)\phi^{\dagger}(y)|0\rangle \\ &= \langle 0|\phi^{+}(x)\underbrace{\phi^{\dagger+}(y)|0\rangle}_{=0}+\langle 0|\phi^{+}(x)\underbrace{\phi^{\dagger-}(y)|0\rangle}_{=(\,factor)|\phi\rangle}+\langle 0|\phi^{-}(x)\underbrace{\phi^{\dagger+}(y)|0\rangle}_{=0}+\langle 0|\phi^{-}(x)\underbrace{\phi^{\dagger-}(y)|0\rangle}_{=(\,factor)|\phi\rangle} \\ &= \langle 0|\underbrace{\phi^{+}(x)\phi^{\dagger-}(y)|0\rangle}_{=(\,factor)|0\rangle}+\underbrace{\langle 0|\phi^{-}(x)|\phi\rangle}_{=\langle 0|\bar{\phi}\phi\rangle=0}=\langle 0|\phi^{+}(x)\phi^{\dagger-}(y)|0\rangle . \end{aligned} \tag{3-122}$$

Step 2, first part, finding Feynman propagator for virtual particle (not antiparticle) in terms of a commutator

where "factor" represents the non-operator quantities in each field operator term that are left unchanged when the creation and destruction coefficient operators act on a ket. To the last part of (3-122), we can add zero in the form of

$$0=\langle 0|-\phi^{\dagger-}(y)\underbrace{\phi^{+}(x)|0\rangle}_{=0}. \tag{3-123}$$

By adding a term equal to zero, we can use a commutator

Doing that, we find (3-122) becomes

$$i\Delta_F(x-y)=\langle 0|\phi^{+}(x)\phi^{\dagger-}(y)-\phi^{\dagger-}(y)\phi^{+}(x)|0\rangle=\langle 0|\left[\phi^{+}(x),\phi^{\dagger-}(y)\right]|0\rangle \tag{3-124}$$

Step 2, second part, expressing Feynman propagator for virtual antiparticle in terms of a commutator

In similar fashion, for $t_x < t_y$, the case for a virtual antiparticle, one finds, by doing Prob. 18, that

$$i\Delta_F(x-y)=\langle 0|\left[\phi^{\dagger+}(y),\phi^{-}(x)\right]|0\rangle . \tag{3-125}$$

In summary for Step 2, we have shown the Feynman propagator can be expressed in terms of commutators as

Step 2, summary

$$\begin{aligned} i\Delta_F(x-y) &= \langle 0|\left[\phi^{+}(x),\phi^{\dagger-}(y)\right]|0\rangle \quad \text{if } t_y<t_x \text{ (virtual particle)} \\ &= \langle 0|\left[\phi^{\dagger+}(y),\phi^{-}(x)\right]|0\rangle \quad \text{if } t_x<t_y \text{ (virtual anti-particle).} \end{aligned} \tag{3-126}$$

Step 3: Expressing Commutator Forms of $i\Delta_F$ as Integrals

Define the symbol $i\Delta^+$ as the commutator of the field type *a* solutions (for particle) of the first line of (3-126), i.e.,

Step 3, *first part, find* $i\Delta^+$ = *commutation relation for type a fields* (*particles*) *as an integral*

$$i\Delta^+(x-y)=\left[\phi^+(x),\phi^{\dagger-}(y)\right], \tag{3-127}$$

where the solutions used on the RHS are the integral (continuous) form for the Klein-Gordon solutions (3-37). It is common usage to use a + sign to designate (3-127), rather than the letter *a*, which would be easier to remember. Just think "*a* type field" when you see +. Equation (3-127) is thus

$$\begin{aligned} i\Delta^+(x-y) &= \frac{1}{2(2\pi)^3}\iint\left[a(\mathbf{k}),a^\dagger(\mathbf{k}')\right]\frac{e^{-ikx}e^{ik'y}}{\sqrt{\omega_{\mathbf{k}}\omega_{\mathbf{k}'}}}d^3\mathbf{k}\,d^3\mathbf{k}' \\ &= \frac{1}{2(2\pi)^3}\int\left(\int\frac{e^{ik'y}}{\sqrt{\omega_{\mathbf{k}}\omega_{\mathbf{k}'}}}\delta(\mathbf{k}-\mathbf{k}')\,d^3\mathbf{k}'\right)e^{-ikx}d^3\mathbf{k}, \end{aligned} \tag{3-128}$$

and hence,

$$i\Delta^+(x-y)=\frac{1}{2(2\pi)^3}\int\frac{e^{-ik(x-y)}}{\omega_{\mathbf{k}}}d^3\mathbf{k}. \tag{3-129}$$

Similarly, where a minus sign stands for *b* type fields (since they are associated with antiparticles, the minus makes some sense), we define the commutator in the second line of (3-126),

Step 3, *second part, find* $i\Delta^-$ = *commutation relation for type b fields* (*anti-particles*)

$$\begin{aligned} i\Delta^-(x-y)=\left[\phi^{\dagger+}(y),\phi^-(x)\right] &= \frac{1}{2(2\pi)^3}\iint\left[b(\mathbf{k}),b^\dagger(\mathbf{k}')\right]\frac{e^{ikx}e^{-ik'y}}{\sqrt{\omega_{\mathbf{k}}\omega_{\mathbf{k}'}}}d^3\mathbf{k}\,d^3\mathbf{k}' \\ &= \frac{1}{2(2\pi)^3}\int\frac{e^{ik(x-y)}}{\omega_{\mathbf{k}}}d^3\mathbf{k}, \end{aligned} \tag{3-130}$$

Thus, (note different authors may define the symbols $i\Delta^+$ and $i\Delta^-$ somewhat differently)

Step 3, *final part, combine above parts into one symbol* $i\Delta^\pm$

$$i\Delta^\pm(x-y)=\frac{1}{2(2\pi)^3}\int\frac{e^{\mp ik(x-y)}}{\omega_{\mathbf{k}}}d^3\mathbf{k}. \tag{3-131}$$

Note that though our earlier expressions for $i\Delta$ and $i\Delta^\pm$, contained operators that operated on the ket part of the VEV in (3-126), because the commutator of these operators in (3-128) (and similarly, in (3-130)) is a number, $i\Delta^\pm$ are simply numbers, not operators. Since the expectation value of a number is a number, $i\Delta_F$ of (3-126) is only that, a number. (To be precise it is a numeric *function*, not an operator *function*.) The bottom line is: We don't have to worry about operators, their effects, or VEV brackets any more, but can simply evaluate the Feynman propagator $i\Delta_F$ as a numeric mathematical relation.

Our VEV of operators expression of the propagator has become a simple numeric function

Step 4: Expressing the Two Real Integrals $i\Delta^\pm$ as Contour Integrals

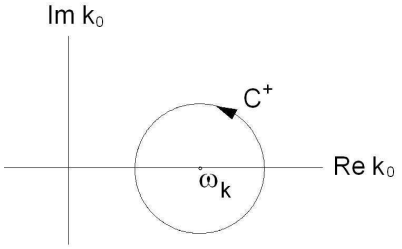


Figure 3-4. Contour Integral for Real, Positive Frequency

It will prove advantageous if we express (3-131) as contour integrals. Before doing so, we first review complex integral theory.

Review of integral in the complex plane

Consider the complex plane for a function *f* of the complex variable k_0, i.e., $f(k_0)$. Here, the symbol k_0 is not a pole (poles are usually designated with null subscript), but represents a complex number generalization of the zeroth component (the energy) of 4-momentum *k*. We concern ourselves with the particular case where k_0 takes on the real value $\omega_{\mathbf{k}}$.

From complex variable theory,

$$f(\omega_{\mathbf{k}}) = \frac{1}{i2\pi}\int_{C^+} \frac{f(k_0)}{k_0 - \omega_{\mathbf{k}}} dk_0 \,. \tag{3-132}$$

Step 4, first part, express $i\Delta^+$ as a contour integral

Now, re-express (3-129) as a regular (not complex plane) integral as

$$i\Delta^+(x-y) = \frac{1}{(2\pi)^3}\int e^{i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})} \underbrace{\left\{\frac{e^{-i\omega_{\mathbf{k}}(t_x - t_y)}}{2\omega_{\mathbf{k}}}\right\}}_{f(\omega_{\mathbf{k}})} d^3\mathbf{k} \,, \tag{3-133}$$

where we take the bracketed quantity as equal to $f(\omega_{\mathbf{k}})$, and where

$$f(k_0) = \frac{e^{-ik_0(t_x - t_y)}}{k_0 + \omega_{\mathbf{k}}} \,. \tag{3-134}$$

We can then use (3-134) in (3-132) to re-express $f(\omega_{\mathbf{k}})$ in terms of a contour integral. Using this for the bracket in (3-133), we find (3-133) becomes

$$\begin{aligned} i\Delta^+(x-y) &= \frac{1}{(2\pi)^3}\int e^{i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})} \left\{\frac{1}{i2\pi}\int_{C^+} \frac{f(k_0)}{k_0 - \omega_{\mathbf{k}}} dk_0\right\} d^3\mathbf{k} \\ &= \frac{1}{(2\pi)^3}\int e^{i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})} \left\{\frac{1}{i2\pi}\int_{C^+} \frac{e^{-ik_0(t_x - t_y)}}{(k_0 - \omega_{\mathbf{k}})(k_0 + \omega_{\mathbf{k}})} dk_0\right\} d^3\mathbf{k} \\ &= \frac{-i}{(2\pi)^4}\int_{C^+} \frac{e^{-ik(x-y)}}{(k_0)^2 - (\omega_{\mathbf{k}})^2} d^4k \,. \end{aligned} \tag{3-135}$$

where the integral notation now implies integration over four dimensions of the 4-momentum, with the 3-momentum part from $-\infty$ to $+\infty$ in real space and the energy part a contour integral in complex space. Note that the integral does not "blow up" because $k_0 \neq \omega_{\mathbf{k}}$ over the contour integral. We are using a mathematical trick that works, though it jars our usual understanding that, for real particles, the zeroth component of 4-momentum equals energy. k_0 has at this point become, for us, a variable that generally does not equal energy $\omega_{\mathbf{k}}$.

We modify (3-135) a little by noting what is always true mathematically for any four vector, and thus true for 4-momentum components,

Modifying terms in our result a little

$$k^2 = (k_0)^2 - (\mathbf{k})^2 \quad \rightarrow \quad (k_0)^2 = k^2 + (\mathbf{k})^2 \tag{3-136}$$

and what is physically true relativistically for rest mass, energy, and 3-momentum (see(3-2)),

$$\omega_{\mathbf{k}}^2 - (\mathbf{k})^2 = \mu^2 \quad \rightarrow \quad \omega_{\mathbf{k}}^2 = \mu^2 + (\mathbf{k})^2 \,. \tag{3-137}$$

Substitute the RH expressions of (3-136) and (3-137) into the last line of (3-135) to get

$$i\Delta^+(x-y) = \frac{-i}{(2\pi)^4}\int_{C^+} \frac{e^{-ik(x-y)}}{k^2 - \mu^2} d^4k \,. \tag{3-138}$$

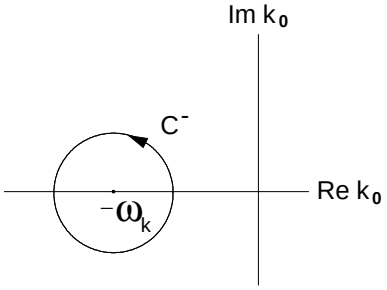


Figure 3-5. Contour Integral for Real, Negative Frequency

Step 4, second part, express $i\Delta^-$ as a contour integral

For $i\Delta^-(x-y)$, we carry out similar steps except that the contour integral (still c.c.w.) is now about $-\omega_{\mathbf{k}}$. In Appendix C, we carry out these steps. When all is said and done, we find the only differences from (3-138) to be the sign and the contour, which is now about the negative frequency value and designated by C^-.

$$i\Delta^-(x-y) = \frac{i}{(2\pi)^4}\int_{C^-}\frac{e^{-ik(x-y)}}{k^2-\mu^2}d^4k. \tag{3-139}$$

Step 5: Re-express $i\Delta_F$ in Most Convenient Form

Step 5, expressing Feynman propagator as integral over real, not complex space

We would like two things more: 1) express the propagator as a single function so we don't have to keep track (while we are integrating over spacetime and doing other things) of whether the virtual field is a particle or antiparticle (i.e., whether to use the Δ^+ or Δ^- function), and 2) have all our integrations over real numbers rather than deal with contour integrals.

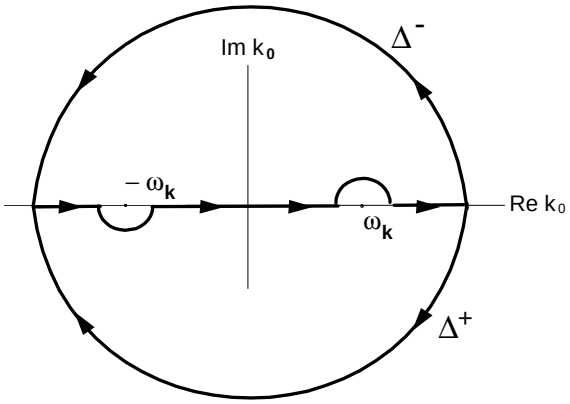


Figure 3-6. Contour Integrals for Δ^- and $-\Delta^+$

To do this, consider Fig. 3-6, where we have shown two contour integrals and the quantity being integrated in both cases is the integrand of (3-139). The top loop represents Δ^- (we have left out a factor of $(2\pi)^4$ in the figure) and encloses $-\omega_{\mathbf{k}}$ with a ccw path. The lower loop encloses $+\omega_{\mathbf{k}}$, but since it has a cw integration path, the result will have a sign change, and hence equals the ccw integration in (3-138). Thus, the lower loop represents Δ^+.

This means we can define the Feynman propagator Δ_F of (3-138) and (3-139) as proportional to the same integral over the two different loops of Fig. 3-6. We say "proportional" because we also have to include the concomitant integration over the 3D space of **k** not shown in Fig. 3-6, as well as the various constants involved.

Two different contours for the Feynman propagator written with same integral, different meaning for path C_F

So we can then re-write the Feynman propagator of (3-138) and (3-139) with Fig. 3-6, as

$$i\Delta_F(x-y) = \frac{i}{(2\pi)^4}\int_{C_F}\frac{e^{-ik(x-y)}}{k^2-\mu^2}d^4k\ , \tag{3-140}$$

where the C_F on the integral defines the route we take in the plane of Fig. 3-6.

Now, consider enlarging the outer hemispheric parts of the two loops in Fig. 3-6, so they extend essentially to infinity. The value of the contour integrals over them will remain unchanged. But the k^2 value in the denominator of (3-140) will become so large that any contribution to the integral over those parts of the path will become negligible. (See Appendix D.) Thus, we can effectively take the integral (3-140) as extending only along the real axis from $-\infty$ to $+\infty$ as in Fig. 3- 7.

Extending all parts of contours to ∞ except along real axis

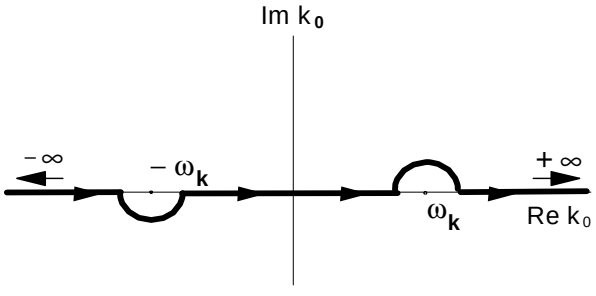


Fig. 3-7. Contour C_F for Δ_F

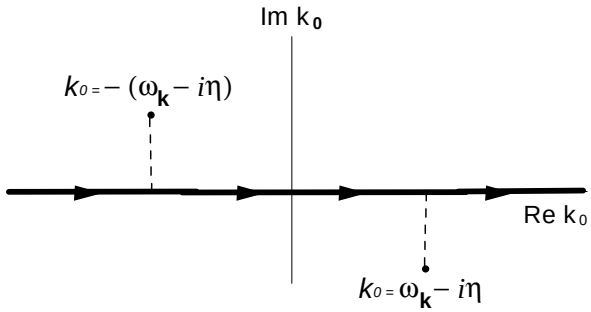


Fig. 3-8. Contour and Displaced Poles for Δ_F

We can further simplify by moving the poles an infinitesimal distance η off the real axis as shown in Fig. 3-8 and deform the contour so that it is all along the real axis. In the limit as $\eta \to 0$, the integral will have the same value, though we must now include this slight pole shift in the propagator expression (3-140). We do this by recalling from (3-136) and (3-137) that we used

Instead of integrating around poles on the axis, move poles slightly off the axis

$$k^2 - \mu^2 = (k_0)^2 - (\omega_{\mathbf{k}})^2 \tag{3-141}$$

to obtain the denominator of (3-140), so we must temporarily restate (3-140) using the right hand side of (3-141), then shift the poles. Thus, (3-140) becomes

$$i\Delta_F(x-y) = \frac{i}{(2\pi)^4}\int_{-\infty}^{+\infty} \frac{e^{-ik(x-y)}}{(k_0)^2 - (\omega_{\mathbf{k}} - i\eta)^2}\, d^4k. \tag{3-142}$$

Yields a single integral over real space representing both virtual particle & antiparticle, the most convenient form for the Feynman propagator

If we then use (3-141) again, ignore second order terms in η, and take ε = 2ηω**k**, we have our final result for the Feynman scalar propagator

$$\boxed{i\Delta_F(x-y) = \frac{i}{(2\pi)^4}\int_{-\infty}^{+\infty} \frac{e^{-ik(x-y)}}{k^2 - \mu^2 + i\varepsilon}\, d^4k.} \tag{3-143}$$

Note the advantages of this form. We now have a single mathematical relationship that automatically describes both a particle propagating from *y* to *x and* an antiparticle propagating from *x* to *y*. We also have done away with the cumbersome contour integrals in favor of a simple 4D integral over the entire real (not complex) 4-momentum space. In principle, we can evaluate this integral then take ε to zero after the integration is carried out.

Summary of Steps 1 to 5

Summary of propagator derivation

Steps 1 to 4 for the virtual particle Feynman propagator were

Steps 1 *to* 4

$$\begin{aligned} i\Delta_F(x-y) &= \langle 0|T\{\phi(x)\phi^\dagger(y)\}|0\rangle = \langle 0|\phi(x)\phi^\dagger(y)|0\rangle \quad \text{if } t_y < t_x \text{ (particle)} \\ &= \langle 0|\underbrace{\left[\phi^+(x), \phi^{\dagger -}(y)\right]}_{i\Delta^+(x-y),\ \text{a number}}|0\rangle = i\Delta^+(x-y) = \frac{1}{2(2\pi)^3}\int \frac{e^{-ik(x-y)}}{\omega_{\mathbf{k}}} d^3\mathbf{k} \\ &= \frac{-i}{(2\pi)^4}\int_{C^+} \frac{e^{-ik(x-y)}}{k^2 - \mu^2}\, d^4k\ , \end{aligned} \tag{3-144}$$

and for the virtual anti-particle Feynman propagator,

$$\begin{aligned} i\Delta_F(x-y) &= \langle 0|T\{\phi(x)\phi^\dagger(y)\}|0\rangle = \langle 0|\phi^\dagger(y)\phi(x)|0\rangle \quad \text{if } t_x < t_y \text{ (anti-particle)} \\ &= \langle 0|\underbrace{\left[\phi^{\dagger +}(y), \phi^-(x)\right]}_{i\Delta^-(x-y),\ \text{a number}}|0\rangle = i\Delta^-(x-y) = \frac{1}{2(2\pi)^3}\int \frac{e^{ik(x-y)}}{\omega_{\mathbf{k}}} d^3\mathbf{k} \\ &= \frac{i}{(2\pi)^4}\int_{C^-} \frac{e^{-ik(x-y)}}{k^2 - \mu^2}\, d^4k\ . \end{aligned} \tag{3-145}$$

The two contour integrals of (3-144) and (3-145) were combined in Step 5 to yield the single integral over real space of (3-143).

Step 5

3.13.3 Comments on the Propagator and Its Derivation

The Propagator and Interaction Theory

Our definition of Feynman propagator here will pop up in our formal derivation of interaction theory

The derivation above was formulated with an eye to interaction theory. In that theory, amplitudes are derived for various kinds of interactions between various particles. The square of the magnitude of each amplitude turns out to be the probability of that particular interaction (transition) occurring. These transition amplitudes each depend on the initial real particles, the final real particles, and the virtual particle(s) that mediate the transition. It turns out that the factor in the amplitude representing the virtual particle contribution is identical to the Feynman propagator Δ_F as we defined it in the VEV of the time ordering operator (3-121). Thus, it is also equal to (3-143), so we can simply plug the RHS of (3-143) into the overall transition amplitude as part of our analysis.

This is one reason we started with the relation $\phi\phi^\dagger$ to create and destroy a virtual scalar particle, rather than what one might initially expect, the simpler creation and destruction operator relation

$a(\mathbf{k})a^\dagger(\mathbf{k})$. Our heuristic approach was tailored to match what we knew would be coming in the mathematical development of interaction theory.

Feynman diags, & our derivation, imply creation/ destruction at a point, but more properly, waves created/destroyed & they spread out over space.

Meaning of Spacetime Points *y* and *x*

In Fig. 3-3a, we imply the virtual particle is created at *y* and destroyed at *x*. In Feynman diagrams virtual particles are depicted in this way, and at least one real incoming particle can be thought of as being destroyed at *y*, as in Fig. 1-1 of Chap. 1, pg.2, with a virtual particle created simultaneously at *y*. At *x* the virtual particle is destroyed, with the simultaneous creation of at least one outgoing real particle at *x*.

To be precise, it is more correct to think of the incoming, outgoing, and virtual particles as moving waves spread out in space. What we calculate for a given *y* and *x* is the probability density for the interaction as a function of the coordinates *y* and *x*. If **y** and **x** are closer, one would find the probability density for the interaction to occur is greater; if farther away, the probability density is less. Integrating over all **x** and **y** gives the total probability for observing the interaction.

We are really finding probability density as function of **x**, **y**

Momentum Space Form of the Propagator

From (3-143), we can readily write down the 4-momentum space form of the propagator, the Fourier transform of (3-143), which will be very useful,

$$\boxed{\Delta_F(k) = \frac{1}{k^2 - \mu^2 + i\varepsilon}}. \tag{3-146}$$

Momentum space form of the propagator

Earlier version was physical space form

Green's Functions, Correlation Functions, and Propagators

Feynman propagators have the form of functions known in mathematics as Green's functions, and you will sometimes see them referred to in that way. You may also see them referred to as correlation functions for free fields, because there is a correlation implied between events *x* and *y*.

Feynman propagator = Green function or correlation function

3.14 Chapter Summary

Scalars and Relativistic Quantum Mechanics (RQM)

Do Prob. 20 to create your own Wholeness Chart summary of scalars and RQM as in Sect. 3.1.

Scalars and Quantum Field Theory (QFT)

This part of the chapter is key. Know it, and you know most of the basic principles in QFT. Spin ½ and spin 1 field theory closely parallel that of scalars, so most of the conceptual battle is waged in this Chap. 3.

Free scalar QFT is summarized in the second column of Wholeness Chart 5-4 at the end of Chap. 5. If you can, more or less, reproduce that Wholeness Chart column without looking at it (that is, derive the essence of QFT), you have achieved something few have achieved.

QFT Grounded in 2nd Quantization

It is important to understand how the entire theory springs out of the two 2nd quantization postulates. All the operators (number, Hamiltonian, creation/destruction, 3-momentum, charge, etc) are a direct result of these postulates. So is the vacuum energy. Wholeness Chart 5-4 can help to make that transparent.

In particular, starting with the classical Lagrangian density (or Hamiltonian density), the commutation postulate gives us the rest of the theory, step-by-step, as illustrated below (where we use only *a* type particles to save space).

Although the steps shown below are specifically for scalars, the developments of QFT for spin ½ and spin 1 particles follow precisely the same conceptual steps.

Steps to QFT

$$[\phi^r(\mathbf{x},t),\ \pi_s(\mathbf{y},t)] = i\delta^r{}_s\,\delta(\mathbf{x}-\mathbf{y}) \ \rightarrow\ [a(\mathbf{k}),\ a^\dagger(\mathbf{k}')] = \delta_{\mathbf{k}\mathbf{k}'} \ \rightarrow\ H_0{}^0 \text{ \& vacuum energy } \rightarrow$$

$$N_a(\mathbf{k}) = a^\dagger(\mathbf{k})a(\mathbf{k}) \text{ as number operator } \rightarrow\ a^\dagger(\mathbf{k}),\ a(\mathbf{k}) \text{ as creation/destruction operators}$$

$$\downarrow \qquad\qquad \swarrow \quad \searrow$$

form of observable operators (for real particles) | the propagator (for virtual particles) | interaction theory – to be studied (for real & virtual particles)

Kinds of Operators in QFT

In QFT there are two kinds of operators. One kind is the usual one from NRQM and RQM representing the dynamical variables of classical theory, such as the Hamiltonian (energy), the 3-momentum operator, charge, etc. The other kind comprises creation and destruction operators.

The first kind, when operating on an eigenstate, re-produces the original state multiplied by an eigenvalue. The second kind changes the state to another state (raising or lowering the number of particles in the state.) The second kind comprises the coefficients $a(\mathbf{k})$, $a^\dagger(\mathbf{k})$, $b(\mathbf{k})$, $b^\dagger(\mathbf{k})$, as well as the fields of which they are a part, ϕ and $\phi^\dagger$. Note that operators of this kind do not have eigenvalues, since their operation on a state changes that state, rather than re-producing it (times an eigenvalue), and hence they are generally *not* observable.

Wholeness Chart 3-4. Different Kinds of Operators in QFT

	Examples	**Effect on Eigenstate**	**Observable?**
Dynamical Variable Operators	H, $\mathbf{P}$, Q, $N_a(\mathbf{k})$	eigenvalue times original eigenstate	Yes
Raising and Lowering Operators	$a^\dagger(\mathbf{k})$, $a(\mathbf{k})$, $b^\dagger(\mathbf{k})$, $b(\mathbf{k})$	new eigenstate, one more/less particle	No
Fields	ϕ and $\phi^\dagger$	as above	No

Odds and Ends

For a summary of bosons vs fermions, and Fock space, see Wholeness Charts 3-1 and 3-2.

3.15 Appendix A: Klein-Gordon Equation from H.P. Equation of Motion

3.15.1 Background Math Needed for Delta Function Relation

From Arfken and Weber, *Mathematical Methods for Physicists*, 4th ed (Academic Press 1995), pg 85,

$$\int \frac{d\delta(x'-a)}{dx'} f(x')\,dx' = -\int \frac{df(x')}{dx'}\delta(x'-a)\,dx' = -\left.\frac{df(x)}{dx}\right|_{x=a}, \tag{3-147}$$

where in our case we will have

$$x' \to \mathbf{x}' \quad a \to \mathbf{x} \quad f(x') \to \nabla'\phi(\mathbf{x}') \quad \frac{d\delta(x-a)}{dx} \to \nabla'\delta(\mathbf{x}'-\mathbf{x}), \tag{3-148}$$

so that (3-147) becomes

$$\int \nabla'\delta(\mathbf{x}'-\mathbf{x})\cdot\nabla'\phi(\mathbf{x}',t)\,d\mathbf{x}' = -\nabla\cdot\nabla\phi(\mathbf{x},t). \tag{3-149}$$

3.15.2 Deriving the Scalar Field Equation

The Heisenberg equation of motion for any operator is

$$i\frac{\partial}{\partial t}\mathcal{O} = [\mathcal{O}, H], \tag{3-150}$$

and for a complex scalar field, this is

$$i\frac{\partial}{\partial t}\phi = [\phi, H]. \tag{3-151}$$

Thus, using (3-33) for $\mathcal{H}$ to find $H = \int \mathcal{H} d^3x$, we have

$$i\frac{\partial}{\partial t}\phi(\mathbf{x},t) = \left[\phi(\mathbf{x},t), \int d^3\mathbf{x}'\left\{ \underbrace{\pi^\dagger\pi}_{\substack{\text{only non}\\ \text{-zero result}}} + \nabla'\phi^\dagger\cdot\nabla'\phi + \mu^2\phi^\dagger\phi \right\}\right] \tag{3-152}$$

where the quantities inside the integral are all functions of $\mathbf{x}'$ and t. Since $\phi(\mathbf{x},t)$ is not a function of $\mathbf{x}'$, we can evaluate the commutator inside the integral. The second and third terms inside the integral of (3-152) commute with ϕ, and thus drop out. Writing out the independent variable dependence only when needed for clarity, and using the field commutation relations for ϕ and π (reproduced below from Chap. 2, Wholeness Chart 2-5, pg. 31, last box in RH column) of

$$\left[\pi_s,\phi^r\right]=-i\delta^r{}_s\delta(\mathbf{x}'-\mathbf{x}), \tag{3-153}$$

in the second line below, where it says "subs", we find (3-152) becomes

$$\begin{aligned} i\frac{\partial}{\partial t}\phi(\mathbf{x},t)&=\int d^3\mathbf{x}'\left[\phi(\mathbf{x},t),\pi^\dagger(\mathbf{x}',t)\pi(\mathbf{x}',t)\right]=\int d^3\mathbf{x}'\Big\{\underbrace{\phi\pi^\dagger}_{\text{commute}}\pi-\pi^\dagger\underbrace{\pi\phi}_{\text{subs}}\Big\}\\ &=\int d^3\mathbf{x}'\Big\{\underbrace{\pi^\dagger\phi\pi-\pi^\dagger\phi\pi}_{0}+\pi^\dagger(\mathbf{x}',t)i\delta(\mathbf{x}'-\mathbf{x})\Big\}\\ &=i\pi^\dagger(\mathbf{x},t). \end{aligned} \tag{3-154}$$

Next, using (3-150) when the operator is the complex conjugate of the canonical momentum,

$$i\frac{\partial}{\partial t}\pi^\dagger(\mathbf{x},t)=\Big[\underbrace{\pi^\dagger(\mathbf{x},t)}_{\substack{\text{function}\\ \text{of }\mathbf{x}}},\int d^3\mathbf{x}'\Big\{\underbrace{\underbrace{\pi^\dagger\pi}_{\substack{\to 0\text{ in}\\ \text{commutator}}}+\nabla'\phi^\dagger\cdot\nabla'\phi+\mu^2\phi^\dagger\phi}_{\text{functions of }\mathbf{x}'}\Big\}\Big]\ . \tag{3-155}$$

Note that $\nabla'\pi^\dagger(\mathbf{x},t)=0$, because the derivative of a function of $\mathbf{x}$ is with respect to a primed $\mathbf{x}'$, and we can move $\pi^\dagger$ inside and outside of any quantity the 3D spatial derivative operates on. We use this several times in what follows. We then focus on the second term in (3-155) and substitute (3-149) in the third line below where it says "use math relation above". That second term is

$$\begin{aligned} &\int d^3\mathbf{x}'\Big\{\underbrace{\pi^\dagger\nabla'\phi^\dagger}_{\nabla'(\pi^\dagger\phi^\dagger)}\cdot\nabla'\phi-\nabla'\phi^\dagger\cdot\underbrace{(\nabla'\phi)\pi^\dagger}_{\nabla'\underbrace{(\phi\pi^\dagger)}_{\text{commute}}}\Big\}=\int d^3\mathbf{x}'\Big\{\nabla'\underbrace{(\pi^\dagger\phi^\dagger)}_{\substack{\text{use com}\\ \text{relations}}}\cdot\nabla'\phi-\underbrace{\nabla'\phi^\dagger\cdot\nabla'(\pi^\dagger\phi)}_{\nabla'(\phi^\dagger\pi^\dagger)\cdot\nabla'\phi}\Big\}\\ &=\int d^3\mathbf{x}'\Big\{\nabla'(\phi^\dagger\pi^\dagger)\cdot\nabla'\phi-\underbrace{\nabla' i\delta(\mathbf{x}'-\mathbf{x})\cdot\nabla'\phi}_{\text{use math relation above}}-\nabla'(\phi^\dagger\pi^\dagger)\cdot\nabla'\phi\Big\}=i\nabla^2\phi(\mathbf{x},t). \end{aligned} \tag{3-156}$$

By doing Prob. 6 at the end of the chapter, the reader can verify that evaluation of the third term in the RHS of (3-155), using similar (but simpler) steps, leads to

$$i\frac{\partial}{\partial t}\pi^\dagger(\mathbf{x},t)=i\left(\nabla^2-\mu^2\right)\phi(\mathbf{x},t). \tag{3-157}$$

Substituting the time derivative of (3-154) into (3-157), one gets the Klein-Gordon equation

$$\frac{\partial^2}{\partial t^2}\phi=\left(\nabla^2-\mu^2\right)\phi, \tag{3-158}$$

thus showing that the equation of motion of a scalar field in the Heisenberg picture, expressed in terms of commutation relations, is equivalent to the Klein-Gordon equation.

3.16 Appendix B: Vacuum Quanta and Harmonic Oscillators

One might argue that two vacuum quanta traveling waves of energy $\frac{1}{2}\hbar\omega_{\mathbf{k}}$ with 3-momenta $\mathbf{k}$ and $-\mathbf{k}$ could be superimposed to yield a vacuum standing wave, i.e., a harmonic oscillator (distributed in space). But the total energy of the standing wave would then be $\hbar\omega$, which is not the ground state of a quantum oscillator, and thus the parallel disappears. Further, the wave form would still be sinusoidal in nature, not that of a Hermite polynomial. Nor is any potential involved.

One might instead argue that the two traveling wave eigenstates are superimposed to comprise a general quantum state, wherein the probability of measuring each of the states is ½. In this case, the expectation value for energy of the standing wave would be $\frac{1}{2}(\frac{1}{2}\hbar\omega_{\mathbf{k}})+\frac{1}{2}(\frac{1}{2}\hbar\omega_{\mathbf{k}})=\frac{1}{2}\hbar\omega_{\mathbf{k}}$. But, this violates (3-55), which tells us the energy must be $\hbar\omega_{\mathbf{k}}$. Thus, this interpretation is inconsistent with contemporary QFT. The lack of Hermite polynomial form and potential points apply here, as well.

Still further, this logic implies the entire vacuum state is one general state comprised of all the ½ energy eigenstates. This would mean the expectation value for the energy of the vacuum is the average energy of all those eigenstates, not the sum of them. So, if one assumes (which is common) an upper limit on these energies of the Planck scale energy (see Appendix A of Chap. 10), the total vacuum energy of the universe could then not exceed the Planck energy, which is about that of a very small bit of dust and hardly anything to make a fuss about.

3.17 Appendix C: Propagator Derivation Step 4 for Δ^-

We derive (3-139) from the scalar propagator derivation step 3, (3-131),

$$i\Delta^-(x-y)=\frac{1}{2(2\pi)^3}\int\frac{e^{ik(x-y)}}{\omega_{\mathbf{k}}}d^3\mathbf{k}=\frac{1}{2(2\pi)^3}\int\frac{e^{i\omega_{\mathbf{k}}(t_x-t_y)}}{\omega_{\mathbf{k}}}e^{-i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})}d^3\mathbf{k}$$

$$=\frac{1}{2(2\pi)^3}\int_{-\infty}^{\infty}\underbrace{\frac{e^{i\omega_{\mathbf{k}}(t_x-t_y)}}{\omega_{\mathbf{k}}}}_{\substack{\text{even, same value}\\ \text{for }\mathbf{k}\text{ or }-\mathbf{k}}}\underbrace{\underbrace{\left(cos(-\mathbf{k})\bullet(\mathbf{x}-\mathbf{y})\right)}_{\substack{\text{even, same value}\\ \text{for }\mathbf{k}\text{ or }-\mathbf{k}}}+\underbrace{i\,sin(-\mathbf{k})\bullet(\mathbf{x}-\mathbf{y})\right)}_{\substack{\text{odd, integral of this}\\ \text{contribution = 0 for }\mathbf{k}\text{ or }-\mathbf{k}}}}_{\text{integral same if take this }=cos\,\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})+i\,sin\,\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})=e^{i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})}}d^3\mathbf{k} \qquad (3\text{-}159)$$

$$=\frac{1}{(2\pi)^3}\int\underbrace{\frac{e^{i\omega_{\mathbf{k}}(t_x-t_y)}}{2\omega_{\mathbf{k}}}}_{=g(-\omega_{\mathbf{k}})}e^{i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})}d^3\mathbf{k}\ .$$

From complex variable theory

$$g(-\omega_{\mathbf{k}})=\frac{1}{i2\pi}\int_{C^-}\frac{g(k_0)}{k_0-(-\omega_{\mathbf{k}})}dk_0\quad\text{with}\quad g(k_0)=\frac{e^{-ik_0(t_x-t_y)}}{-k_0+\omega_{\mathbf{k}}}\ , \qquad (3\text{-}160)$$

which we can check is correct (i.e., equals the underbracket quantity in the last part of (3-159)) via

$$g(-\omega_{\mathbf{k}})=\frac{e^{-i(-\omega_{\mathbf{k}})(t_x-t_y)}}{-(-\omega_{\mathbf{k}})+\omega_{\mathbf{k}}}\overset{\text{checks}}{=}\frac{e^{i\omega_{\mathbf{k}}(t_x-t_y)}}{2\omega_{\mathbf{k}}}\ . \qquad (3\text{-}161)$$

Putting the RH quantity in (3-160) into the LH quantity in (3-160), we have

$$g(-\omega_{\mathbf{k}})=\frac{1}{i2\pi}\int_{C^-}\frac{1}{k_0+\omega_{\mathbf{k}}}\frac{e^{-ik_0(t_x-t_y)}}{-k_0+\omega_{\mathbf{k}}}dk_0=\frac{i}{2\pi}\int_{C^-}\frac{e^{-ik_0(t_x-t_y)}}{(k_0)^2-(\omega_{\mathbf{k}})^2}dk_0\ . \qquad (3\text{-}162)$$

Using (3-162) in the last part of (3-159) results in

$$i\Delta^-(x-y)=\frac{i}{(2\pi)^4}\int_{C^-}\underbrace{\frac{e^{-ik(x-y)}}{(k_0)^2-(\omega_{\mathbf{k}})^2}}_{=k^2-\mu^2}d^4k\ , \qquad (3\text{-}163)$$

where the underbracket part comes from (3-136) and (3-137). (3-163) is (3-139).

3.18 Appendix D: Enlarging the Integration Path of Fig. 3-6

The integral of Fig. 3-6 expressed in (3-140), can be written as

$$i\Delta_F(x-y)=\frac{i}{(2\pi)^4}\int_{C_F}\frac{e^{-ik_0(t_x-t_y)}e^{i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})}}{k^2-\mu^2}dk_0d^3x=\frac{i}{(2\pi)^4}\int_{C_F}\frac{e^{-i(\mathrm{Re}k_0+i\mathrm{Im}k_0)(t_x-t_y)}e^{i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})}}{k^2-\mu^2}dk_0d^3$$

$$=\frac{i}{(2\pi)^4}\int_{C_F}\frac{e^{-i(\mathrm{Re}k_0)(t_x-t_y)}e^{(\mathrm{Im}k_0)(t_x-t_y)}e^{i\mathbf{k}\bullet(\mathbf{x}-\mathbf{y})}}{k^2-\mu^2}dk_0d^3x. \qquad (3\text{-}164)$$

The factor with Rek_0 in the exponent oscillates and will be swamped by the denominator wherever $k_0^2\to\infty$. But factor with Imk_0 in the exponent is real, so we have to be careful about it.

For the lower half plane of Fig. 3-6, the integral $i\Delta_F(x\text{-}y)$ represents $\Delta^+(x\text{-}y)$, where $t_x > t_y$ and $\text{Im}k_0$ is negative. That means in (3-164) the factor with $\text{Im}k_0$ has a negative value in the exponent and will go to zero as $\text{Im}k_0 \to -\infty$.

For the upper half plane of Fig. 3-6, the integral $i\Delta_F(x\text{-}y)$ represents $\Delta^-(x\text{-}y)$, where $t_y > t_x$ and $\text{Im}k_0$ is positive. That means in that half of the plane that the factor with $\text{Im}k_0$ has a negative value in the exponent too and will go to zero as $\text{Im}k_0 \to +\infty$.

Thus the integral over the contour vanishes whenever ${k_0}^2 \to \infty$.

3.19 Problems

1. Substitute (3-9) into the non-relativistic Schrödinger equation (3-1), and also the relativistic Klein-Gordon equation (3-8), to prove to yourself that only terms with exponential form $-i(E_n t - \mathbf{p}_n\cdot\mathbf{x})/\hbar$ solve the Schrödinger equation, but all terms in (3-9) solve the Klein-Gordon equation. Do you see that the single time derivative in the former equation, and the second order time derivative in the latter, are responsible for this?

2. Prove that the orthonormality conditions (3-15) of states $\phi_{\mathbf{k},A}$ also apply to states $\phi_{\mathbf{k},B\dagger}$.

3. Repeat steps (3-24) and (3-27) using the terms with coefficients $B_{\mathbf{k}}^{\dagger}$ in (3-12) instead of those with $A_{\mathbf{k}}$. You should find total probability of negative unity.

4. Express the Klein-Gordon equations (3-35) and their discrete solutions (3-36) in cgs units (i.e., with c and $\hbar \neq 1$) and plug the latter into the former to show that $\mu^2 = m^2c^2/\hbar^2$.

5. Prove that the continuous solutions (3-37) solve the Klein-Gordon equations.

6. Show that the 3rd term in (3-155) of the Appendix equals $-i\mu^2\phi(\mathbf{x},t)$.

7. Derive the commutators for the continuous solutions to the Klein-Gordon field equation from the second postulate of 2nd quantization. (Warning: This problem may not be worth the significant investment in time needed.)

8. Starting with the mass term in (3-48), derive (3-53).

9. Find the VEV (vacuum expectation value) of the free field scalar Hamiltonian.

10. Show that $a^{\dagger}(\mathbf{k})$ creates an a type particle with 3-momentum $\mathbf{k}$, $b(\mathbf{k})$ destroys a b type particle with 3-momentum $\mathbf{k}$, and $b^{\dagger}(\mathbf{k})$ creates a b type particle with 3-momentum $\mathbf{k}$. Follow steps similar to those in (3-71) to (3-74).

11. Show $a(\mathbf{k})\left|n_{\mathbf{k}}\right\rangle = \sqrt{n_{\mathbf{k}}}\left|n_{\mathbf{k}}-1\right\rangle$. Does it follow in a heart beat that $b(\mathbf{k})\left|\bar{n}_{\mathbf{k}}\right\rangle = \sqrt{\bar{n}_{\mathbf{k}}}\left|\bar{n}_{\mathbf{k}}-1\right\rangle$?

12. Substitute the free field solutions (3-36) to the Klein-Gordon equation into the probability density operator relation (3-89) and then insert that into (3-90) to find the effective probability density operator expressed in terms of number operators (3-91). It will help you in doing so to note that for any term where $\mathbf{k} \neq \mathbf{k}'$, the destruction and creation operators will cause the ket to be different from the bra, so the resulting term in the expectation value $\bar{\rho}$ will be zero. Hence, those terms can be ignored in determining an effective ρ.

    Note that the result you get is restricted to situations where all particles (in the ket) are in $\mathbf{k}$ eigenstates, which is almost invariably the case in QFT problems and applications. With particles in general (non $\mathbf{k}$ eigen) states, ρ becomes more complicated.

13. Using (3-100), the expression for 3-momentum in terms of the fields and their conjugate momenta, and the Klein-Gordon field equation solutions, prove (3-101), the number operator form of the 3-momentum operator.

14. For the state $\left|2\phi_{\mathbf{k}_1}, 3\bar{\phi}_{\mathbf{k}_1}, \bar{\phi}_{\mathbf{k}_2}\right\rangle$, determine the expectation value of $\mathbf{P}$, the 3-momentum operator.

15. Show that for real (not complex) scalar fields, in order for π to be equal to $\dot{\phi}$, the constant *K* in the scalar Lagrangian density (3-30) must be ½. In general, in QFT, for real fields, we take *K*=½.

16. Show that if instead of the 2nd quantization, postulate #2 of commutator relations (3-40), we had anti-commutators between the field and its conjugate momentum, i.e.,

$$\left[\phi^r(\mathbf{x},t),\pi_s(\mathbf{y},t)\right]_+ = \phi^r\pi_s + \pi_s\phi^r = i\delta^r{}_s\delta(\mathbf{x}-\mathbf{y}) \tag{3-165}$$

then the coefficient commutators would be anti-commutator relations, i.e.,

$$\left[a(\mathbf{k}),a^\dagger(\mathbf{k}')\right]_+ = -\left[b(\mathbf{k}),b^\dagger(\mathbf{k}')\right]_+ = \delta_{\mathbf{k}\mathbf{k}'}\ . \tag{3-166}$$

(Hint: Just use opposite signs in (3-43) 2nd row and then in last two rows inside the bracket just before the last equal sign. Then, all commutators in (3-45) to (3-47) become anti-commutators.

17. Find the transition amplitude operating on the vacuum when a virtual anti-particle is propagated as shown in Fig. 3-3b. Use symbols for numeric factors resulting from creation and destruction operators acting on the vacuum and other states.

18. Prove (3-125).

19. Reproduce the essence, with the best detail you can muster, of the Spin 0 column in Wholeness Chart 5-4 without looking at it. That is, prove to yourself that you know how the free field part of QFT is developed.

20. Create your own Wholeness Chart summary of RQM, as presented in Sect. 3.1. Take each sub-section heading of Sect. 3.1 as a block in the left hand column of your chart. Put the main result(s) of that section in the block just to its right in the next column. In between main results insert blocks with short notes on how one gets from the material above to the result in the block below. If there are other comments you wish to add, put them in another column to the right of the others.

Chapter 4 *print vers 3/6/13 copyright of Robert D. Klauber*

Spinors: Spin ½ Fields

Niels Bohr: "What are you working on Mr. Dirac?"
Paul Dirac: "I'm trying to take the square root of something"

4.0 Preliminaries

While it may seem humorous to think of a physics Nobel laureate struggling over a square root problem, Dirac's meaning here was actually quite deep.

Dirac sought a first order relativistic Schrödinger equation

The quotes above purportedly came during a break at a 1927 conference Bohr and Dirac attended. Dirac later recalled that he continued on by saying he was trying to find a relativistic quantum theory of the electron, and Bohr commented, "But Klein has already solved that problem." Dirac then tried to explain he was not satisfied with the (Klein-Gordon) solution because it involved a 2nd order equation in time. That led to negative energy solutions, and he sought a 1st order equation like the non-relativistic Schrödinger equation. But the conference reconvened just then, and the discussion ended.

4.0.1 Background

Recall from Chap. 3, Sects. 3.0.1 (pg. 40) and 3.1.2 (pg. 42) that we had to use H^2 to develop our relativistic wave equation, because the relativistic Hamiltonian H entailed the operator $\partial_i\partial_i$ under a square root sign, and that had no meaning. Dirac wanted to find a meaningful H to use in a relativistic Schrödinger equation of form

In other words, he sought a wave equation in H, not H^2

$$H\psi = i\frac{\partial}{\partial t}\psi\,, \tag{4-1}$$

rather than

$$H^2\phi = -\frac{\partial^2}{\partial t^2}\phi \qquad \text{(Klein-Gordon eq)}\,. \tag{4-2}$$

It is no secret that he succeeded, and his famous result, published in early 1928, is now known as the Dirac equation. We will study it in depth in this chapter.

His equation turned out to be specifically for spin ½ particles, not all particles

It wasn't too long after Dirac's discovery of the correct form for (4-1), that people realized (4-2) actually describes scalars; and (4-1), spin ½ fermions, such as the electron. The mathematical nature of the Dirac equation (4-1) provided a good indication for this. That is, (4-1) turns out (as we will see) to be a matrix equation with H being a square matrix quantity and ψ, a column matrix.

In NRQM, we represented up and down spin of particles via wave functions that had a two component column matrix "tacked on". $(1,0)^T$ represented spin up; and $(0,1)^T$, spin down. So, if ψ in (4-1) in RQM (and QFT) turns out to be a column matrix (and it does), then we could make a good bet that it will represent spinors, rather than scalars. We would be smart to make such a bet, as we would end up winning it.

The Dirac equation is a matrix equation

Interestingly, the column matrix solutions ψ to (4-1) turn out to have four components, rather than two. Given that the relativistic (scalar) solutions to the relativistic wave equation we found in Chap. 3 provided us with antiparticles, which essentially doubled our total number of fields/particles, this should not be too surprising. Four spin components is just what we need to represent particles with up or down spin (2 components) plus antiparticles with up or down spin (2 more components.)

4.0.2 Chapter Overview

Our approach to spin ½ fermions in this chapter will parallel that for spin 0 bosons. You may find it helpful to compare and contrast the bulleted material below with that of the Chapter Overview for scalars at the beginning of Chap. 3, pg. 41.

Spinor theory development parallels scalar theory

RQM first,

where we will look at

RQM overview (spinors)

- the lack of a classical theory of fermions (no macroscopic fermionic behavior observed) and thus, being unable to use a classical H in 1st quantization,
- deducing the Dirac equation, a relativistic Schrödinger equation in H, not H^2,
- solutions (states in RQM) to the Dirac equation,
- probability density and its connection to the normalization constant in the solutions,
- negative energies and the Dirac equation solutions,
- how the Dirac solutions (unexpectedly at first) represent spin ½ particles, and
- spin and the spin operator acting on the solutions (which we didn't have with scalars).

Then QFT,

QFT overview (spinors)

- noting the lack of classical, macroscopic fermionic fields and thus, being unable to use a classical $\mathcal{H}$ in 2nd quantization,
- assuming the RQM Dirac equation as the QFT field equation, with the same solution form,
- using the (Dirac) field equation to deduce the QFT $\mathcal{L}$ for spinors (the reverse route from the scalar case), and employing the Legendre transformation to get $\mathcal{H}$,
- assuming solution coefficients obey anti-commutation (instead of commutation) relations,
- determining relevant operators in QFT: $H = \int \mathcal{H}\, d^3x$, number, creation/destruction, etc.,
- showing this approach avoids real particle negative energy states,
- seeing how the vacuum is filled with spinor quanta of energy $-\frac{1}{2}\hbar\omega$,
- deriving other operators (probability density, 3-momentum, charge, spin), and
- showing spinors are fermions, and they won't work with commutation relations.

And then,

- finding the spinor Feynman propagator.

Free (no force) Fields

Still only free particles/fields in this chapter

As in Chap. 3, we look herein only at free spinors.

4.1 Relativistic Quantum Mechanics for Spinors

4.1.1 No Classical Spinor Fields: Can We Quantize?

No classical theory for spin ½ particles/fields, because fermions can't occupy same state

In Chap. 3, Wholeness Chart 3.1 (pg. 65), we recalled that, via the Pauli exclusion principle, fermions cannot occupy the same state within the same macro system. So, whereas photons (bosons) can occupy the same state and a lot of them can therefore reinforce one another to produce a macroscopic electromagnetic field, spinors (fermions) cannot do so. In other words, we have no classical macroscopic spinor fields to sense, interact with, and study experimentally. And thus, we have no classical theory of spinors.

First quantization started with the classical Hamiltonian (or equivalently, the Lagrangian) and used that as the quantum Hamiltonian. But we have no classical spinor theory and thus no classical spinor Hamiltonian. Precisely parallel statements can be made for 2nd quantization. There is simply no classical theory with spinor Hamiltonian and Lagrangian densities.

So we can't do 1st or 2nd quantization

So we can't do 1st or 2nd quantization for spinors in the way it was advertised earlier, i.e., as THE way to obtain a good quantum theory. (My apologies for the false advertising, but you would have been confused at the time, otherwise.)

Dirac found another way

So how do we deduce a relativistic spinor quantum theory? We answer this question in the next section by showing how Dirac did it (though he was actually trying to do something else.)

4.1.2 Dirac's Approach to RQM: Another History Lesson

Dirac's primary goal was a 1st order relativistic Schrödinger equation, and he postulated that if it existed, it must have the general form (where, as before, we use the ket form symbolism for the wave equation solution in particle quantum theory)

General form a 1st order RQM equation must have

$$i\frac{\partial}{\partial t}|\psi\rangle = H|\psi\rangle = (\boldsymbol{\alpha}\bullet\mathbf{p} + \beta m)|\psi\rangle . \tag{4-3}$$

In (4-3), **p** is particle three momentum, and the vector $\boldsymbol{\alpha}$ and the scalar β would have to be determined. Thus, the equation would be first order in the time derivative (and hopefully yield only positive energy solutions). Also, the relativistic free particle *H* would be a linear function of both **p** and mass *m*. The key question then is 'what are $\boldsymbol{\alpha}$ and β?' in order for this equation to be true.

Square of this equation must equal K-G eq

To find the answer, Dirac reasoned that H^2 and $|\psi\rangle$ must also satisfy the usual relativistic energy momentum relation (and therefore the Klein-Gordon equation)

$$-\frac{\partial^2}{\partial t^2}|\psi\rangle = H^2|\psi\rangle = (\mathbf{p}^2 + m^2)|\psi\rangle . \tag{4-4}$$

Squaring the operators in (4-3) and inserting the results into (4-4), we get

$$\begin{aligned} -\frac{\partial^2}{\partial t^2}|\psi\rangle = H^2|\psi\rangle &= (\alpha_i p_i + \beta m)(\alpha_j p_j + \beta m)|\psi\rangle \\ &= \left(\alpha_i^2 p_i^2 + \underbrace{\left(\overbrace{\alpha_i\alpha_j + \alpha_j\alpha_i}^{i>j \text{ here}} \right)}_{\text{must}=0,} p_i p_j + \underbrace{(\alpha_i\beta + \beta\alpha_i)}_{\text{must}=0} p_i m + \beta^2 m^2 \right)|\psi\rangle , \end{aligned} \tag{4-5}$$

This squaring restricts form of terms in general equation

where comparison with the RHS of (4-4) shows the bracketed quantities in the lower line above must equal zero. That comparison also shows that $\alpha_i^2 = 1$ and $\beta^2 = 1$. In summary, where anti-commutators are defined as $[\alpha_i,\alpha_j]_+ = \alpha_i\alpha_j + \alpha_j\alpha_i$,

$$\begin{aligned} &\left[\alpha_i,\alpha_j\right]_+ = [\alpha_i,\beta]_+ = 0 \quad i \neq j \quad \alpha_1,\alpha_2,\alpha_3,\beta \text{ all anti-commute with each other,} \\ &(\alpha_1)^2 = (\alpha_2)^2 = (\alpha_3)^2 = (\beta)^2 = 1 \text{ (the identity matrix).} \end{aligned} \tag{4-6}$$

The α_i, β thus must be matrices with certain properties

If α_i and β were numbers they would have to commute and could not possibly anti-commute. Hence, they can only be matrices. Since these matrices are operators operating on $|\psi\rangle$, then $|\psi\rangle$ itself must be a multicomponent object (i.e., a column matrix, at least.)

Using (4-6), one can show that the α_i and β matrices are traceless, hermitian, have ±1 eigenvalues, and must have an even dimension of at least four. It will save time if you can simply accept these results. If not, then please prove them to yourself. I do note that I, myself, have never done so.

Dirac & Pauli found a set of 4X4 matrices that worked

Choosing the minimum dimension case (four), Dirac and Pauli came up with a set of matrices which solve all of the above conditions (specifically (4-6)) which is now called the standard (or Dirac-Pauli) representation, and which we will study in some depth in this chapter. There are, however, other possible choices for α_i and β that satisfy the same conditions. Two of these, called the Weyl and Majorana representations, are also four dimensional and can be convenient for some advanced applications, but we will ignore them herein.

The 4D abstract space of the solutions is called spinor space

Square matrices in a 4D space must be 4X4, and thus from (4-3), if $|\psi\rangle$ is a column matrix (a vector), it must have four components (a 4D vector). Take care to note that the 4D space we are talking about here is *not* the four dimensional physical space of relativity theory, but an abstract space, often called spinor space.

The matrices Dirac and Pauli found for spinor space are

$$\beta=\begin{bmatrix}1&&&\\&1&&\\&&-1&\\&&&-1\end{bmatrix}\quad \alpha_1=\begin{bmatrix}&&&1\\&&1&\\&1&&\\1&&&\end{bmatrix}\quad \alpha_2=\begin{bmatrix}&&&-i\\&&i&\\&-i&&\\i&&&\end{bmatrix}\quad \alpha_3=\begin{bmatrix}&&1&\\&&&-1\\1&&&\\&-1&&\end{bmatrix}, \tag{4-7}$$

Form of the α_i, β matrices

where blank components equal zero. (4-7) is commonly written using the 2X2 Pauli matrices σ_i as

$$\beta=\begin{bmatrix}I&0\\0&-I\end{bmatrix}\quad \alpha_1=\begin{bmatrix}0&\sigma_1\\\sigma_1&0\end{bmatrix}\quad \alpha_2=\begin{bmatrix}0&\sigma_2\\\sigma_2&0\end{bmatrix}\quad \alpha_3=\begin{bmatrix}0&\sigma_3\\\sigma_3&0\end{bmatrix}, \tag{4-8}$$

where 0 represents the 2X2 null matrix.

Solutions to Dirac equation also solve K-G equation

Note that the Klein-Gordon equation can be considered the "square" of the Dirac equation and hence any solution $|\psi\rangle$ which solves the Dirac equation also solves the Klein-Gordon equation.

4.1.3 More Convenient Way to Express the Matrices

Dirac matrices γ^μ, found from α_i and β, are better to work with

The Dirac equation can be expressed in a more convenient way by premultiplying (4-3) by β. To help when we do that, we define four matrices, called Dirac matrices or gamma matrices, as

$$\gamma^0=\beta\qquad \gamma^1=\beta\alpha_1\qquad \gamma^2=\beta\alpha_2\qquad \gamma^3=\beta\alpha_3\ , \tag{4-9}$$

where you can do Prob. 2 to show these equal

$$\gamma^0=\begin{bmatrix}1&&&\\&1&&\\&&-1&\\&&&-1\end{bmatrix}\quad \gamma^1=\begin{bmatrix}&&&1\\&&1&\\&-1&&\\-1&&&\end{bmatrix}\quad \gamma^2=\begin{bmatrix}&&&-i\\&&i&\\&-i&&\\i&&&\end{bmatrix}\quad \gamma^3=\begin{bmatrix}&&1&\\&&&-1\\-1&&&\\&1&&\end{bmatrix}, \tag{4-10}$$

Form of Dirac matrices

or commonly, as

$$\gamma^0=\begin{bmatrix}I&0\\0&-I\end{bmatrix}\quad \gamma^1=\begin{bmatrix}0&\sigma_1\\-\sigma_1&0\end{bmatrix}\quad \gamma^2=\begin{bmatrix}0&\sigma_2\\-\sigma_2&0\end{bmatrix}\quad \gamma^3=\begin{bmatrix}0&\sigma_3\\-\sigma_3&0\end{bmatrix}. \tag{4-11}$$

From henceforth, we will do virtually nothing with the α_i and β matrices, and focus on the γ^μ matrices, instead.

Note the Hermiticity conditions (which you can prove by doing Prob. 3),

Complex conjugate transpose relations for Dirac matrices

$$\gamma^{\mu\dagger}=\gamma^0\gamma^\mu\gamma^0\ . \tag{4-12}$$

4.1.4 The Dirac Equation Expressed with Dirac Matrices

Dirac's original 1st order equation (4-3) in terms of $\boldsymbol{\alpha}$ and β, pre-multiplied by β, takes on the form

$$i\underbrace{\beta}_{\gamma^0}\frac{\partial}{\partial t}|\psi\rangle=\left(\underbrace{\beta\alpha_i}_{\gamma^i}p_i+\underbrace{\beta^2}_{\mathrm{I}}m\right)|\psi\rangle=\left(-i\gamma^i\frac{\partial}{\partial x^i}+m\right)|\psi\rangle, \tag{4-13}$$

or rearranged as what is formally called the Dirac equation

Dirac equation in terms of Dirac matrices & all indices written out

$$\boxed{\sum_{\eta=1}^{4}\left(\sum_{\mu=0}^{3}i\left(\gamma^\mu\right)_{\kappa\eta}\partial_\mu-m\delta_{\kappa\eta}\right)|\psi\rangle_\eta=0\qquad \kappa=1,2,3,4}, \tag{4-14}$$

where we have written out the 4X4 spinor space indices in κ and η, and the summation signs, in order to make it explicitly clear what is going on in spinor space. Note that the Dirac equation is actually *four separate non-matrix equations*, one for each value of the index κ. And each of these equations entails a sum of matrix components (sum over μ), each post multiplied by one the four components (in η index) of the column vector $|\psi\rangle$. Yes, it seems complicated. But also yes, it works. And also, yes, it is considered beautiful by many.

Dirac equation is actually four non-matrix equations

You will get used to the complication in time. When you do, you should gain an appreciation for the beauty, as well. In the words of the equation's discoverer,

> *"The research worker, in his efforts to express the fundamental laws of Nature in mathematical form, should strive mainly for mathematical beauty. He should take simplicity into consideration in a subordinate way to beauty ... It often happens that the requirements of simplicity and beauty are the same, but where they clash, the latter must take precedence. "*
>
> — Paul A. M. Dirac

Common, short hand form of Dirac equation

You should do Prob. 4 to provide some practice with (4-14), and then note that the common way to write the Dirac equation is to hide the spinor space indices in κ and η, i.e.,

$$\boxed{\left(i\gamma^{\mu}\partial_{\mu}-m\right)|\psi\rangle=0}\ , \tag{4-15}$$

where you have to be vigilant to remember the implicit 4X4 spinor space matrix/column nature of (4-15) as expressed explicitly in (4-14).

Slash notation also very common in Dirac equation

Another notation commonly used, which is the most streamlined of all, is

$$\not{\partial}=\gamma^{\mu}\partial_{\mu}\quad \text{so, the Dirac equation}\rightarrow\left(i\not{\partial}-m\right)|\psi\rangle=0\ . \tag{4-16}$$

We note in passing that

$$m\rightarrow\frac{mc}{\hbar}\quad \text{in non-natural units in the Dirac equation}\ . \tag{4-17}$$

4.1.5 Solutions to the Dirac Equation

Writing out Dirac equation as an eigenvalue problem

(4-15) can be considered an eigenvalue problem with eigenvalue zero, or by re-expressing it as

$$i\gamma^{\mu}\partial_{\mu}|\psi\rangle=m|\psi\rangle\ , \tag{4-18}$$

as an eigenvalue problem with eigenvalue *m*. Writing out (4-18), we have

$$i\left(\gamma^{0}\partial_{0}+\gamma^{1}\partial_{1}+\gamma^{2}\partial_{2}+\gamma^{3}\partial_{3}\right)|\psi\rangle=m|\psi\rangle=$$

$$i\left(\begin{bmatrix}1&&&\\&1&&\\&&-1&\\&&&-1\end{bmatrix}\partial_{0}+\begin{bmatrix}&&&1\\&&1&\\&-1&&\\-1&&&\end{bmatrix}\partial_{1}+\begin{bmatrix}&&&-i\\&&i&\\&-i&&\\i&&&\end{bmatrix}\partial_{2}+\begin{bmatrix}&&1&\\&&&-1\\-1&&&\\&1&&\end{bmatrix}\partial_{3}\right)\begin{vmatrix}\psi_1\\\psi_2\\\psi_3\\\psi_4\end{vmatrix}\rangle \tag{4-19}$$

$$=i\left(\begin{bmatrix}\partial_0&0&\partial_3&\partial_1-i\partial_2\\0&\partial_0&\partial_1+i\partial_2&-\partial_3\\-\partial_3&-\partial_1+i\partial_2&-\partial_0&0\\-\partial_1-i\partial_2&\partial_3&0&-\partial_0\end{bmatrix}\right)\begin{vmatrix}\psi_1\\\psi_2\\\psi_3\\\psi_4\end{vmatrix}\rangle=m\begin{vmatrix}\psi_1\\\psi_2\\\psi_3\\\psi_4\end{vmatrix}\rangle.$$

Note that the numeric subscripts on the ∂ symbols refer to derivatives with respect to time and space, whereas the numeric subscripts on the components of $|\psi\rangle$ refer to the respective components of the ket in spinor space.

Since (4-19) is a 4X4 matrix eigenvalue problem, there must be four solutions $|\psi^{(n)}\rangle$, where *n* = 1,2,3,4, with each such solution having four spinor space components. We will not go through the tedium of solving (4-19) directly. Rather, I will provide the solutions, and you will do Prob. 5 to prove to yourself, by substitution, that they are indeed valid solutions.

The Dirac equation solutions in the Dirac-Pauli (standard) representation are

$$\left|\psi^{(1)}\right\rangle = \underbrace{\sqrt{\frac{E+m}{2m}}\begin{pmatrix}1\\0\\\dfrac{p^3}{E+m}\\\dfrac{p^1+ip^2}{E+m}\end{pmatrix}}_{\text{spinor } u_1 = \text{ part of solution in 4D spinor space}} \underbrace{e^{-ipx}}_{\substack{\text{4D}\\\text{physical}\\\text{space}\\\text{part}}} = u_1 e^{-ipx} \qquad \left|\psi^{(2)}\right\rangle = \underbrace{\sqrt{\frac{E+m}{2m}}\begin{pmatrix}0\\1\\\dfrac{p^1-ip^2}{E+m}\\\dfrac{-p^3}{E+m}\end{pmatrix}}_{\text{spinor } u_2} e^{-ipx} = u_2 e^{-ipx}$$

$$\left|\psi^{(3)}\right\rangle = \underbrace{\sqrt{\frac{E+m}{2m}}\begin{pmatrix}\dfrac{p^3}{E+m}\\\dfrac{p^1+ip^2}{E+m}\\1\\0\end{pmatrix}}_{\text{spinor } v_2} e^{ipx} = v_2 e^{ipx} \qquad \left|\psi^{(4)}\right\rangle = \underbrace{\sqrt{\frac{E+m}{2m}}\begin{pmatrix}\dfrac{p^1-ip^2}{E+m}\\\dfrac{-p^3}{E+m}\\0\\1\end{pmatrix}}_{\text{spinor } v_1} e^{ipx} = v_1 e^{ipx}\,. \tag{4-20}$$

*Solutions to Dirac equation (discrete, plane waves, **p** eigenstates)*

Yes, again, these are more complicated than solutions we have dealt with in the past, but you will get used to them with time. Note several things in (4-20). Any constant instead of $\sqrt{(E+m)/2m}$ would suffice, but that choice was made because things will work better later on with it, as we will see. The symbol E is always a positive number of magnitude equal to the energy. p^i is positive if it points in the positive direction of its respective axis. These are plane wave solutions. We have defined new symbols $u_r(\mathbf{p})$ and $v_r(\mathbf{p})$ (r =1,2), which are the column vectors multiplied by the constant shown, are functions only of **p** for a given m (since $E = \sqrt{\mathbf{p}^2 + m^2}$), and go by the name spinors, or four-spinors. Note that the particles represented by the $|\psi^{(n)}\rangle$ are also often called spinors. We will show shortly that r values represent different spin states (for example, u_1 represents spin up, and u_2 represents spin down in the particle at-rest system.) As you might expect, we will find the solutions containing $v_r(\mathbf{p})$ are associated with antiparticles; and those with $u_r(\mathbf{p})$, with particles. More on this later, but for now, take care to note the reverse order numbering on $v_{2,1}$ from $u_{1,2}$, which is customary.

Column vector parts of solutions called spinors

The solutions (4-20) are eigenstates of **p**, since every measurement of 3-momentum of the particles they represent would result in the value **p**. They are also eigenstates of energy, since, for given m, a free particle of 3-momentum **p** has a fixed E. [1]

Inner Product of Spinors

We will find the inner product of each spinor in (4-20) with itself a valuable thing to know. For $u_1(\mathbf{p})$,

[1] Similar to that mentioned in a footnote in Chap. 3 for scalars, there are additional solutions to the Dirac equation having exponents of form $\pm i(E_{\mathbf{p}}t + \mathbf{p}\cdot\mathbf{x})$, instead of $\pm i(E_{\mathbf{p}}t - \mathbf{p}\cdot\mathbf{x})$, but these have been widely ignored. The possible impact on QFT of including these solutions in the theory is discussed in R. D. Klauber, "Mechanism for Vanishing Zero-Point Energy", http://arxiv.org/abs/astro-ph/0309679 (2003).

$$u_1^\dagger(\mathbf{p})u_1(\mathbf{p}) = \frac{E+m}{2m}\begin{pmatrix}1 & 0 & \dfrac{p^3}{E+m} & \dfrac{p^1-ip^2}{E+m}\end{pmatrix}\begin{pmatrix}1\\0\\\dfrac{p^3}{E+m}\\\dfrac{p^1+ip^2}{E+m}\end{pmatrix} = \frac{E+m}{2m}\left(1+\frac{\mathbf{p}^2}{(E+m)^2}\right) \tag{4-21}$$

Inner product of spinors with themselves

$$= \frac{E+m}{2m}\left(\frac{(E+m)^2+\mathbf{p}^2}{(E+m)^2}\right) = \frac{E^2\overbrace{+m^2}^{E^2-\mathbf{p}^2}+2Em+\mathbf{p}^2}{2m(E+m)} = \frac{2E^2+2Em}{2m(E+m)} = \frac{E}{m}.$$

By doing Prob. 6, you can feel comfortable with the general result (underline means no summation),

$$u_{\underline{r}}^\dagger(\mathbf{p})u_{\underline{r}}(\mathbf{p}) = v_{\underline{r}}^\dagger(\mathbf{p})v_{\underline{r}}(\mathbf{p}) = \frac{E}{m}, \tag{4-22}$$

Spinor magnitude = $\sqrt{E/m}$

which we got because of our original choice of the constant in (4-20). (You may want to ruminate, by looking at (4-21), on how our spinor inner product would have been a little more complicated if we had chosen unity as our original constant in (4-20). We make our choices for arbitrary constants in order to conform with custom. They are what are commonly used.)

Orthogonality of Spinors

The inner product of u_1 and u_2 is

$$u_1^\dagger(\mathbf{p})u_2(\mathbf{p}) = \frac{E+m}{2m}\begin{pmatrix}1 & 0 & \dfrac{p^3}{E+m} & \dfrac{p^1-ip^2}{E+m}\end{pmatrix}\begin{pmatrix}0\\1\\\dfrac{p^1-ip^2}{E+m}\\\dfrac{-p^3}{E+m}\end{pmatrix} \tag{4-23}$$

Orthogonality of spinors

$$= \frac{E+m}{2m}\left(\frac{p^3(p^1-ip^2)-p^3(p^1-ip^2)}{(E+m)^2}\right) = 0.$$

By doing Prob. 7, you can prove to yourself that

$$u_r^\dagger(\mathbf{p})u_s(\mathbf{p}) = v_r^\dagger(\mathbf{p})v_s(\mathbf{p}) = 0 \qquad r \neq s\,. \tag{4-24}$$

We can combine (4-22), (4-24), and the first part of Prob. 8 into the general result

$$\boxed{u_r^\dagger(\mathbf{p})u_s(\mathbf{p}) = v_r^\dagger(\mathbf{p})v_s(\mathbf{p}) = \frac{E}{m}\delta_{rs} \qquad u_r^\dagger(\mathbf{p})v_s(-\mathbf{p}) = 0.} \tag{4-25}$$

General spinor inner product relations

Orthogonality of Eigensolutions

We can show that the eigensolutions (4-20) to the Dirac equation are orthogonal in the usual quantum mechanical way. For example, for first two solutions of (4-20)

$$\left\langle \psi^{(1)} \middle| \psi^{(2)} \right\rangle = \int \psi_{state}^{(1)\,\dagger}\psi_{state}^{(2)}d^3x = \int u_1^\dagger(\mathbf{p})e^{+ipx}u_2(\mathbf{p})e^{-ipx}d^3x = \underbrace{u_1^\dagger(\mathbf{p})u_2(\mathbf{p})}_{=0}\underbrace{\int e^{+ipx}e^{-ipx}d^3x}_{V}. \tag{4-26}$$

For the first and third solutions,

$$\left\langle \psi^{(1)} \middle| \psi^{(3)} \right\rangle = \int u_1^\dagger(\mathbf{p})e^{+ipx}v_2(\mathbf{p})e^{+ipx}d^3x = \underbrace{u_1^\dagger(\mathbf{p})v_2(\mathbf{p})}_{\neq 0}\underbrace{\int e^{+ipx}e^{+ipx}d^3x}_{=0}. \tag{4-27}$$

By continuing with different solution pairs (do Prob. 9 for practice), one can prove that

$$\left\langle \psi^{(m)} \middle| \psi^{(n)} \right\rangle = 0 \text{ for } m \neq n\,. \tag{4-28}$$

Dirac eigen solutions are orthogonal

Given that we should know mathematically that eigenvector solutions in any eigenvalue problem are generally orthogonal, and that the Dirac equation represents an eigenvalue problem, this should not be too surprising.

General Solution to Dirac Equation

The most general solution to the Dirac equation is a sum (or integral) of all eigenstates, each having a (typically complex) coefficient representing the amount of that eigenstate in the total general state. The discrete plane wave general solution is then

General solution = sum of eigenstate solutions

$$\psi_{state} = |\psi\rangle = \sum_{r,\mathbf{p}} \sqrt{\frac{m}{VE_{\mathbf{p}}}} \left(C_r(\mathbf{p}) u_r(\mathbf{p}) e^{-ipx} + D_r^{\dagger}(\mathbf{p}) v_r(\mathbf{p}) e^{ipx} \right), \tag{4-29}$$

where $C_r(\mathbf{p})$ and $D_r^{\dagger}(\mathbf{p})$ are the coefficients. We will explain later the reason for the unusual normalization constant chosen in terms of mass, energy, and volume, though from (4-22) and the normalization we did in Chap. 3, you may be sensing what is coming.

4.1.6 The Adjoint Dirac Equation

Unlike the Klein-Gordon equation, the Dirac equation is a matrix equation. So, rather than a complex conjugate form of the wave equation, we need to consider taking a complex conjugate transpose of that equation, and of its solutions (4-20). But, as it turns out, the theory works more coherently if rather than $|\psi\rangle^{\dagger} = \langle\psi| = \psi^{\dagger}_{state}$, we define and use the adjoint

Definition of adjoint

$$\bar{\psi}_{state} = \psi^{\dagger}_{state}\gamma^0 = |\psi\rangle^{\dagger}\gamma^0 = \langle\psi|\gamma^0 = \langle\bar{\psi}|, \tag{4-30}$$

where an inner product between the row vector $|\psi\rangle^{\dagger} = \langle\psi| = \psi^{\dagger}_{state}$ and the gamma matrix are implied, and we hope the symbolism has not become unwieldy. You should do Prob. 10 to show yourself what the four row vectors $\langle\bar{\psi}^{(n)}|$ look like.

The adjoint Dirac equation is obtained by taking the complex conjugate transpose of the Dirac equation (4-15), post multiplying by γ^0, and using the Hermiticity conditions along with the adjoint definition. (Do this in Prob. 11.) The result is

The adjoint Dirac equation

$$i\partial_{\mu}\langle\bar{\psi}|\gamma^{\mu} + m\langle\bar{\psi}| = 0 \quad . \tag{4-31}$$

Simply by deriving (4-31), we have proven that the adjoint (4-30) solves the adjoint Dirac equation, as long as $|\psi\rangle$ solves the Dirac equation. By doing Prob. 12, you can justify it to yourself in a more "hands on" way.

Adjoint spinors are defined as the row vectors

Adjoint spinor definitions

$$\bar{u}_r = u_r^{\dagger}\gamma^0 \qquad \bar{v}_r = v_r^{\dagger}\gamma^0, \tag{4-32}$$

which, with (4-29) and (4-30), gives us the discrete plane wave adjoint general solution form

General adjoint solution

$$\bar{\psi}_{state} = \langle\bar{\psi}| = \sum_{r,\mathbf{p}} \sqrt{\frac{m}{VE_{\mathbf{p}}}} \left(D_r(\mathbf{p}) \bar{v}_r(\mathbf{p}) e^{-ipx} + C_r^{\dagger}(\mathbf{p}) \bar{u}_r(\mathbf{p}) e^{ipx} \right). \tag{4-33}$$

4.1.7 Probability Density for Dirac Fermions in RQM

Probability and the Four-Current Using the Dirac Equation

Recall from Chap. 3 (see Box 3-1, pg. 45, and Sect. 3.1.4, pg. 44) that, in any theory, we can typically use the governing equation (and often its complex conjugate transpose) to find a conserved quantity, which in quantum theory can be total probability. We did this in NRQM and for scalars in RQM. So, we try the same general approach for spin ½ particles, but note that researchers found early on that the adjoint Dirac equation was better for this purpose than simply the complex conjugate of the Dirac equation.

Dirac equation and its adjoint yield a 4-divergence of a 4-current = 0

This is actually much simpler to do for the first order Dirac equation and its adjoint than it was for the second order Klein-Gordon equation, so we leave it as Prob. 13 for the reader. The result of that problem is

$$\partial_\mu j^\mu = 0 \qquad j^\mu = (\rho, \mathbf{j}) = \bar{\psi}_{state}\,\gamma^\mu \psi_{state} = \langle \bar{\psi} | \gamma^\mu | \psi \rangle_{\text{not integ}}, \tag{4-34}$$

Form of the 4-current

where the subscript "not integ" means we are not integrating over space in the bracket shown, contrary to what the bracket symbol typically implies in the position basis. (4-34) means the total quantity

$$\begin{aligned} \int_V j^0 d^3x &= \int_V \rho d^3x = \int_V \bar{\psi}_{state}\,\gamma^0 \psi_{state} d^3x = \langle \bar{\psi} | \gamma^0 | \psi \rangle \\ &= \int_V \psi^\dagger_{state} \underbrace{\gamma^0 \gamma^0}_{\mathrm{I}} \psi_{state} d^3x = \langle \psi | \psi \rangle = Q' \end{aligned} \tag{4-35}$$

The conserved quantity related to the 4-current

is conserved for V = all space, and ρ is the density value corresponding to Q'. Naturally, we would like ρ to be a probability density and Q' to be total probability, as we had in NRQM.

Probability for the Dirac Discrete Solutions

For a single particle state in RQM, we assume at first that the solution (4-29) has only terms with coefficients C_r (i.e., only has spinors of form u_r), i.e., the general state $|\psi\rangle$ contains no eigenstates with coefficients D_r (i.e., no spinors of form v_r). Our ρ of (4-34) is then

Is that conserved quantity probability?

$$\begin{aligned} \rho &= \left(\sum_{r,\mathbf{p}} \sqrt{\frac{m}{VE_\mathbf{p}}} C_r^\dagger(\mathbf{p}) \underbrace{\bar{u}_r(\mathbf{p})}_{u_r^\dagger \gamma^0(\mathbf{p})} e^{ipx} \right) \gamma^0 \left(\sum_{r',\mathbf{p}'} \sqrt{\frac{m}{VE_{\mathbf{p}'}}} C_{r'}(\mathbf{p}') u_{r'}(\mathbf{p}') e^{-ip'x} \right) \\ &= \left(\sum_{r,\mathbf{p}} \sqrt{\frac{m}{VE_\mathbf{p}}} C_r^\dagger(\mathbf{p}) u_r^\dagger(\mathbf{p}) e^{ipx} \right) \left(\sum_{r',\mathbf{p}'} \sqrt{\frac{m}{VE_{\mathbf{p}'}}} C_{r'}(\mathbf{p}') u_{r'}(\mathbf{p}') e^{-ip'x} \right). \end{aligned} \tag{4-36}$$

If (4-36) is probability density, its integral over all space must equal 1. In such integration all terms with $\mathbf{p}' \neq \mathbf{p}$ go to zero due to the exponential term, and all remaining terms with $r' \neq r$ go to zero from the spinor orthogonality relation (4-24). We end up with

$$\int \rho d^3x = \sum_{r,\mathbf{p}} \frac{m}{VE_\mathbf{p}} \left(C_r^\dagger(\mathbf{p}) C_r(\mathbf{p}) \underbrace{u_r^\dagger(\mathbf{p}) u_r(\mathbf{p})}_{E_\mathbf{p}/m} \underbrace{\int e^{-ipx} e^{ipx} d^3x}_{V} \right) = \sum_{r,\mathbf{p}} C_r^\dagger(\mathbf{p}) C_r(\mathbf{p}) = \sum_{r,\mathbf{p}} |C_r(\mathbf{p})|^2 = 1. \tag{4-37}$$

For C_r type particles, the conserved quantity can be interpreted as probability

Thus, by taking ρ as probability density, $|C_r(\mathbf{p})|^2$ is the probability of measuring the particle with spin state r and 3-momentum eigenstate $\mathbf{p}$, similar to what the coefficients of eigenstates represented in NRQM, and for scalars in RQM (see Chap. 3, Sect. 3.1.4, pgs. 44-47).

Normalization Factors

(4-37) is the reason we used $\sqrt{m/VE_\mathbf{p}}$ as our normalization factor in solutions (4-20). A different such factor would not have resulted in unity on the RHS of (4-37).

We chose constant in solutions so conserved quantity would equal 1

4.1.8 Negative Energies for D Type Dirac Fermions

Do Prob. 14 to prove to yourself that a Dirac fermion represented by a solution with exponential form $-ipx$ (that has spinor u_r and coefficient $C_r(\mathbf{p})$), has positive energy; and one represented by the solution form ipx (spinor v_r and coefficient $D_r(\mathbf{p})$), has negative energy.

D_r type particles have negative energy

Thus, Dirac did not solve one of the problems he originally set out to solve. Half of his solutions represent particles with negative energies, just like the Klein-Gordon solutions did.

4.1.9 Probability for D Type Dirac Fermions

Repeating the process of (4-36) for a D type particle instead of a C type, we find

$$\int \rho d^3x = \sum_{r,\mathbf{p}} \frac{m}{VE_\mathbf{p}} \left(D_r(\mathbf{p}) D_r^\dagger(\mathbf{p}) \underbrace{v_r^\dagger(\mathbf{p}) v_r(\mathbf{p})}_{E_\mathbf{p}/m} \underbrace{\int e^{-ipx} e^{ipx} d^3x}_{V} \right) = \sum_{r,\mathbf{p}} D_r(\mathbf{p}) D_r^\dagger(\mathbf{p}) = \sum_{r,\mathbf{p}} |D_r(\mathbf{p})|^2 = 1 \tag{4-38}$$

Probability for D_r type particle

and a total positive probability (unlike the negative probability in the scalar case) that we can set equal to 1.

4.1.10 A Great Use for the Dirac Approach: Spin

Dirac approach of great value for spin ½

So though Dirac's approach still resulted in negative energies, he did solve two of the problems he set out to solve (a first order equation and no negative probabilities.) Well beyond that, however, he provided physics with something of enormous value. His equation, and its solutions, allow us to model relativistic particles with spin ½, which could not be done before. Importantly, this turned out to give correct relativistic solutions for the hydrogen atom, which the Klein-Gordon equation did not.

RQM Spin Operator

Before delving into relativistic spin, you should read over and understand Box 4-1, Review of Spin in NRQM. That box is summarized in the second column of Wholeness Chart 4-1, pg 101.

In RQM, we have recently seen that each of the solutions Dirac came up with has a column matrix (spinor) built in. But it is a four, not two, component column matrix. It wasn't long before people realized that Dirac's four solutions (4-20) represented two kinds of particles, regular particles and antiparticles. The four components of each solution were different for each, and represented the four states of spin up and spin down for particles, and spin up and spin down for antiparticles.

The spin operator can be derived formally from tensor analysis of the energy-momentum and angular momentum tensors in 4D spacetime, but that is pretty complicated.[2] We will simply state the RQM spin operator Σ_i as follows, then compare it to the NRQM version of the spin operator, and work some examples, to justify the form we have assumed.

RQM spin operator

$$\Sigma_i = \frac{\hbar}{2}\begin{bmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{bmatrix} \rightarrow \Sigma_1 = \frac{\hbar}{2}\begin{bmatrix} & 1 & & \\ 1 & & & \\ & & & 1 \\ & & 1 & \end{bmatrix} \Sigma_2 = \frac{\hbar}{2}\begin{bmatrix} & -i & & \\ i & & & \\ & & & -i \\ & & i & \end{bmatrix} \Sigma_3 = \frac{\hbar}{2}\begin{bmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix}. \quad (4\text{-}39)$$

Note (4-39) is a 3D object in physical space (3 components *i* of the angular momentum), but each of the components in that space is itself a 4X4 matrix in relativistic 4D spinor space, rather than non-relativistic 2D spinor space of NRQM. In relativity, spin has three spatial components, as it did non-relativistically. But, relativistically, each component must act on a 4D column vector in spinor space. So, if we were to guess at a 4D matrix spin operator, then (4-39), where we formed 4D matrices from our 2D Pauli matrices, would be a good first guess. Compare (4-39) to (B4-1.1) of Box 4-1. And, as we will see, this guess turns out to be correct.

As in NRQM, we can choose the direction of our *z* axis however we like, and it was easier in NRQM if we lined it up in the direction of spin of our particle, so spin would be either up (plus *z* direction), or down (minus *z* direction). We'll do something similar in RQM, so we'll focus for the present on the *z* direction spin component operator Σ_3.

Dirac Fermion Spins: Stationary Particle Examples

Consider a Dirac particle which, for simplicity, is simply sitting in front of us and not moving. So **p**=0 in our frame, and the solutions (4-20) become much simplified. What then, are their respective spins? For the first such solution, we have

First Dirac solution, if particle stationary, is spin up eigenstate of RQM spin operator

$$\Sigma_3 \left|\psi^{(1)}\right\rangle = \frac{\hbar}{2}\begin{bmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix} \underbrace{\sqrt{\frac{m+m}{2m}}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{u_1} e^{-ipx} = \frac{\hbar}{2}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} e^{-ipx} = \frac{\hbar}{2}\left|\psi^{(1)}\right\rangle \quad \hbar = 1 \text{ in n.u.} \qquad (4\text{-}40)$$

and the spin is + $\hbar$/2, indicating that our first Dirac solution is for spin up.

---

[2] See F. Mandl and G. Shaw, *Quantum Field Theory*, 1st ed. (John Wiley 1984), Chap 2, pgs. 38-39 and Chap 4, pg. 65.

Box 4-1. Review of Spin in Non Relativistic Quantum Mechanics

NRQM Spin Operator

In NRQM our spin operator S_i was the Pauli matrices σ_i times the factor $\hbar/2$,

$$S_i = \frac{\hbar}{2}\sigma_i \;\rightarrow\; S_1 = \frac{\hbar}{2}\sigma_1 = \frac{\hbar}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad S_2 = \frac{\hbar}{2}\sigma_2 = \frac{\hbar}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad S_3 = \frac{\hbar}{2}\sigma_3 = \frac{\hbar}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{B4-1.1}$$

Spin (non-orbital angular momentum) S_i in physical space acted like a vector with three components, as it does classically. A particle, or object, can have spin components in any of the three dimensions. In spinor space, however, spin was represented by a 2X2 matrix, one for each 3D component.

Eigenstates of the Spin Operator Components

z direction

Thus, a wave function (ket) with a column matrix representing spin up in the *z* direction $(1,0)^{\mathrm{T}}$ had spin $\hbar/2$; and one in the down direction $(0,1)^{\mathrm{T}}$ had spin $-\hbar/2$.

$$\begin{aligned} S_3\left|\psi_{\mathbf{p},up}\right\rangle &= S_3\left(A(\mathbf{p})e^{-ipx}\begin{bmatrix}1\\0\end{bmatrix}\right) = A(\mathbf{p})e^{-ipx}\frac{\hbar}{2}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = A(\mathbf{p})e^{-ipx}\frac{\hbar}{2}\begin{bmatrix}1\\0\end{bmatrix} = \frac{\hbar}{2}\left|\psi_{\mathbf{p},up}\right\rangle \\ S_3\left|\psi_{\mathbf{p},down}\right\rangle &= S_3\left(A'(\mathbf{p})e^{-ipx}\begin{bmatrix}0\\1\end{bmatrix}\right) = -\frac{\hbar}{2}\left|\psi_{\mathbf{p},down}\right\rangle . \end{aligned} \tag{B4-1.2}$$

The spin up and spin down states are eigenstates of S_3, the *z* direction component of the spin operator S_i.

x direction

$(1,1)^{\mathrm{T}}$ and $(1,-1)^{\mathrm{T}}$ are readily shown to be eigenstates of the *x* component of the spin operator, i.e., S_1, via

$$\begin{aligned} S_1\left|\psi_{\mathbf{p},+x\,\mathrm{spin}}\right\rangle &= S_1\left(B(\mathbf{p})e^{-ipx}\begin{bmatrix}1\\1\end{bmatrix}\right) = B(\mathbf{p})e^{-ipx}\frac{\hbar}{2}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = B(\mathbf{p})e^{-ipx}\frac{\hbar}{2}\begin{bmatrix}1\\1\end{bmatrix} = \frac{\hbar}{2}\left|\psi_{\mathbf{p},+x\,\mathrm{spin}}\right\rangle \\ S_1\left|\psi_{\mathbf{p},-x\,\mathrm{spin}}\right\rangle &= B'(\mathbf{p})e^{-ipx}\frac{\hbar}{2}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\begin{bmatrix}1\\-1\end{bmatrix} = B'(\mathbf{p})e^{-ipx}\frac{\hbar}{2}\begin{bmatrix}-1\\1\end{bmatrix} = -\frac{\hbar}{2}\left|\psi_{\mathbf{p},-x\,\mathrm{spin}}\right\rangle . \end{aligned} \tag{B4-1.3}$$

Note that we can construct the *x* component eigenstates from linear combinations of the *z* component eigenstates, i.e.

$$\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix} \qquad \begin{bmatrix}1\\-1\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix} - \begin{bmatrix}0\\1\end{bmatrix} . \tag{B4-1.4}$$

y direction

Do Prob. 15 to prove to yourself that $(i,-1)^{\mathrm{T}}$ and $(i,1)^{\mathrm{T}}$ are eigenstates of the *y* component of the spin operator, i.e., S_2, and that they can be constructed from linear combinations of the up and down (*z* direction) eigenstates.

Up and Down (z Direction) Eigenstates Span the 2D Spinor Space (and are Basis Vectors of that Space)

Any spin state (not necessarily pointed in one of the *x*,*y*, or *z* directions) as visualized in 3D physical space is composed of components in the *x*,*y*, and *z* directions of the 3D spin vector. This is a simple law of vector components.

Each of these three components for the *x*,*y*,*z* directions is represented in 2D spinor space as a different column matrix. But, importantly, all possible such column matrices can be expressed as linear combinations of the up and down (*z* direction) eigenstate column matrices. We showed this for the *x* direction in (B4-1.4). You showed it for *y*, if you did the problem suggested above. (We didn't normalize (B4-1.4). The normalized forms are $\left(1/\sqrt{2}\right)(1,1)^{\mathrm{T}}$ and $\left(1/\sqrt{2}\right)(1,-1)^{\mathrm{T}}$.)

Since the *x* and *y* direction spin states can be constructed via linear combination of the *z* direction eigenspin states, all possible spin states can be composed of various combinations of the *z* direction eigenspin states, i.e., of the up and down states $(1,0)^{\mathrm{T}}$ and $(0,1)^{\mathrm{T}}$. In mathematical language, we say $(1,0)^{\mathrm{T}}$ and $(0,1)^{\mathrm{T}}$ are basis vectors that span spinor space.

General Solution Includes All Possible Spin States

$$\left|\psi\right\rangle = \sum_{\mathbf{p}}\left(C_+(\mathbf{p})e^{-ipx}\begin{bmatrix}1\\0\end{bmatrix} + C_-(\mathbf{p})e^{-ipx}\begin{bmatrix}0\\1\end{bmatrix}\right) \tag{B4-1.5}$$

The general solution (B4-1.5) includes all possible spin states. For example, for given $\mathbf{p}$, with $C_-(\mathbf{p}) = -C_+(\mathbf{p})$, spin is in the negative *x* direction (see (B4-1.3) above. Choosing coefficients correctly yields any given particle spin direction.

Bottom Line: If we develop our theory for up and down states, it will be applicable to all possible other spin states, too.

Now do Prob. 16 to prove to yourself that the other three solutions, for a stationary particle (or antiparticle) represent, respectively, spin down ($-\hbar/2$), spin up, and spin down states.

Other Dirac solutions, if particle stationary, are one spin up and two spin down states

We included the symbol $\hbar$ in our spin discussion so far, so that you could see the role it played in non-natural unit formulations, but from now on, we will assume natural units, such that $\hbar = 1$. And we will refer to spin magnitudes as ½ , which is more common than $\hbar/2$.

Dirac Fermion Spins: Moving Particles Discussion

Due to relativity, a moving particle's spin is more complicated

As soon as we have a moving particle, things get more complicated, as we have to include non-zero p^i values in our solutions (4-20). This complication, which we didn't have in NRQM, is due to relativistic effects.

Classical Relativistic Effects on a Spinning Object

Classically, relativistic translation alters **L**

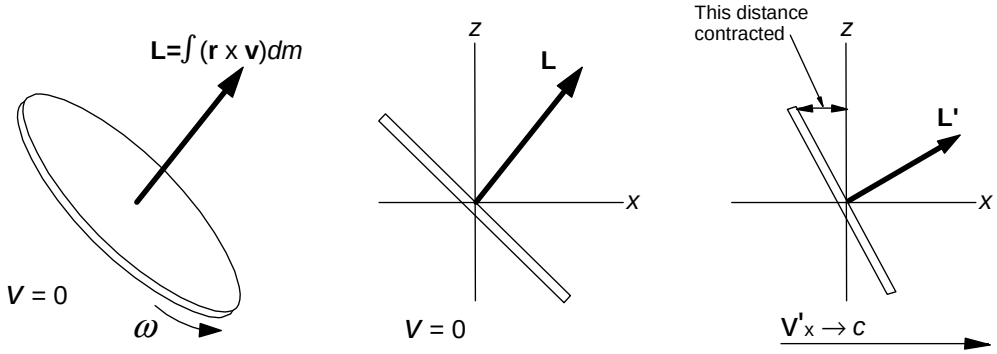
Read Box 4-2 to gain some appreciation for the effect on a classical, macroscopic object's angular momentum direction when viewed from a frame in which the object has translational speed approaching that of light.

Box 4-2. Classical Macroscopic Spinning Object Translating at Relativistic Speed

In 4D relativistic theory, angular momentum is a 2nd order tensor, but it can be treated simply as a vector formed from the integral over a rotating body of $dm(\mathbf{r} \times \mathbf{v}_t) = dm(\mathbf{r} \times (\boldsymbol{\omega} \times \mathbf{r}))$, where symbols should be obvious. When a macroscopic object like a spinning disk, as shown below, moves close to the speed of light, distances contract in the direction of the velocity, and this makes the plane of the disk appear to turn. (See figure below.)

The closer the disk gets to the speed of light, the more the disk surface appears in the observer's frame to align normal to the velocity direction. In the rest frame translating with the disk itself, the disk still appears aligned in the original way.

In the observer's frame, though, the angular momentum **L** appears to turn toward the direction of the velocity. The greater the speed, the greater this turning. At light speed, **L** and **v** become parallel.



Spinning disk, no translation | Same disk, side view | Same disk, high velocity

For **L** in the opposite direction and the same positive *x* direction velocity, **L** would point more and more towards the *negative x* direction, as speed increased.

Quantum Mechanics Experiments: Spin Magnitude Unchanged at Any Speed

Experimentally, quantum particle spin magnitude always the same

We know from all the testing on high speed quantum particles like electrons, that spin (particle angular momentum) does not change in magnitude at all, no matter what speed we see the particle carrying it as having. Electrons always have spin of $\hbar/2$ (½ in natural units), no matter what their velocity.

Quantum mechanically, then, at high speed, a particle's angular momentum (spin) magnitude remains unchanged, but its direction appears to us in our frame to realign itself closer to that of the

translational velocity vector. See Fig. 4-1. As velocity approaches *c*, the angular momentum (spin) approaches the velocity direction (or directly opposite direction.) For massless particles like photons, which travel at *c*, the spin is always aligned parallel with the velocity vector, either in the direction of velocity or in the opposite direction. Such a state is called a pure helicity state.

As $v \rightarrow c$, quantum particle spin axis re-aligns closer to the **v** *direction. For $v = c$, the particle is in a pure helicity state*

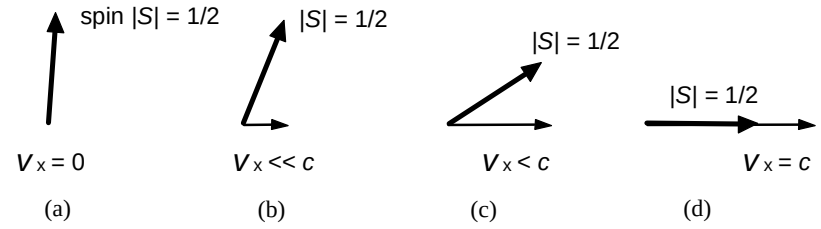


Figure 4-1. Effect of Transverse Velocity on Dirac Particle Spin

Math form of Dirac spinor changes with **p** *as physical spin direction re-aligns*

Mathematically, these kinds of relativistic complications are incorporated into the form of the spinors $u_r(\mathbf{p})$ and $v_r(\mathbf{p})$ (by their dependence on 3-momentum and thus ultimately, on velocity, as we will see below) and by how they are combined to form more general spin states. But it is nonetheless good to have some physical appreciation for why spin angular momentum direction can, for relativistic cases, depend upon linear momentum **p**.

Spinor form actually dependent on just velocity, not 3-momentum

Spinor Components Really Dependent on Velocity

With reference to the prior paragraph, note that the spinor components are actually dependent on particle velocity, rather than momentum, by the following logic.

Energy and momentum are expressed (in non-natural units to make it easier to understand)

$$E = \frac{mc^2}{\sqrt{1 - v^2/c^2}} \qquad p^i = \frac{mv^i}{\sqrt{1 - v^2/c^2}}, \tag{4-41}$$

so in the coefficient and spinor components of the Dirac spinor (4-20) the mass *m* drops out. This leaves them a function solely of velocity (and thus in line with our knowledge, and with Box 4-2, that relativistic effects are dependent on velocity.)

Dirac Spinors: What Happens When $\mathbf{v} \neq 0$

Note what happens to the spin as seen by us, for an electron whose spin is represented solely by u_1, but has $p^1 \neq 0$, with $p^2 = p^3 = 0$ in our frame (the lab.)

$$\Sigma_3 \left|\psi^{(1)}\right\rangle = \frac{1}{2}\begin{bmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix} \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \dfrac{p^1}{E+m} \end{pmatrix} e^{-ipx} = \frac{1}{2}\sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ \dfrac{-p^1}{E+m} \end{pmatrix} e^{-ipx} \neq \frac{1}{2}\left|\psi^{(1)}\right\rangle. \tag{4-42}$$

u_1 = spin up eigenstate if particle at-rest. u_1 not generally a spin up eigenstate for moving particle

So such a moving electron is not in a spin up eigenstate as seen in the lab. This relates to our discussion above about translation affecting the direction of the spin vector axis. Here, the mathematics bears this out. (We will have more to say on this shortly.) u_1 for a non-translating electron has spin up, but u_1 for an electron with high transverse velocity is not an up eigenstate.

Now consider u_1 representing an electron traveling in the *z* direction instead of the *x* direction.

But if particle motion is in direction of at-rest spin axis, it is also a spin eigenstate when moving

$$\Sigma_3 \left|\psi^{(1)}\right\rangle = \frac{1}{2}\begin{bmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix} \sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 \\ 0 \\ \dfrac{p^3}{E+m} \\ 0 \end{pmatrix} e^{-ipx} = \frac{1}{2}\sqrt{\frac{E+m}{2m}} \begin{pmatrix} 1 \\ 0 \\ \dfrac{p^3}{E+m} \\ 0 \end{pmatrix} e^{-ipx} = \frac{1}{2}\left|\psi^{(1)}\right\rangle. \tag{4-43}$$

This electron, represented by u_1, is an up eigenstate as it moves, just as it was when it was at rest. Relativistically, this makes sense, as the plane of a spinning disk with **L** aligned in the direction of **p** would not appear to turn as **p** increased from zero to a relativistic value.

In general, boosts in the spin axis direction leave u_1, u_2, v_2 and v_1 in the same spin eigenstates as they would be at rest. Boosts in other directions take them out of those spin eigenstates. We will look a bit at the math behind this, for general particle states, shortly, but doing Probs. 17 and 18 now will enable you to get some feeling for how this happens with two particular Dirac particle states.

The At-Rest Coordinate System

The at-rest coord system is the lab frame with particle decelerated to **v** = 0 *and z axis parallel to spin*

The four solutions (4-20) (for given **p**) and their associated spinors $u_r(\mathbf{p})$ and $v_r(\mathbf{p})$ can be thought of as up/down spin eigenstates if the respective particles represented by each were decelerated to be at rest in the lab. We will call the coordinate system, fixed in the lab, in which the particle would be decelerated to be at rest, and its spin would be in the *z* direction, the at-rest coordinate system. (See Appendix B, Sect. **Error! Reference source not found.**, pg. **Error! Bookmark not defined.** discussion on this definition.) Note this is a lab frame and is *different from the rest frame coordinate system*, which is fixed to the particle.

Dirac Fermion Spins: How to Think About Them

Think of r index on spinors as representing z spin eigenstate for at-rest coord system

$u_1(\mathbf{p})$ and $v_2(\mathbf{p})$ in this system would then represent spin up stationary states; $u_2(\mathbf{p})$ and $v_1(\mathbf{p})$ represent spin down stationary states. We will virtually never work with electrons that are not moving, so our spinors will always look more complicated than the one in (4-40). But it can help us mentally as we work with such states, that are often not in a spin eigenstate in the lab, to think of them as spin eigenstates for the at-rest system, i.e., for a particle decelerated to **v** = 0 in the lab.

The Four-Spinors Span the 4D Spinor Space

In NRQM 2D spinor space, as we showed in Box 4-1, the two eigenspinors $(1,0)^T$ and $(0,1)^T$ (spin up and spin down, respectively) spanned the space of all possible spins (all possible spins can be constructed from linear combinations of those two eigenspin states, so they are basis vectors for 2D spinor space.) The general solution state (B4-1.5), a linear superposition of spin up and spin down eigenstates, contains all possible spin direction states. Each possible distinct general particle state has different coefficients C_+ and C_- for the eigenspin solutions summed to form that general state. This results in a different spin direction for the particle for each different set of coefficients. Thus, in NRQM, all possible spin directions can be represented by general solution (B4-1.5).

By analogy, we can surmise that the four Dirac spinors u_1, u_2, v_2 and v_1 of (4-20) span the RQM 4D spinor space of all possible spins and momenta, and thus, are basis vectors for that space. Our RQM general solution (4-29) contains within it all possible relativistic spin states.

The four Dirac spinors span the 4D spinor space of RQM

More mathematically, we should know that a 4D space is spanned by four column vectors, each of four components, where these vectors are all independent of one another. Generally, the vector solutions of an eigenvalue problem, which is what the Dirac equation solutions are, are independent and complete, and thus we can conclude, span the space. They can be used as basis vectors.

Further, the stationary particle case has spinors of form (ignoring the normalization factor in front) $(1,0,0,0)^T$, $(0,1,0,0)^T$, $(0,0,1,0)^T$, and $(0,0,0,1)^T$. These are obviously independent (and orthogonal). That independence is not changed by a boost (giving the particle a velocity relative to our frame.)

General RQM Solution Contains All Possible Spin Directions

RQM spinor general solution encompasses all possible spin directions

Hence, similar to NRQM, our RQM general solution (4-29) for spinors contains all possible spin states within it. Different coefficients $C_1(\mathbf{p})$ and $C_2(\mathbf{p})$ will yield different spin states for *C* type particles. And different coefficients $D^\dagger_1(\mathbf{p})$ and $D^\dagger_2(\mathbf{p})$ will yield different spin states for *D* type particles.

To see how this works, look at Fig. 4-1, and consider how each of the four states shown therein can be represented by their respective terms in the general particle state solution (4-29). In general, for *j* = *a,b,c,d*, the four states shown (for a *C* type particle) in Fig. 4-1 are

An example

$$\left|\psi_{(j)}\right\rangle = \sqrt{\frac{m}{VE_{\mathbf{p}_j}}}\left(C_1(\mathbf{p}_j)u_1(\mathbf{p}_j) + C_2(\mathbf{p}_j)u_2(\mathbf{p}_j)\right)e^{-ip_j x}\,. \qquad (4\text{-}44)$$

State (a) there is effectively spin up with $\mathbf{p}_a = 0$, where from (4-20),

$$\left|\psi_{(a)}\right\rangle = \sqrt{\frac{m}{VE_{\mathbf{p}a}}}C_1(0)u_1(0)e^{-ip_a x} = \sqrt{\frac{m}{VE_{\mathbf{p}a}}}\underbrace{\sqrt{\frac{E_{\mathbf{p}_a}+m}{2m}}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}}_{u_1(0)}e^{-ip_a x}, \tag{4-45}$$

which from (4-40) is an eigenstate of Σ_3. So, for state (a), (4-44) has effectively, $C_1 = 1$ and $C_2 = 0$.

For the last state (*d*), where the particle is traveling at the speed of light, (4-44) becomes an eigenstate of Σ_1 (see Prob. 18) with eigenvalue ½. That is,

$$\begin{aligned}\left|\psi_{(d)}\right\rangle &= \sqrt{\frac{m}{VE_{\mathbf{p}d}}}C_1(\infty)u_1(\infty)e^{-ip_d x} + \sqrt{\frac{m}{VE_{\mathbf{p}d}}}C_2(\infty)u_2(\infty)e^{-ip_d x}\\ &= \sqrt{\frac{m}{VE_{\mathbf{p}d}}}\left(\underbrace{\sqrt{\frac{E_{\mathbf{p}_d}+m}{2m}}\begin{pmatrix}1\\0\\0\\1\end{pmatrix}}_{u_1(\infty)} + \underbrace{\sqrt{\frac{E_{\mathbf{p}_d}+m}{2m}}\begin{pmatrix}0\\1\\1\\0\end{pmatrix}}_{u_2(\infty)}\right)e^{-ip_d x} = \sqrt{\frac{m}{VE_{\mathbf{p}d}}}\underbrace{\sqrt{\frac{E_{\mathbf{p}_d}+m}{2m}}\begin{pmatrix}1\\1\\1\\1\end{pmatrix}}_{\text{eigenstate of }\Sigma_1}e^{-ip_d x},\end{aligned} \tag{4-46}$$

where here, we must have $C_1 = 1$ and $C_2 = 1$ (in the normalized version, $C_1 = C_2 = 1/\sqrt{2}$) to get the proper eigenstate on the RHS of the lower row in (4-46).

For "in between" states (*b*) and (*c*), C_1 and C_2 would have other values. We won't get into how those values are determined here. We only want to make the following point.

__p__ determines $u_{1,2}$ and then spin is represented by correct linear combination of u_1 and u_2

The $u_{1,2}$ values are determined in a given problem solely by **p** (or equivalently, **v**.) A particle has given **p**, the $u_{1,2}$ are determined, and they serve as our spinor space basis vectors. Then, for a given spin alignment in our physical world problem, we have to choose the correct values for $C_{1,2}$ to mathematically represent that spin state, for the given **p**. The spin basis vectors $u_{1,2}$ span spinor space, but we have to determine how much of each we need for their linear superposition to equal the spin state we are dealing with. Parallel logic holds for *D* type particles and $v_{2,1}$.

So we see that although our development of the theory focused on only the two states of spin up and spin down (for stationary particle = at-rest frame particle) states, the theory is applicable to all possible spin direction states.

<u>One State That is Impossible</u>

Note that we can never have a relativistic state where the spin vector and **p** are at right angles.

<u>Dirac Spinors Become NRQM Spinors in Low Speed Limit</u>

RQM spinors → NRQM spinors, as $v \to 0$.

Classical relativistic mechanics approaches Newtonian mechanics in the limit where speed is much less than that of light. In parallel fashion, we should expect that our RQM solutions approach NRQM solutions as $v \to 0$, as well. By doing Prob. 19, you can gain an understanding of how this does indeed happen with Dirac spinors. If it didn't, our solutions could not be correct, so this is one more check on the theory we have developed.

<u>Spinors vs Particle Spin Dependence on Velocity</u>

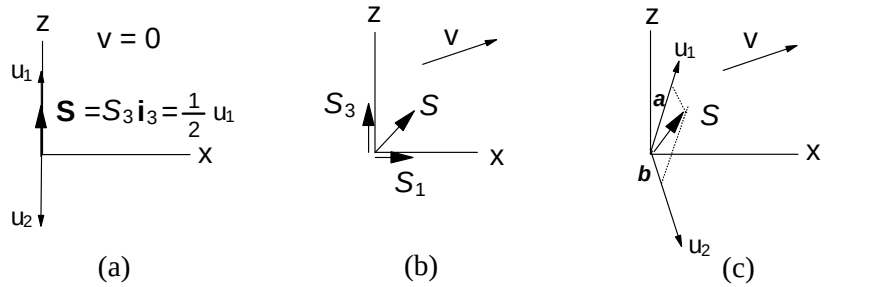
As all of this relativistic spin stuff is a bit mind bending, the following heuristic description of it is offered to help in conceptualizing what is actually going on physically.

Fig. 4-2 illustrates spin for a Dirac fermion in the at-rest system (particle at-rest in the lab) and a boosted particle system in terms of its *x* and *z* components in the lab, and also in terms of its spinor u_1 and u_2 components.

Note that u_1 and u_2 actually exist in spinor space (they are spinor space basis vectors in that space), but they correspond to directions in physical space. For example, in the at-rest system, u_1 represents spin up and so can be visualized as a spatial vector that points in the +*z* direction. Similarly, in the at-rest system, u_2 represents spin down, so can be visualized as a vector pointing in the –*z* direction. (See Fig 4-2a.)

Representing the u_r as vectors is a heuristic oversimplification though, and in fact is not really correct, as operations like spinor addition work a little differently than vector addition. (See Winter[3].) However, temporarily visualizing them as such can aid in our understanding of how they and spin behave, relative to the at-rest coordinate system, for varying particle velocities.

u_1 and u_2 can be visualized heuristically as directions in physical space



(a) Particle at rest system = lab system

(b) Boosted particle, spin x and z components in lab

(c) Boosted particle, spin u_1 and u_2 components in lab

Figure 4-2. Heuristic Look at Spin in the At-Rest System and in a Particle Boosted System

In the particle boosted system (Fig 4-2b and c), u_1 and u_2, as we learned, no longer represent spin in the + and – z directions, but can be visualized as vectors pointing in other directions, where such directions depend on the particular boost velocity **v**.

So both the spin direction of **S** and the spin space basis vectors u_1 and u_2, as visualized in physical space, change direction with boost velocity **v**. (Compare Fig. 4-2a to Fig. 4-2c).

At-rest system

In the at-rest system, spin is aligned with the z axis, and in our example it is the positive direction of the z axis. Spin is **S** (= $\mathbf{S}_3$) = ½ $\mathbf{u}_1$ and

$$\mathbf{S} = \mathbf{S}_3 = S_3\,\mathbf{i}_3 \qquad \text{or} \qquad \mathbf{S} = \tfrac{1}{2}\mathbf{u}_1 . \tag{4-47}$$

Particle boosted system

In the boosted system, we can express the spin (no longer aligned with the z axis of the lab) as

*Dependence of u_1 and u_2 directions on **v** is different from spin direction dependence*

$$\mathbf{S} = S_1\mathbf{i}_1 + S_3\mathbf{i}_3 \qquad \text{or} \qquad \mathbf{S} = a\mathbf{u}_1 + b\mathbf{u}_2 , \tag{4-48}$$

(where u_1 and u_2 are no longer aligned with the z axis either).

Note that when the particle is boosted, the directions of u_1 and u_2 change, and so does the direction of the spin **S**. However, they change a little differently. **S** for the at-rest particle is solely composed of u_1, whereas for the boosted particle it has components of both u_1 and u_2. A common new student mistake is to think that **S** and u_1 change in the same manner with **v**, and that the RHS of (4-47), if it holds at-rest, will hold for any velocity.

Again, the vector addition implied in Fig. 4-2c and the RHS of (4-48) is not quite correct for spinors, but hopefully, the underlying concept that the amount of each of u_1 and u_2 in **S** changes with velocity has become clear.

Summarizing Eigen Status of u_1 and u_2

Summarizing the essence of u_1 and u_2

$u_1(\mathbf{p})\,e^{-ipx}$ and $u_2(\mathbf{p})\,e^{-ipx}$ is each always an eigenstate of the Dirac equation (for any **p**).

$u_1(\mathbf{p})\,e^{-ipx}$ and $u_2(\mathbf{p})\,e^{-ipx}$ is each sometimes an eigenstate of z spin, i.e. of Σ_3 (for $\mathbf{p} = 0$ or $= p^3\,\mathbf{i}_3$).

$u_1(\mathbf{p})\,e^{-ipx}$ and $u_2(\mathbf{p})\,e^{-ipx}$ are always basis vectors for any general state $|\psi\rangle$ (for any **p**).

$u_1(\mathbf{p})$ and $u_2(\mathbf{p})$ is each sometimes an eigenstate of z spin, i.e. of Σ_3 (for $\mathbf{p} = 0$ or $= p^3\,\mathbf{i}_3$).

$u_1(\mathbf{p})$ and $u_2(\mathbf{p})$ are always basis vectors in 4D spinor space (for any **p**).

$u_1(\mathbf{p})$ and $u_2(\mathbf{p})$ change orientation, as visualized in physical space, as **p** changes.

Spin **S** (often written in relativity as **Σ)** changes direction with **p**, but differently than u_1 and u_2.

[3] Winter, Rolf G., *Quantum Physics,* Wadsworth (1979), Chap. 9.

Any general spin state u can be represented as a linear combination of u_1 and u_2 (for any **p**).

$$u(\mathbf{p}) = C_1(\mathbf{p})\, u_1(\mathbf{p}) + C_2(\mathbf{p})\, u_2(\mathbf{p})$$

Any general particle state includes a spin part plus a spacetime part (for any **p**).

$$|\psi\rangle = u(\mathbf{p})\, e^{-ipx} = C_1(\mathbf{p})\, u_1(\mathbf{p})\, e^{-ipx} + C_2(\mathbf{p})\, u_2(\mathbf{p})\, e^{-ipx}$$

4.1.11 RQM Helicity Operator

SEE TEXT FOR THIS SECTION

4.1.12 Summary of RQM for Dirac Particles

Do Prob. 21, which asks you to construct your own Wholeness Chart summary for RQM Dirac particles for all but spin and helicity.

Spin and helicity are summarized herein in Wholeness Chart 4-1 for both NRQM and RQM.

Wholeness Chart 4-1. Spin ½ Particle Spin Summary

	NRQM	**RQM**
Physical space	3D space plus time	4D spacetime
Spinor space dimensions	2D	4D
Wave function = ket (plane waves here)	physical space plus time part e^{-ikx}, and spinor space part (2D spinor)	physical spacetime parts e^{-ikx} and e^{+ikx}, and spinor space part (4D spinor)
Spinor forms = basis vectors (Dirac-Pauli representation for RQM)	$\chi_+ = \begin{bmatrix}1\\0\end{bmatrix}$ $\chi_- = \begin{bmatrix}0\\1\end{bmatrix}$	$u_1 = \sqrt{\frac{E+m}{2m}}\begin{pmatrix}1\\0\\\frac{p^3}{E+m}\\\frac{p^1+ip^2}{E+m}\end{pmatrix}$ $u_2 = \sqrt{\frac{E+m}{2m}}\begin{pmatrix}0\\1\\\frac{p^1-ip^2}{E+m}\\\frac{-p^3}{E+m}\end{pmatrix}$ similar forms for v_2 and v_1, here and below
Spinor form dependence	Independent of velocity **v**	Depends on velocity **v** (on **p**)
Spin operator	$S_i = \frac{\hbar}{2}\sigma_i$ σ_i = 2X2 Pauli matrices	$\Sigma_i = \frac{\hbar}{2}\begin{bmatrix}\sigma_i & 0\\0 & \sigma_i\end{bmatrix}$ 4X4 matrices
Spin operator character	3D physical space angular momentum vector similar to linear momentum vector **p**. Both are quantum operators. Spin vector components = matrix operators in spinor space.	
Spinors spin eigenstate?	Yes. Always	Generally no. Yes if **v** = 0, or spin & x^3 aligned
Spin in directions of 3D axes (Linear combinations not normalized here)	For any velocity $\chi_+ \rightarrow +z\text{spin};\ \chi_- \rightarrow -z\text{spin}$ $\chi_+ + \chi_- \rightarrow +x;\ \chi_+ - \chi_- \rightarrow -x$ $i\chi_+ - \chi_- \rightarrow +y;\ i\chi_+ + \chi_- \rightarrow -y$	For **v** = 0, or spin and **v** both in direction indicated $u_1 \rightarrow +z\text{spin};\ \ u_2 \rightarrow -z\text{spin}$ $u_1 + u_2 \rightarrow +x;\ \ u_1 - u_2 \rightarrow -x$ $iu_1 - u_2 \rightarrow +y;\ iu_1 + u_2 \rightarrow -y$
Spin operator eigenvals	$\pm\hbar/2$ + (–) for given σ_i if spin in pos (neg) *i* axis direction	$\pm\hbar/2$ (½ in natural units) + (–) for given Σ_i if spin in pos (neg) *i* axis direction
General spin direction (Can normalize coefficients if desired.)	$\chi = C_+\chi_+ + C_-\chi_-$ C_+, C_- determine spin direction. Choose them to fit problem spin, including spins not along *x,y,z* axes	$u = C_1(\mathbf{p})u_1(\mathbf{p}) + C_2(\mathbf{p})u_2(\mathbf{p})$ C_1, C_2 determine spin direction. Choose them to fit particular problem spin for any given **p** (including spins not along **p** or *x,y,z* axes.)
Helicity	$\sigma_{\mathbf{p}} = \boldsymbol{\sigma}\cdot\frac{\mathbf{p}}{\lvert\mathbf{p}\rvert} = \sigma_i\frac{p^i}{p}$	$\Sigma_{\mathbf{p}} = \boldsymbol{\Sigma}\cdot\frac{\mathbf{p}}{\lvert\mathbf{p}\rvert} = \Sigma_i\frac{p^i}{p}$
Spinors helicity eigenstates?	Generally no. Yes if spin and **p** aligned for χ_+ or χ_-	Generally no. Yes, if spin and **p** aligned for u_1 or u_2
Helicity eigenstates	Any general state where spin parallel to line of action of **p**	
Helicity eigenvalues	$\pm\hbar/2$ (½ in natural units) + if spin in **p** direction; – if in – **p** direction	
General states summary	A state is in a spin eigenstate of S_i (Σ_i) if spin is aligned in + or – direction of *i*th axis. A state is in a helicity eigenstate if spin is aligned with **p** or – **p**. Eigenstate or not, coefficients on the $\chi_{+,-}$ (or $u_{1,2}$) can be chosen to produce the mathematical state for the particular spin direction (and **p**), and thus for the helicity, as well.	

4.2 The Dirac Equation in Quantum Field Theory

4.2.1 Summary Chart

Be sure to use the summary wholeness chart, as you study this chapter

All that we will do in the remainder of this chapter is summarized in Wholeness Chart 5-4 at the end of Chap. 5. I highly recommend following along, step by step, in that chart as we progress from here to the end of this chapter.

4.2.2 From RQM to QFT

No classical theory for spin ½ fields, so we can't use $\mathcal{H}$ or $\mathcal{L}$ for 2nd quantization postulate #1

Recall from Sect. 4.1.1, pg. 85, that because fermions cannot occupy the same state, they cannot reinforce one another to produce a macroscopic fermionic field. So, we have no classical spinor field theory. That is, we have no Hamiltonian density, Lagrangian density, nor wave equation (all of which are essentially equivalent), which we could use for quantizing and deriving a quantum spinor field theory.

But in our scalar theory we found the QFT wave equation for fields (in the Heisenberg picture) was identical in form to the RQM equation for states (in the Schrödinger picture). It is not a great leap of faith, therefore, to assume the same thing will hold true for spin ½ fermions. And, of course, as history and experiment have proven, it does.

So, we take a hint from scalar theory and take the Dirac RQM equation as our QFT field equation

Thus, the Dirac equation for fields (where we will, as with scalar fields, work in the Heisenberg picture), from (4-15) for states, is

$$\left(i\gamma^\mu\partial_\mu - m\right)\psi = 0 \ . \tag{4-49}$$

Its eigensolutions (for given **p**), identical in form to (4-20) but fields now instead of states, are

$$\psi^{(1)} = u_1 e^{-ipx} \quad \psi^{(2)} = u_2 e^{-ipx} \quad \psi^{(3)} = v_2 e^{ipx} \quad \psi^{(4)} = v_1 e^{ipx}, \tag{4-50}$$

whose mathematical behavior we have already learned quite a bit about.

And the RQM adjoint equation as our QFT adjoint equation

The adjoint Dirac equation for fields, from (4-31), is

$$i\partial_\mu\bar{\psi}\gamma^\mu + m\bar{\psi} = 0\,, \tag{4-51}$$

with adjoint eigensolutions

$$\bar{\psi} = \psi^\dagger\gamma^0 \ \rightarrow\ \bar{\psi}^{(1)} = u_1^\dagger\gamma^0 e^{ipx} = \bar{u}_1 e^{ipx} \quad \bar{\psi}^{(2)} = \bar{u}_2 e^{ipx} \quad \bar{\psi}^{(3)} = \bar{v}_2 e^{-ipx} \quad \bar{\psi}^{(4)} = \bar{v}_1 e^{-ipx}\,. \tag{4-52}$$

4.2.3 Summary of General Plane Wave Solutions

For the Dirac and adjoint Dirac equation, the general discrete plane wave solutions are

$$\boxed{\begin{aligned}\psi &= \sum_{r,\mathbf{p}}\sqrt{\frac{m}{VE_\mathbf{p}}}\left(c_r(\mathbf{p})u_r(\mathbf{p})e^{-ipx} + d_r^\dagger(\mathbf{p})v_r(\mathbf{p})e^{ipx}\right)\\ &= \qquad\qquad \psi^+ \qquad + \qquad \psi^-\\ \bar{\psi} &= \sum_{r,\mathbf{p}}\sqrt{\frac{m}{VE_\mathbf{p}}}\left(d_r(\mathbf{p})\bar{v}_r(\mathbf{p})e^{-ipx} + c_r^\dagger(\mathbf{p})\bar{u}_r(\mathbf{p})e^{ipx}\right)\\ &= \qquad\qquad \bar{\psi}^+ \qquad + \qquad \bar{\psi}^-,\end{aligned}} \tag{4-53}$$

General plane wave solutions to Dirac field equation

and the general continuous plane wave solutions are

$$\begin{aligned}\psi &= \sum_r\sqrt{\frac{m}{(2\pi)^3}}\int\frac{d^3\mathbf{p}}{\sqrt{E_\mathbf{p}}}\left(c_r(\mathbf{p})u_r(\mathbf{p})e^{-ipx} + d_r^\dagger(\mathbf{p})v_r(\mathbf{p})e^{ipx}\right)\\ \bar{\psi} &= \sum_r\sqrt{\frac{m}{(2\pi)^3}}\int\frac{d^3\mathbf{p}}{\sqrt{E_\mathbf{p}}}\left(d_r(\mathbf{p})\bar{v}_r(\mathbf{p})e^{-ipx} + c_r^\dagger(\mathbf{p})\bar{u}_r(\mathbf{p})e^{ipx}\right).\end{aligned} \tag{4-54}$$

Note that for fields we use lower case for the coefficients, as we did for fields in the scalar treatment. You may be anticipating that we want to distinguish them from the RQM coefficients (upper case) because we will find the QFT coefficients to be operators, rather than mere numbers. If so, you will turn out to be right, as we will see.

4.2.4 The Dirac Lagrangian, Conjugate Momentum, and Hamiltonian

From the Dirac equation (4-49), and trial and error, we can deduce the Lagrangian (density) for free spinor fields to be

Dirac Lagrangian (density) $\mathcal{L}$ for free fields, deduced from Dirac equation

$$\mathcal{L}_0^{1/2} = \bar{\psi}\left(i\gamma^\alpha \partial_\alpha - m\right)\psi \ , \tag{4-55}$$

which can be checked by plugging into the Euler-Lagrange equation,

$$\frac{d}{dx^\mu}\left(\frac{\partial \mathcal{L}}{\partial \phi^n{}_{,\mu}}\right) - \frac{\partial \mathcal{L}}{\partial \phi^n} = 0\,, \quad \text{with} \qquad \phi^1 = \bar{\psi};\ \ \phi^2 = \psi;\ \ \mathcal{L} = \mathcal{L}_0^{1/2}\,. \tag{4-56}$$

By doing Prob.22, you can prove to yourself that for n = 1 above, we get the Dirac equation, and for n = 2, we get the adjoint Dirac equation.

Conjugate momenta for ψ and $\bar{\psi}$ are

Dirac conjugate momenta

$$\pi^{1/2} = \frac{\partial \mathcal{L}_0^{1/2}}{\partial \psi_{,0}} = i\bar{\psi}\gamma^0 = i\psi^\dagger\gamma^0\gamma^0 = i\psi^\dagger \qquad \bar{\pi}^{1/2} = \frac{\partial \mathcal{L}_0^{1/2}}{\partial \bar{\psi}_{,0}} = 0\,, \tag{4-57}$$

where the adjoint momentum on the RHS might be a little surprising, but is true.

The Dirac Hamiltonian density can be found from the Legendre transformation as

Dirac Hamiltonian density $\mathcal{H}$ from $\mathcal{L}$

$$\begin{aligned}\mathcal{H}_0^{1/2} &= \pi^{1/2}\dot{\psi} + \bar{\pi}^{1/2}\dot{\bar{\psi}} - \mathcal{L}_0^{1/2} = i\psi^\dagger\dot{\psi} - \mathcal{L}_0^{1/2} = i\underbrace{\psi^\dagger\gamma^0}_{\bar{\psi}}\gamma^0\dot{\psi} - \mathcal{L}_0^{1/2} \\ &= i\bar{\psi}\gamma^0\dot{\psi} - i\bar{\psi}\gamma^0\dot{\psi} - i\bar{\psi}\gamma^i\partial_i\psi + m\bar{\psi}\psi = -i\bar{\psi}\gamma^i\partial_i\psi + m\bar{\psi}\psi\ .\end{aligned} \tag{4-58}$$

4.3 Anti-commutation Relations for Dirac Fields

4.3.1 No Spinor Poisson Brackets: Try Assuming Coefficient Commutators

No classical spinor field Poisson brackets → guess that Dirac coeffs have commutation relations like scalars

Since we don't have macroscopic spinor fields, and thus no associated Poisson brackets, we can't really carry out second quantization postulate #2, in which we took Poisson brackets over into commutators. But just as we took a hint in Sect. 4.2.2 from the scalar fields case for postulate #1 (i.e., we assumed that the QFT wave equation was the same as the RQM wave equation in both scalar and spin ½ cases), we can take a similar hint for postulate #2. That is, we can postulate that the Dirac solution coefficients $c_r(\mathbf{p})$ and $d_r(\mathbf{p})$ obey the same sort of commutation relations that the Klein-Gordon solution coefficients $a(\mathbf{k})$ and $b(\mathbf{k})$ did.

That guess is wrong

But, when early researchers did this, they soon found that such *commutation relations did not work for Dirac fields*. They did not produce a viable theory that matched the real world. (We will see this near the end of the chapter, but for now, just accept it.)

4.3.2 Dirac Coefficient Anti-commutation Relations Do Work

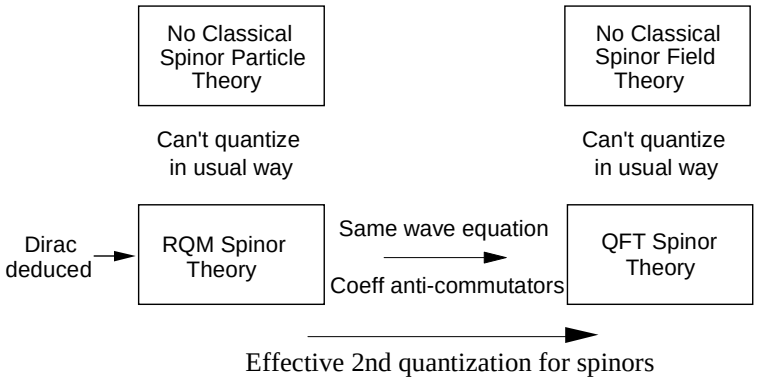
A new guess: Dirac coeffs obey anti-commutation relations.

However, it was soon found that coefficient anti-commutation relations, parallel in form to the scalar coefficient commutation relations, did work. These are

$$\boxed{\begin{gathered}\left[c_r(\mathbf{p}), c_s^\dagger(\mathbf{p}')\right]_+ = \left[d_r(\mathbf{p}), d_s^\dagger(\mathbf{p}')\right]_+ = \delta_{rs}\delta_{\mathbf{pp}'} \text{ (discrete)};\ = \delta_{rs}\delta(\mathbf{p}-\mathbf{p}') \text{ (continuous)}. \\ \text{All other anti-commutators between coefficients equal zero.}\end{gathered}} \tag{4-59}$$

This guess is our postulate #2 for 2nd quantization of spinor fields

We will thus use (4-59) as our postulate #2 of 2nd quantization for Dirac fields. See Fig. 4-3. And before too long, we will prove to ourselves that indeed, this postulate does give us a viable spinor field theory.



The right guess and the same Dirac wave equation yields QFT for spinors

Figure 4-3. The Route to QFT for Spin ½ Fields

4.4 The Dirac Hamiltonian in QFT

SEE TEXT FOR REST OF THIS CHAPTER

4.5 Problems

1. Show that *i*) (4-7) solve the relations (4-6), and *ii*) they also fulfill the requirements of being traceless, hermitian (matrix equals its complex conjugate transpose), and β and α_1 have ± 1 eigenvalues. Note that eigenvectors for β are any constant times any of $(1,0,0,0)^T$, $(0,1,0,0)^T$, $(0,0,1,0)^T$, and $(0,0,0,1)^T$. Eigenvectors for α_1 are any constant times any of $(1,0,0,1)^T$, $(0,1,1,0)^T$, $(1,0,0,-1)^T$, and $(0,1,-1,0)^T$. Do you believe α_2 and α_3 would also fulfill all these relations and requirements, if we took the trouble to examine them?

2. Show that (4-9) equals (4-10) and (4-11).

3. Prove the Hermiticity conditions $\gamma^{\mu\dagger} = \gamma^0 \gamma^\mu \gamma^0$ by substitution, plugging, and chugging.

4. Use the form (4-14) of the Dirac equation (reproduced below), insert the Dirac gamma matrices (4-10), and show each term in the summation over μ. Write the whole thing as a matrix equation (which should be a sum of matrices where that sum is post-multiplied by the column matrix $|\psi\rangle$). Note that a different derivative is taken on the whole vector (column matrix) $|\psi\rangle$ for each Dirac gamma matrix. Note also that the resulting matrix equation has four rows, and thus, may be considered as four separate scalar equations, one for each index κ. Each such equation has all four components of $|\psi\rangle$ in it.

$$\sum_{\eta=1}^{4} \left(\sum_{\mu=0}^{3} i \left(\gamma^\mu\right)_{\kappa\eta} \partial_\mu - m\delta_{\kappa\eta} \right) |\psi\rangle_\eta = 0 \qquad \kappa = 1,2,3,4 \qquad (4\text{-}14)$$

5. Show via substitution that the solutions (4-20) satisfy the Dirac equation (4-15).

6. Find the inner products of at least two of $u_2^\dagger(\mathbf{p})u_2(\mathbf{p})$, $v_2^\dagger(\mathbf{p})v_2(\mathbf{p})$, and $v_1^\dagger(\mathbf{p})v_1(\mathbf{p})$.

7. Find the inner product of $v_1^\dagger(\mathbf{p})v_2(\mathbf{p})$.

8. Show that the inner product of $u_1^\dagger(\mathbf{p})v_2(-\mathbf{p}) = 0$ and that $u_1^\dagger(\mathbf{p})u_2(-\mathbf{p}) = 0$.

9. Show that for the solutions (4-20), $\langle \psi^{(1)} | \psi^{(4)} \rangle = 0$ and $\langle \psi^{(3)} | \psi^{(4)} \rangle = 0$

10. Use (4-10) and (4-20) to show yourself what the four adjoints $\langle \bar{\psi}^{(n)} |$ look like.

11. Derive the adjoint Dirac equation (4-31).

12. Pick one of the adjoints $\langle \bar{\psi}^{(n)} |$ you found in Prob. 10, plug it into the adjoint Dirac equation, and chug the math to show it solves that equation.

13. Derive the 4-current related to the conserved quantity associated with Dirac particles that might be interpreted as probability. (Hint: Pre-multiply the Dirac equation by $\bar{\psi}$, post-multiply the adjoint Dirac equation by ψ, add the two, then group the resulting terms such that the four derivative ∂_μ is taken with respect to a quantity in brackets. That quantity will be the four-current with zero four-divergence in spacetime.) You should get the four current j^μ of the RHS of (4-34) multiplied by the constant i, but the constant is irrelevant in $\partial_\mu j^\mu = 0$, so, we drop it.

14. Operate on the four solutions (4-20) with the Hamiltonian $H = i\partial/\partial t$ to show that those with coefficient $D_r(\mathbf{p})$ have negative energies and those with $C_r(\mathbf{p})$ have positive energies.

15. Show that the NRQM states $(i,-1)^{\mathrm{T}} e^{-ipx}$ and $(i,1)^{\mathrm{T}} e^{-ipx}$ are eigenstates of the y component of the NRQM spin operator, i.e., S_2, and that they can be constructed from linear combinations of the NRQM up and down (z direction) eigenstates.

16. Take $\mathbf{p} = 0$ and operate on the last three of the solutions (4-20) with the relativistic spin operator Σ_3 to show that they are spin eigenstates of Σ_3 when they represent particles that are not moving.

17. Pick one or two, other than the first, of the solutions (4-20) and show that for $p^1 \neq 0$, $p^2 = p^3 = 0$, a particle, which would be in a z direction (up) spin eigenstate if $\mathbf{p}$ were zero, will not be in that up spin eigenstate. Then show, for the same solutions chosen, but with $p^1 = p^2 = 0$, $p^3 \neq 0$ as seen in the lab, that the particle is in the same spin eigenstate as it would be if $\mathbf{p}$ were zero.

18. Show that as a Dirac particle of spinor form $u = u_1 + u_2$ approaches the speed of light in the x direction (i.e., $p^1 \to \infty$, $p^2 << p^1$, and $p^3 << p^1$), then u approaches a spin state in the plus x direction (i.e., the spinor $u = u_1 + u_2$ is an eigenvector of Σ_1 of (4-39) with eigenvalue + ½ (in natural units.) Then show that for $p^2 \approx p^1$ and $p^3 \approx p^1$, this is not true.

19. Take the relativistic solutions $|\psi^{(1)}\rangle$ and $|\psi^{(2)}\rangle$ of (4-20) for the Dirac equation in the non-relativistic limit, i.e., with $v<<1$ ($c=1$ in natural units), and show that they reduce to approximately $(1,0,0,0)^{\mathrm{T}} e^{-ipx}$ and $(0,1,0,0)^{\mathrm{T}} e^{-ipx}$. Do these look remarkably like the wave functions of a spin ½ particle you encountered in NRQM, i.e., spin up state $e^{-ipx}(1,0)^{\mathrm{T}}$ and spin down state $e^{-ipx}(0,1)^{\mathrm{T}}$?

    Operate on each of the two low speed 4-spinor states you found with the 4D spin operator Σ_3 of (4-39). Now, operate on each of the two NRQM 2-spinor states above for up and down spin with $\frac{1}{2}\hbar\sigma_3$. Do you get the same spin eigenstates and eigenvalues? Does the low speed RQM case for C type (i.e., having spinors u_r) particles parallel that of NRQM?

20. Take a Dirac particle in state $|\psi^{(2)}\rangle$ with $p^1 \neq 0$, $p^2 = p^3 = 0$. Without doing any math, is this in a helicity eigenstate or not? Operate on the state with the helicity operator and prove your answer mathematically. Then repeat the exercise for the same state with $p^3 \neq 0$, $p^1 = p^2 = 0$.

21. Construct your own Wholeness Chart summarizing Sect. 4.1, RQM for Spinors. You do not need to summarize spin and helicity, as that is done for you in Wholeness Chart 4-1.

22. Using (4-55), show that for $n = 1$ in (4-56), we get the Dirac equation; and for $n = 2$, we get the adjoint Dirac equation.

Chapter 5 *Print vers: 3/6/13 copyright of Robert D. Klauber*

Vectors: Spin 1 Fields

"Three passions, simple but overwhelmingly strong, have governed my life: the longing for love, the search for knowledge, and unbearable pity for the suffering of mankind. These passions, like great winds, have blown me hither and thither, in a wayward course, over a great ocean …
… I have wished to understand the hearts of men. I have wished to know why the stars shine. And I have tried to apprehend the Pythagorean power by which number holds sway above the flux. A little of this, but not much, I have achieved."

Excerpts from "What I Have Lived For"
by Bertrand Russell

5.0 Preliminaries

Few of you who have come this far in this book do not share, in some part, Russell's passion for knowledge. It is my hope that, as each of us lives his or her life, we can also share in his other two, most noteworthy, passions.

With regard to the "power by which number holds sway above the flux", few have more ably demonstrated that power in their work than James Clerk Maxwell. His casting of the phenomena of electricity and magnetism into one elegant and holistic mathematical structure will forever remain one of the monumental achievements in the history of mankind.

Although his famous Maxwell equations were formulated for a classical world, they play an equally fundamental role quantum mechanically, as we shall soon see.

5.0.1 Background

Maxwell's equations in 4D format are the basis for QFT of photons

Maxwell first published his equations in 1864, well before the advent of special relativity, so they were framed in a distinctly 3D spatial plus time format. And that is how virtually every physics student first learns them. However, as QFT is a distinctly relativistic theory, we will need to work with Maxwell's equations in the more appropriate 4D format. This treatment of electromagnetism is typically reserved for graduate courses, after students have gained some level of comfort with the 3 + 1 dimensional approach. We will review the 4D approach, but hopefully, it is something that readers of this book have already been exposed to, as it serves as the bedrock for QFT of photons (massless spin 1 bosons).

5.0.2 Chapter Overview

Spin 1 boson theory development parallels spin 0 boson theory

Our approach to spin 1 bosons (called vectors, for reasons that will become apparent) in this chapter is three fold, including i) a review of classical electromagnetic theory, ii) RQM for photons, and iii) QFT for photons. As we will see, the second and third of these bear striking parallel to comparable aspects of spin 0 boson theory, and this will help to make our work easier.

Vector bosons, like scalars, can be massive or massless, but since our focus in this book is on quantum electrodynamics (QED), where force is mediated by photons, we will be virtually exclusively concerned with photons, which are massless vector bosons.

The following bulleted points provide an overview of this chapter. You may find it helpful to compare and contrast the material below for RQM and QFT with that of the Chapter Overview for scalars at the beginning of Chap. 3.

A review of classical e/m first,

where we will look at

Classical e/m overview

- the pre-relativistic version of Maxwell's equations, their (3D) vector and scalar potentials, and how they describe classical electromagnetic fields/waves,
- those same equations and e/m fields/waves represented covariantly (i.e., in special relativity) as a single equation for a 4D potential A^μ, and
- the classical relativistic Lagrangian density $\mathcal{L}$ for classical e/m fields.

Then RQM for photons (of which e/m waves are made),

RQM preview (*photons = massless vectors*)

- deducing the quantum Maxwell equation in terms of the 4D potential A^μ by applying 1st quantization to classical theory,
- solutions $|A^\mu\rangle$ to that equation (the 4D potential will represent a photon mathematically), and
- noting that those solutions parallel the Klein-Gordon solutions, and so, much of what we learned for scalars can be carried over directly to vectors with $|\phi\rangle \rightarrow |A^\mu\rangle$.

Then QFT for photons,

QFT preview (*photons = massless vectors*)

- from 2nd quantization, finding the same Maxwell equation, with the same mathematical form for the solutions A^μ, but this time the solutions are quantum fields, not states,
- using the classical relativistic $\mathcal{L}$ for e/m fields and the Legendre transformation to get $\mathcal{H}$ (Hamiltonian density),
- from 2nd quantization, finding the photon field A^μ commutation relations for QFT,
- determining relevant QFT operators for photons by a short cut method: comparing to similar operators for scalars: H, number, creation/destruction, momentum, charge, etc., and
- finding the Feynman propagator for photons by analogy with the scalar propagator.

As in Chaps. 3 and 4, in this chapter, we will deal only with free particles/fields.

Only free photons in this chapter

SEE THE TEXT FOR THE REMAINDER OF THIS CHAPTER

NOTE THAT WHOLENESS CHART SUMMARIZING CHAPS. 3 TO 5 IS INCLUDED BELOW

Summary Chart

Chaps. 3, 4, and 5 are summarized in Wholeness Chart 5-4 that follows.

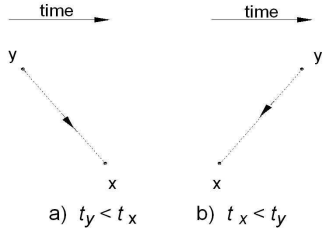
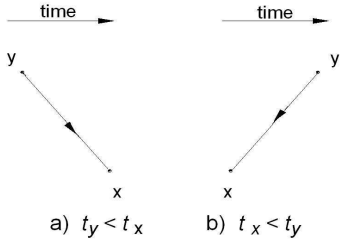
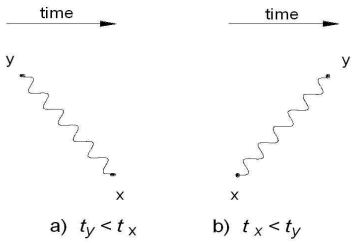
QED/FIELD THEORY OVERVIEW: PART 1

Wholeness Chart 5-4. From Field Equations to Propagators and Observables
Heisenberg Picture, Free Fields

	Spin 0	**Spin ½**	**Spin 1**
Classical Lagrangian density, free	$\mathcal{L}_0^0 = K\left(\partial_\alpha\phi\partial^\alpha\phi - \mu^2\phi\phi\right)$	None. Macroscopic spinor fields not observed.	$\mathcal{L}_0^1 = \underbrace{\frac{\mu^2}{2}A^\mu A_\mu}_{\mu=0 \text{ for photons}} - \frac{1}{2}\left(\partial_\nu A_\mu\right)\left(\partial^\nu A^\mu\right)$
2nd quantization, Postulate #1	Bosons: Quantum field $\mathcal{L}$ (or equivalently, $\mathcal{H}$) same as classical, fields are complex, and K =1. Spinors: Dirac eq from RQM with states → fields. Deduce $\mathcal{L}$ from Dirac eq; $\mathcal{H}$ from Legendre transf.		
QFT Lagrangian density, free	$\mathcal{L}_0^0 = \left(\partial_\alpha\phi^\dagger\partial^\alpha\phi - \mu^2\phi^\dagger\phi\right)$	$\mathcal{L}_0^{1/2} = \bar{\psi}\left(i\partial\!\!\!/ - m\right)\psi \qquad \partial\!\!\!/ = \gamma^\alpha\partial_\alpha$	As above for classical.
	$\mathcal{L}$ ↑ into the Euler-Lagrange equation yields ↓		
Free field eqs	$\left(\partial_\alpha\partial^\alpha + \mu^2\right)\phi = 0$ $\left(\partial_\alpha\partial^\alpha + \mu^2\right)\phi^\dagger = 0$	$(i\gamma^\alpha\partial_\alpha - m)\psi = 0$ $(i\partial_\alpha\bar{\psi}\gamma^\alpha + m\bar{\psi}) = 0 \qquad \bar{\psi} = \psi^\dagger\gamma^0$	$\left(\partial_\alpha\partial^\alpha + \mu^2\right)A^\mu = 0$ photon $\mu = 0$ $A^{\mu\dagger} = A^\mu$ for chargeless (photon)
Conjugate momenta	$\pi_0^0 = \frac{\partial\mathcal{L}_0^0}{\partial\dot{\phi}} = \dot{\phi}^\dagger;\ \pi_0^{0\dagger} = \frac{\partial\mathcal{L}_0^0}{\partial\dot{\phi}^\dagger} = \dot{\phi}$	$\pi^{1/2} = i\psi^\dagger;\ \ \bar{\pi}^{1/2} = 0$	$\pi_\mu^1 = -\dot{A}_\mu$
Hamiltonian density	$\mathcal{H}_0^0 = \pi_0^0\dot{\phi} + \pi_0^{0\dagger}\dot{\phi}^\dagger - \mathcal{L}_0^0$ $= \left(\dot{\phi}\dot{\phi}^\dagger + \nabla\phi^\dagger\cdot\nabla\phi + \mu^2\phi^\dagger\phi\right)$	$\mathcal{H}_0^{1/2} = \pi^{1/2}\dot{\psi} - \mathcal{L}_0^{1/2}$	$\mathcal{H}_0^1 = \pi_\mu^1\dot{A}^\mu - \mathcal{L}_0^1$
Free field solutions	$\phi = \phi^+ + \phi^-$ $\phi^\dagger = \phi^{\dagger+} + \phi^{\dagger-}$	$\psi = \psi^+ + \psi^-$ $\bar{\psi} = \bar{\psi}^+ + \bar{\psi}^-$	$A^\mu = A^{\mu+} + A^{\mu-}$ (photon)
Discrete eigenstates (Plane waves, constrained to volume V)	$\phi(x) = \sum_{\mathbf{k}}\frac{1}{\sqrt{2V\omega_{\mathbf{k}}}}\left(a(\mathbf{k})e^{-ikx} + b^\dagger(\mathbf{k})e^{ikx}\right)$ $\phi^\dagger(x) = \sum_{\mathbf{k}}\frac{1}{\sqrt{2V\omega_{\mathbf{k}}}}\left(b(\mathbf{k})e^{-ikx} + a^\dagger(\mathbf{k})e^{ikx}\right)$	$\psi = \sum_{r,\mathbf{p}}\sqrt{\frac{m}{VE_{\mathbf{p}}}}\left(c_r(\mathbf{p})u_r(\mathbf{p})e^{-ipx} + d_r^\dagger(\mathbf{p})v_r(\mathbf{p})e^{ipx}\right)$ $\bar{\psi} = \sum_{r,\mathbf{p}}\sqrt{\frac{m}{VE_{\mathbf{p}}}}\left(d_r(\mathbf{p})\bar{v}_r(\mathbf{p})e^{-ipx} + c_r^\dagger(\mathbf{p})\bar{u}_r(\mathbf{p})e^{ipx}\right)$	$A^\mu = \sum_{r,\mathbf{k}}\frac{1}{\sqrt{2V\omega_{\mathbf{k}}}}\left(\varepsilon_r^\mu(\mathbf{k})a_r(\mathbf{k})e^{-ikx} + \varepsilon_r^\mu(\mathbf{k})a_r^\dagger(\mathbf{k})e^{ikx}\right)$
Continuous eigenstates (Plane waves, no volume constraint)	$\phi(x) = \int\frac{d\boldsymbol{k}}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}}}}\left(a(\mathbf{k})e^{-ikx} + b^\dagger(\mathbf{k})e^{ikx}\right)$ $\phi^\dagger(x) = \int\frac{d\boldsymbol{k}}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}}}}\left(b(\mathbf{k})e^{-ikx} + a^\dagger(\mathbf{k})e^{ikx}\right)$	$\psi = \sum_r\sqrt{\frac{m}{(2\pi)^3}}\int\frac{d^3\mathbf{p}}{\sqrt{E_{\mathbf{p}}}}\left(c_r(\mathbf{p})u_r(\mathbf{p})e^{-ipx} + d_r^\dagger(\mathbf{p})v_r(\mathbf{p})e^{ipx}\right)$ $\bar{\psi} = \sum_r\sqrt{\frac{m}{(2\pi)^3}}\int\frac{d^3\mathbf{p}}{\sqrt{E_{\mathbf{p}}}}\left(d_r(\mathbf{p})\bar{v}_r(\mathbf{p})e^{-ipx} + c_r^\dagger(\mathbf{p})\bar{u}_r(\mathbf{p})e^{ipx}\right)$ spinor indices on u_r, v_r, and ψ suppressed. r = 1,2.	$A^\mu = \sum_r\frac{1}{\sqrt{2(2\pi)^3}}\int\frac{d\mathbf{k}}{\sqrt{\omega_{\mathbf{k}}}}\left(\varepsilon_r^\mu(\mathbf{k})a_r(\mathbf{k})e^{-ikx} + \varepsilon_r^\mu(\mathbf{k})a_r^\dagger(\mathbf{k})e^{ikx}\right)$ $r = 0,1,2,3$ (4 polarization vectors)

2nd quantization Postulate #2	Bosons: $\left[\phi^r(\mathbf{x},t),\pi_s(\mathbf{y},t)\right]=\left[\phi^r\pi_s-\pi_s\phi^r\right]=i\delta^r{}_s\delta(\mathbf{x}-\mathbf{y})$, ϕ^r = any field, other commutators = 0. Spinors: Coefficient anti-commutation relations parallel coefficient commutation relations for bosons.		
	Bosons: using conjugate momenta expressions in ↑ yields ↓		
Equal time commutators (intermediate step only)	$\left[\phi(\mathbf{x},t),\dot{\phi}^\dagger(\mathbf{y},t)\right]=i\delta(\mathbf{x}-\mathbf{y})$	Not needed for spinor derivation.	$\left[A^\mu(\mathbf{x},t),\dot{A}^\nu(\mathbf{y},t)\right]=-ig^{\mu\nu}\delta(\mathbf{x}-\mathbf{y})$
	Bosons: Using free field solutions in ↑ with 3D Dirac delta function (e.g., for discrete solutions, $\delta(\mathbf{x}-\mathbf{y})=\frac{1}{V}\sum_{n=-\infty}^{+\infty}e^{-i\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{y})}$), and matching terms, yields the coefficient commutators ↓.		
Coefficient commutators	$\left[a(\mathbf{k}),a^\dagger(\mathbf{k}')\right]=\left[b(\mathbf{k}),b^\dagger(\mathbf{k}')\right]$	$\left[c_r(\mathbf{p}),c_s^\dagger(\mathbf{p}')\right]_+=\left[d_r(\mathbf{p}),d_s^\dagger(\mathbf{p}')\right]_+$	$\left[a_r(\mathbf{k}),a_s{}^\dagger(\mathbf{k}')\right]$
discrete	$=\delta_{\mathbf{kk}'}$	$=\delta_{rs}\delta_{\mathbf{pp}'}$	$=\zeta_{\underline{r}}\delta_{\underline{r}\underline{s}}\delta_{\mathbf{kk}'}\quad\zeta_0=-1,\zeta_{1,2,3}=1$
continuous	$=\delta(\mathbf{k}-\mathbf{k}')$	$=\delta_{rs}\delta(\mathbf{p}-\mathbf{p}')$	$=\zeta_{\underline{r}}\delta_{\underline{r}\underline{s}}\delta(\mathbf{k}-\mathbf{k}')$
Other coeffs	All other commutators = 0	All other anti-commutators = 0	All other commutators = 0
The Hamiltonian Operator			
	Substituting the free field solutions into the free Hamiltonian density $\mathcal{H}_0$, integrating $H_0=\int\mathcal{H}_0\,d^3x$, and using the coefficient commutators ↑ in the result, yields ↓. Acting on states with H_0 yields number operators.		
H_0	$\sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(N_a(\mathbf{k})+\frac{1}{2}+N_b(\mathbf{k})+\frac{1}{2}\right)$	$\sum_{\mathbf{p},r}E_{\mathbf{p}}\left(N_r(\mathbf{p})-\frac{1}{2}+\bar{N}_r(\mathbf{p})-\frac{1}{2}\right)$	$\sum_{\mathbf{k},r}\omega_{\mathbf{k}}\left(N_r(\mathbf{k})+\frac{1}{2}\right)$
Number operators	$N_a(\mathbf{k})=a^\dagger(\mathbf{k})a(\mathbf{k})$ $N_b(\mathbf{k})=b^\dagger(\mathbf{k})b(\mathbf{k})$	$N_r(\mathbf{p})=c_r{}^\dagger(\mathbf{p})c_r(\mathbf{p})$ $\bar{N}_r(\mathbf{p})=d_r{}^\dagger(\mathbf{p})d_r(\mathbf{p})$	$N_r(\mathbf{k})=\zeta_r\,a_r{}^\dagger(\mathbf{k})a_r(\mathbf{k})$
Creation and Destruction Operators			
	Evaluating $N_a(\mathbf{k})\,a(\mathbf{k})\,\lvert n_{\mathbf{k}}\rangle$ (similar for other particle types) with ↑ and the coefficient commutators yields ↓		
creation	$a^\dagger(\mathbf{k}),b^\dagger(\mathbf{k})$	$c_r{}^\dagger(\mathbf{p}),d_r{}^\dagger(\mathbf{p})$	$a_r{}^\dagger(\mathbf{k})$
destruction	$a(\mathbf{k}),b(\mathbf{k})$	$c_r(\mathbf{p}),d_r(\mathbf{p})$	$a_r(\mathbf{k})$
Normaliz factors lowering	$a(\mathbf{k})\lvert n_k\rangle=\sqrt{n_k}\,\lvert n_k-1\rangle$	$c_r(\mathbf{p})\lvert\psi_{r,\mathbf{p}}\rangle=\lvert 0\rangle$	as with scalars
raising	$a^\dagger(\mathbf{k})\lvert n_k\rangle=\sqrt{n_k+1}\,\lvert n_k+1\rangle$	$c_r^\dagger(\mathbf{p})\lvert 0\rangle=\lvert\psi_{r,\mathbf{p}}\rangle$	as with scalars
tot particle num	$N(\phi)=\sum_{\mathbf{k}}\left(N_a(\mathbf{k})-N_b(\mathbf{k})\right)$	$N(\psi)=\sum_{\mathbf{p},r}\left(N_r(\mathbf{p})-\bar{N}_r(\mathbf{p})\right)$	$N\left(A^\mu\right)=\sum_{\mathbf{k},r}N_r(\mathbf{k})$
tot particle num: lowering	$\phi=\phi^+ +\phi^-$	$\psi=\psi^+ +\psi^-$	$A^{\mu+}$
raising	$\phi^\dagger=\phi^{\dagger+}+\phi^{\dagger-}$	$\bar{\psi}=\bar{\psi}^+ +\bar{\psi}^-$	$A^{\mu-}$

Four Currents and Probability			
Four currents (operators) $j^{\mu},_{\mu} = 0$	$j^{\mu} = (\rho, \mathbf{j}) = -i\left(\phi^{\dagger},^{\mu}\phi - \phi,^{\mu}\phi^{\dagger}\right)$	$j^{\mu} = (\rho, \mathbf{j}) = \bar{\psi}\gamma^{\mu}\psi$	$j^{\mu} = -i\left(A_{\alpha},^{\mu\dagger}A^{\alpha} - A_{\alpha},^{\mu}A^{\alpha\dagger}\right)$ = 0 for photons $\left(A_{\alpha}{}^{\dagger} = A_{\alpha}\right)$
	Emphasis in field theory is usually on the number of particles ($N(\mathbf{k})$ operator), and particle probability densities are rarely used. For completeness, however, and to make the connection with quantum mechanics, they are included below. (Antiparticles would have negative values of those below!)		
Single particle probability density (not operator)	$\rho(\mathbf{x},t) = \langle\phi(\mathbf{x}',t)\vert\, j^0(\mathbf{x},t)\,\vert\phi(\mathbf{x}',t)\rangle$ Note integration over $\mathbf{x}'$, not $\mathbf{x}$ For type *a* plane wave, $\rho = \frac{1}{V}$	As at left, but with Dirac j^0 above.	= 0 for chargeless particles.
Charge, not probability	Scalar type *b* particle → negative ρ. Photons → $\rho = 0$. Led to conclusion that j^0 is really proportional to *charge* probability density.		
Observables			
	Observable operators like total energy, three momentum, and charge are found by integrating corresponding density operators over all 3-space. (For spin ½, electrons assumed below with $q = -e$)		
H	$P_0 = \sum_{\mathbf{k}} \omega_{\mathbf{k}} \left(N_a(\mathbf{k}) + N_b(\mathbf{k})\right)$	$P_0 = \sum_{\mathbf{p},r} E_{\mathbf{p}} \left(N_r(\mathbf{p}) + \bar{N}_r(\mathbf{p})\right)$	$P_0 = \sum_{\mathbf{k},r} \omega_{\mathbf{k}} N_r(\mathbf{k})$
P_i = 3-momentum	$\mathbf{P} = \sum_{\mathbf{k}} \mathbf{k}\left(N_a(\mathbf{k}) + N_b(\mathbf{k})\right)$	$\mathbf{P} = \sum_{\mathbf{p},r} \mathbf{p}\left(N_r(\mathbf{p}) + \bar{N}_r(\mathbf{p})\right)$	$\mathbf{P} = \sum_{\mathbf{k},r} \mathbf{k}\, N_r(\mathbf{k})$
s^{μ}	$q\, j^{\mu} = q(\rho, \mathbf{j})$	$q\left(j^{\mu} - (\text{constant})\right) \rightarrow \partial_{\mu} s^{\mu} = 0$	0 for photons
Q	$\int s^0 d^3x = q\sum_{\mathbf{k}}\left(N_a(\mathbf{k}) - N_b(\mathbf{k})\right)$	$\int s^0 d^3x = -e\sum_{\mathbf{p},r}\left(N_r(\mathbf{p}) - \bar{N}_r(\mathbf{p})\right)$	0 for photons
Spin operator for RQM states and QFT fields	N/A	$\mathbf{\Sigma} = \Sigma_i = \frac{1}{2}\begin{bmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{bmatrix}$ $i = 1,2,3$ σ_i = 2D Pauli matrices	magnitude = 1 for photons,
Helicity operator for RQM states and QFT fields	N/A	$\frac{\mathbf{\Sigma}\cdot\mathbf{p}}{\vert\mathbf{p}\vert}$	helicity eigenstates
Spin operator for QFT states	N/A	$\int \psi^{\dagger}\mathbf{\Sigma}\psi d^3x$	magnitude = 1 for photons,
Helicity operator for QFT states	N/A	$\int \psi^{\dagger}\left(\frac{\mathbf{\Sigma}\cdot\mathbf{p}}{\vert\mathbf{p}\vert}\right)\psi d^3x$	helicity eigenstates

<table>
<tr><th colspan="4">Bosons, Fermions, and Commutators</th></tr>
<tr><td></td><td colspan="3">Operations on states with creation, destruction, and number operators above yield the properties below.</td></tr>
<tr><td>Properties of states:</td><td>$n_a(\mathbf{k}) = 0,1,2,\ldots,\infty$
So spin 0 states bosonic.</td><td>$n_r(\mathbf{p}) = 0,1$ only
So spin ½ states fermionic.</td><td>$n_r(\mathbf{k}) = 0,1,2,\ldots,\infty$
So spin 1 states bosonic.</td></tr>
<tr><td>Bosons can only employ commutators
Fermions can only employ anti-commutators</td><td>If anti-commutators used instead of commutators with Klein-Gordon equation solutions, then observable (not counting vacuum energy) Hamiltonian operator would have form
$H_0{}^0 = 0$ and $H_0{}^0 \lvert\phi_{\mathbf{k}}\rangle = 0$,i.e.,
all scalar particles would have zero energy.
Hence, we cannot use anticommutators with spin 0 bosons.</td><td>Commutators lead to 2 or more identical particle states co-existing in same multiparticle state. Anti-commutators lead to only one given single particle state per multi-particle state.
Therefore, commutators cannot be used with spin ½ fermions.
This is further proof that we need commutators with bosons.</td><td>Same as spin 0.</td></tr>
<tr><th colspan="4">The Feynman Propagator</th></tr>
<tr><td></td><td colspan="3">Creation and destruction of free particles (& antiparticles) and their propagation visualized below.</td></tr>
<tr><td>Feynman diagrams</td><td> 
</td><td> 
</td><td> 
</td></tr>
<tr><td>Step 1
Time ordered operator T</td><td colspan="3">If $t_y < t_x$, $T\{\phi(x)\phi^\dagger(y)\} = \phi(x)\phi^\dagger(y)$, i.e., the $\phi^\dagger(y)$ operates first, and should be placed on the right.
If $t_x < t_y$, $T\{\phi(x)\phi^\dagger(y)\} = \phi^\dagger(y)\phi(x)$, i.e, the $\phi(x)$ operates first, and should be placed on the right.
Note that $\phi(x)$ commutes with $\phi^\dagger(y)$ for $x \neq y$. [Fermion fields anti-commute.]</td></tr>
<tr><td>Transition amplitude
(double density in x and y)</td><td>$\langle 0 \lvert T\{\phi(x)\phi^\dagger(y)\} \rvert 0\rangle = i\Delta_F$</td><td>$\langle 0 \lvert T\{\psi_\alpha(x)\bar{\psi}_\beta(y)\} \rvert 0\rangle = iS_{F\alpha\beta}$</td><td>$\langle 0 \lvert T\{A^\mu(x)A^\nu(y)\} \rvert 0\rangle = iD_F{}^{\mu\nu}$</td></tr>
<tr><td></td><td colspan="3">The above vacuum expectation values (transition amplitudes) represent both
1) creation of a particle at y, destruction at x, and
2) creation of an antiparticle at x, destruction at y
} transition amplitude = Feynman propagator</td></tr>
<tr><td>Step 2
Propagator in terms of two commutators</td><td colspan="3">By adding a term equal to zero to the Feynman propagator above, it can be expressed as vacuum expectation values (VEVs) of two commutators</td></tr>
<tr><td></td><td>$i\Delta_F(x-y) =$
$\langle 0 \lvert \left[\phi^+(x), \phi^{\dagger -}(y)\right] \rvert 0\rangle \quad t_y < t_x$
$\langle 0 \lvert \left[\phi^{\dagger +}(y), \phi^-(x)\right] \rvert 0\rangle \quad t_x < t_y$</td><td>$iS_{F\alpha\beta}(x-y) =$
$\langle 0 \lvert \left[\psi_\alpha^+(x), \bar{\psi}_\beta^-(y)\right]_+ \rvert 0\rangle \quad t_y < t_x$
$-\langle 0 \lvert \left[\bar{\psi}_\beta^+(y), \psi_\alpha^-(x)\right]_+ \rvert 0\rangle \quad t_x < t_y$</td><td>$iD_F{}^{\mu\nu}(x-y) =$
$\langle 0 \lvert \left[A^{\mu +}(x), A^{\nu -}(y)\right] \rvert 0\rangle \quad t_y < t_x$
$\langle 0 \lvert \left[A^{\nu +}(y), A^{\mu -}(x)\right] \rvert 0\rangle \quad t_x < t_y$</td></tr>
</table>

Step 3 As 3-momentum integrals	With the coefficient commutation relations, the above two commutators (for each spin type) can be expressed as two integrals over 3-momentum space		
Definition of symbols for commutators	$\left[\phi^{+}(x),\phi^{\dagger-}(y)\right]=i\Delta^{+}(x-y)$ $\left[\phi^{\dagger+}(y),\phi^{-}(x)\right]=i\Delta^{-}(x-y)$	$\left[\psi_{\alpha}^{+}(x),\bar{\psi}_{\beta}^{-}(y)\right]_{+}=iS_{\alpha\beta}^{+}(x-y)$ $-\left[\bar{\psi}_{\beta}^{+}(y),\psi_{\alpha}^{-}(x)\right]_{+}=iS_{\alpha\beta}^{-}(x-y)$	$\left[A^{\mu+}(x),A^{\nu-}(y)\right]=iD^{\mu\nu+}(x-y)$ $\left[A^{\nu+}(y),A^{\mu-}(x)\right]=iD^{\mu\nu-}(x-y)$
	$i\Delta^{\pm}=\frac{1}{2(2\pi)^3}\int\frac{e^{\mp ik(x-y)}}{\omega_k}d^3\mathbf{k}$	$iS^{\pm}=\frac{\pm 1}{2(2\pi)^3}\int\frac{(\not{p}\pm m)e^{\mp ip(x-y)}}{E_p}d^3\mathbf{p}$	$iD^{\mu\nu\pm}=-g^{\mu\nu}i\Delta^{\pm}$
	Δ^{+}, S^{+}, $D^{\mu\nu+}$ represent particles; Δ^{-}, S^{-}, $D^{\mu\nu-}$ represent anti-particles. Symbols $S^{\pm}=S^{\pm}{}_{\alpha\beta}$ Although fields such as ϕ are operators, because of their coefficient commutation relations, each integral above is a number, not an operator. The expectation value of a number X is simply the same number X. ($\langle 0\|X\|0\rangle=X\langle 0\|0\rangle=X$. So, the Feynman propagator will also be simply a number (no brackets needed.)		
Step 4 As contour integrals	Contour integral theory (integration in the complex plane) permits the above two integrals (for each spin type) over real 3-momentum space to be expressed as contour integrals.		
	$i\Delta^{\pm}=$ $\frac{\mp i}{(2\pi)^4}\int_{C^{\pm}}\frac{e^{-ik(x-y)}}{k^2-\mu^2}d^4k$	$iS^{\pm}=$ $\frac{\mp i}{(2\pi)^4}\int_{C^{\pm}}\frac{(\not{p}+m)e^{-ip(x-y)}}{p^2-m^2}d^4p$	$iD^{\mu\nu\pm}=$ $\frac{\mp ig^{\mu\nu}}{(2\pi)^4}\int_{C^{\pm}}\frac{e^{-ik(x-y)}}{k^2-\underbrace{\mu^2}_{\text{photon}=0}}d^4k$
Step 5 As one integral	Taking certain limits with contour integrals in the complex plane yields a single form for the Feynman propagator that works for any time ordering and will prove more convenient.		
in physical space	$\Delta_F(x-y)=$ $\frac{1}{(2\pi)^4}\int\frac{e^{-ik(x-y)}}{k^2-\mu^2+i\varepsilon}d^4k$	$S_{F\alpha\beta}(x-y)=$ $\frac{1}{(2\pi)^4}\int\frac{(\not{p}+m)e^{-ip(x-y)}}{p^2-m^2+i\varepsilon}d^4p$	$D_F^{\mu\nu}(x-y)=$ $\frac{-g^{\mu\nu}}{(2\pi)^4}\int\frac{e^{-ik(x-y)}}{k^2+i\varepsilon}d^4k$
in momentum space	$\Delta_F(k)=\frac{1}{k^2-\mu^2+i\varepsilon}$	$S_{F\alpha\beta}(p)=\frac{\not{p}+m}{p^2-m^2+i\varepsilon}$	$D_F^{\mu\nu}(k)=\frac{-g^{\mu\nu}}{k^2+i\varepsilon}$

Chapter 6 *print vers 3/6/13 copyright of Robert D. Klauber*

Symmetry, Invariance, and Conservation for Free Fields

"The time has come", the walrus said, "to speak of many things,
of symmetries, Lagrangians, and changeless transformings."

Re-rendering of Lewis Carroll
by R. Klauber

6.0 Preliminaries

My apologies to Lewis Carroll for the liberties taken with his great work, but the Jabberwockian, oxymoron-like phrase "changeless transforming" will come to have deep significance for us. We will find it central to our understanding of symmetry in general, and more specifically, in our study of quantum field theory.

6.0.1 Background

Symmetry is one of the most aesthetically captivating and philosophically meaningful concepts known to mankind. Rooted originally in the arts, it has evolved and re-emerged in our modern age as a unified and holistic structural basis for all of science.

But if so, what then, particularly in mathematical terms, is it? If, in a work of art, it is a quality, perhaps somewhat abstract and related closely to feeling and emotion, how does it relate to physics? Can it be defined precisely?

We begin our answer to these questions after the chapter preview below.

6.0.2 Chapter Overview

First, an introduction to symmetry,

where we will look at

- a simple definition of symmetry without math,
- examples of symmetry, and
- a mathematical definition of symmetry.

A simple definition of symmetry with examples

Then, symmetry in classical physics, including

- laws of nature symmetric under Lorentz transformation, i.e., laws are invariant in spacetime (same for all observers)
- symmetry in the Lagrangian $L \rightarrow$ a related quantity is conserved

Symmetry in classical mechanics

Then, symmetry in quantum field theory, including

- field equations symmetric under Lorentz transformation, i.e., they are invariant in spacetime (same for all observers)
- symmetry in the Lagrangian density $\mathcal{L} \rightarrow$ a related quantity is conserved
- symmetry, gauges, and gauge theories

Symmetry in QFT

Free vs interacting fields

We will deal primarily with free particles and fields in this chapter, but the principles will apply in general, as we shall see when we investigate interactions.

Symmetry principles apply to free and interacting cases, but only free in this chapter

6.1 Introduction to Symmetry

6.1.1 Symmetry Simplified

Each of us has some intuitive feel for what symmetry is, though most might, at least at first, have some difficulty coming up with a very precise definition. Certainly snowflakes have symmetry, and so do cylinders and beach balls. A map of New York probably does not. Just what exactly is it that we sense about an object that causes us to deem it symmetric?

To see what that certain something is, imagine yourself looking at a real life version of the cylinder depicted in the figure below. Then imagine closing your eyes for a moment, and during the time you can't see, someone else rotates the cylinder about the vertical axis shown in the figure. When you open your eyes is there any way you could tell that the rotation had taken place? The answer, of course, is no, but what does that mean?

It means that even though something changed (the rotational position of the cylinder), something else remained unchanged. The form we perceive, the wholeness that is the cylinder, looks exactly the same. The act of moving or "transforming" the cylinder simultaneously exhibits the qualities of both change (transformation) and non-change (invariance).

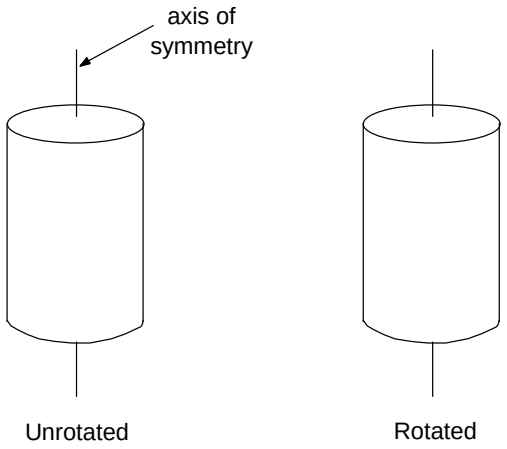


Figure 6-1. Symmetry of a Cylinder

Symmetry is the propensity for non-change with change

So what then is *symmetry*? It is simply the propensity *for non-change with change, for invariance under transformation.* In many cases, such as this one, it is a relationship between the whole and the parts in which the whole can exhibit changelessness while the component parts change. In virtually every case, it involves superficial change with more profound non-change.

Different degrees of non-change with change mean different degrees of symmetry

Symmetry manifests to greater or lesser degrees. A sphere, for instance, has more symmetry than a cylinder because it possesses innumerable (rather than only one) possible axes about which it could be rotated and still appear the same. A snowflake has even less symmetry than a cylinder, since there are only six discrete positions into which it could be rotated where no change could be discerned. A baseball glove has no symmetry whatever. There are absolutely no ways it could be rotated without looking distinctly different.

Different kinds of symmetry: rotational, translational,

Symmetry extends beyond rotation. Consider an infinite length horizontal line. Translate it 10 meters to the right. It still looks the same. It has translational symmetry. Consider the human body where the right half is reflected to the left, and the left half reflected to the right. It still looks the same (to good approximation.) To high degree, our bodies have mirror, or reflection, symmetry.

There are continuous symmetries, like the cylinder of Fig. 6-1, a sphere, or the infinite straight line discussed above. For these, transformation is continuous. And there are discrete symmetries,

like the snowflake, an infinite picket fence, or any reflection symmetry. For these, the transformation only maintains an invariant quality in certain discrete positions.

Symmetries can be continuous or discrete

Extrapolating these ideas beyond mere geometry and rotation, we can begin to understand why symmetry is considered so meaningful and fascinating. Non-change with change permeates many diverse phenomena. In many works of visual art, such as those of Escher or Indian mandalas, this principle is evident. In architecture, it has been pervasive throughout the ages. In music, the refrain, typically the essence of a song, remains the same, while other lyrics change. And that certain something we sense in the work of a great master is typically there throughout all of his or her individual pieces. We know that a Bach sonata, even if we have never heard it before, is by Bach. We know a Picasso painting, even if we have never seen it before, is by Picasso. We sense symmetry.

Symmetry plays a major role in the arts and elsewhere

6.1.2 Symmetry Mathematically

Mathematically, symmetry comprises invariance under transformation

In mathematical terms, the rotations, translations, and reflections we discussed in the previous section are known as transformations. Any transformation, by definition, is a change of something. If the transformation is symmetric, something else remains unchanged, or in math terms, invariant. Not all transformations are symmetric, of course. We will look at some mathematical examples below, but first we need to note one more thing.

Transformation is change of object with observer fixed or vice versa.

The transformation depicted in Fig. 6-1 can be understood either as a rotation of the cylinder in one direction while we remain fixed (an active transformation, by name), or alternatively, as a rotation of our viewing frame of reference in the other direction while the cylinder remains fixed (a passive transformation). The same thing is true for snowflakes, the translation of a straight line, and more. Transformations typically involve a *change of perspective*, a change in the relationship between the observer and the thing being observed.

Changing observer = changing coordinate system, most useful interpretation

Mathematically, when we change our position of observation, it is equivalent to using a new, different reference frame and coordinate system, oriented differently from, and/or displaced relative to, the original. So a transformation can be viewed simply as a change of coordinate system, and this is often represented as a shifting from unprimed to primed coordinates. We will focus on this (passive transformation) interpretation, the most common one in physics, and most relevant to QFT.

Example #1

So how about some simple examples? For starters, see Fig. 6-2, where on the left hand side we show the function

$$f\left(x^1,x^2\right)=\left(x^1\right)^2+\left(x^2\right)^2. \tag{6-1}$$

Example of a function symmetric under rotation transformation

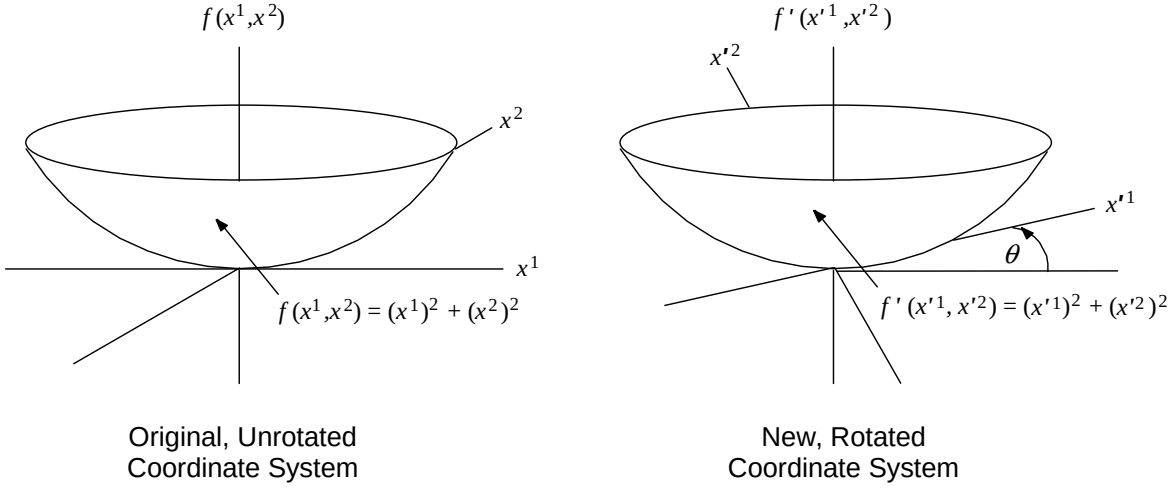


Figure 6-2. Example of a Function Symmetric Under Coordinate Transformation

We then change to a coordinate system rotated relative to the first, where our transformation from the first set of coordinates to the second is

$$x'^1 = x^1 \cos\theta + x^2 \sin\theta \qquad x'^2 = -x^1 \sin\theta + x^2 \cos\theta\,, \tag{6-2}$$

with the inverse transformation being

$$x^1 = x'^1 \cos\theta - x'^2 \sin\theta \qquad x^2 = x'^1 \sin\theta + x'^2 \cos\theta\,. \tag{6-3}$$

In matrix form, these are

$$\begin{bmatrix} x'^1 \\ x'^2 \end{bmatrix} = \underbrace{\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}}_{T} \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} \qquad \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} = \underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}_{T^{-1}=T^{\mathrm{T}}} \begin{bmatrix} x'^1 \\ x'^2 \end{bmatrix}, \tag{6-4}$$

2D *rotation transformation*

where we designate the transformation by *T*, whose inverse is its own transpose.

Substituting (6-3) into (6-1) to express our function in the new system primed coordinates yields

$$\begin{aligned} f\left(x^1, x^2\right) &= \left(x^1\right)^2 + \left(x^2\right)^2 = \\ f'\left(x'^1, x'^2\right) &= \left(x'^1 \cos\theta - x'^2 \sin\theta\right)^2 + \left(x'^1 \sin\theta + x'^2 \cos\theta\right)^2 \\ &= \left(x'^1\right)^2 \left(\cos^2\theta + \sin^2\theta\right) + \left(x'^2\right)^2 \left(\cos^2\theta + \sin^2\theta\right) = \left(x'^1\right)^2 + \left(x'^2\right)^2 = f\left(x'^1, x'^2\right). \end{aligned} \tag{6-5}$$

Function has same form in original or primed coordinates → it is symmetric under the transformation

The function has exactly the same form in both coordinate systems, exactly the same form whether we express it in terms of the unprimed or primed coordinates. Given Fig. 6-2, this should not be much of a surprise.

The prime on f' is used to indicate it has, in general, different functional form from *f*, which is the case for non-symmetric functions. But since the function *f* here is symmetric under the transformation, the functional form of *f* and f' is the same, so we drop the prime. This can be more easily understood with the following example.

Example #2

Consider the function

$$g\left(x^1, x^2\right) = \left(x^2\right)^2. \tag{6-6}$$

Example of function not symmetric under rotation transformation

Express (6-6) in the primed coordinate system by substituting (6-3) into it, and we get

$$g = \left(x^2\right)^2 = \left(x'^1 \sin\theta + x'^2 \cos\theta\right)^2 = \left(x'^1\right)^2 \sin^2\theta + \left(x'^2\right)^2 \cos^2\theta + 2x'^1 x'^2 \sin\theta\cos\theta \neq \left(x'^2\right)^2. \tag{6-7}$$

Thus, *g* has different form in the two systems and is *not* symmetric under the transformation *T*.

$$g\left(x^1, x^2\right) = g'\left(x'^1, x'^2\right) \neq g\left(x'^1, x'^2\right) \quad \text{but} \quad f\left(x^1, x^2\right) = f'\left(x'^1, x'^2\right) = f\left(x'^1, x'^2\right). \tag{6-8}$$

The transformed form of *g*, represented by g', has the same value at the same physical point, but it is not the same form in terms of the primed coordinates as *g* was in terms of the unprimed coordinates. But f', the transformed form of *f*, did have the same form in terms of both sets of coordinates, and thus, we dropped the prime on *f* on the RHS of (6-8).

In spite of its non-symmetry under rotation, *g* is symmetric under a different kind of transformation, the translation to a coordinate system which is displaced relative to the first along the x^1 axis, i.e., $x^1 \rightarrow x'^1 = x^1 + \text{constant}$, or

But same function is symmetric under a translation transformation

$$\begin{bmatrix} x'^1 \\ x'^2 \end{bmatrix} = \begin{bmatrix} x^1 \\ x^2 \end{bmatrix} + \begin{bmatrix} K \\ 0 \end{bmatrix} \qquad K = \text{constant}\,. \tag{6-9}$$

Substitution of (6-9) into (6-6) yields $g'(x'^1, x'^2) = (x'^2)^2$, having the same form in both systems.

If a coordinate is missing from f, then f is symmetric with respect to change of that coordinate

Lessons from the Examples

From Example #2, we can deduce the general rule that if a coordinate is missing in a given function, that function is invariant under a transformation solely in the direction of that coordinate

(and also under multiplication of the coordinate by a constant, which will be less important for us.) The function is symmetric with respect to that transformation.

Value of a scalar function at a physical point stays same under any transformation

In both examples and in general, the value of a particular function at a given physical point in space is the same under any transformation, symmetric or not. The new coordinates are simply a new way to designate that particular point with different numbers, but it is the same point in space, and hence must have the same numeric value for functional there. If f or g were a physical entity, like pressure, simply changing our coordinates would not change the value of the pressure at any given point in space, even though the numbers describing that point's location are different.

Form of a scalar function stays same under a symmetry transformation

So under *any* transformation of coordinate axes, the value at a physical point of every possible scalar function is invariant. Under a *symmetry* transformation the form of the function also is invariant. Under a non-symmetry transformation, the form of the function looks different in terms of the new coordinates, and we represent that functional difference with a prime on the function label.

Scalars are Invariant, Vectors are Covariant

Vector components change under transformation

Consider a 2D position vector in physical space represented in the unprimed coordinates of Example #1 by $x^i = (x^1, x^2)$. Under the rotation transformation T, this becomes $x'^i = (x'^1, x'^2) \neq (x^1, x^2)$. A different (i.e., non-invariant) set of coordinates represents the exact same vector. But it is the same vector at the same physical location, and in fact, has the same length in each coordinate system equal to

But vector length & direction in physical space unchanged for any coord system

$$|\mathbf{x}| = |x^i| = \sqrt{(x^1)^2 + (x^2)^2} = \sqrt{(x'^1)^2 + (x'^2)^2} = |x'^i|. \tag{6-10}$$

So the scalar value at the point (equal to the length of the position vector at that point) is the same in both systems, but the coordinate values are not.

Vectors are covariant under coordinate transformation

It is generally true of every vector **v**, not just the position vector shown here, that its physical, measurable length (a scalar value) remains unchanged under any coordinate transformation, but its component values change. This is called covariance. Scalar values are invariant under coordinate transformation; vector components are covariant. (Don't confuse this use of the word "covariant" with our use of the terms covariant and contravariant coordinates.)

Vector transformation symmetric if components unchanged

Parallel to scalars, if the vector components remain unchanged under a given transformation, then that transformation is a symmetry transformation, i.e., $v'^j(x'^i) = v^j(x^i)$. One example is the **E** field around a point charge, which points radially outward from the point, described in a coordinate system with origin at the point. Rotating to a new coordinate system, we find the same functional dependence of the **E** field on the new coordinates. See Prob. 7.

All of above true for 4D *and other spaces, as well*

All of these conclusions are valid for any dimension space, and in particular for our purposes, the 4D spacetime of relativity theory. They are also valid for systems of generalized coordinates, not just Cartesian like those shown here, and for both particles and fields. Probs. 1 through 6 and Wholeness Chart 6-1 can help you gain more comfort with these concepts.

Wholeness Chart 6-1. Symmetry Summary

	Non-Symmetric Transformation	**Symmetric Transformation**
Coordinate values change?	Yes	Yes
Scalar value at a physical point the same?	Yes	Yes
Form of function invariant?	No	Yes
Vector magnitude and direction at a physical point the same?	Yes	Yes
Vector components invariant?	No	Yes
Vector components vary covariantly?	Yes	No, invariant
General rule: If a function h is not a function of the jth coordinate x^j, then h is symmetric under the transformation $x^j \rightarrow x^j$ + constant		

6.2 Symmetry in Classical Mechanics

6.2.1 Invariance of the Laws of Nature

Einstein showed laws of nature symmetric (invariant) in spacetime

Symmetry turns out to play an extraordinary role in the physics of our creation. Albert Einstein, in possibly the most far reaching of any scientific discovery, provided the first insight into the universe's innate symmetry. He showed, via his theories of relativity, that even though the visible world of changing objects appears different at different places, in different times, to different observers, the physical laws of nature governing those objects remain invariant regardless of when, where, or how they are perceived. The laws of physics, acting on a subtler, more holistic level of creation, exhibit changelessness in the midst of change and are said to be *symmetric* throughout spacetime.

We review this discovery by Einstein and the classical mechanics leading up to it below. This should not be new material for most readers, but since it forms a good part of the foundation for understanding the ramifications of symmetry in physics, I provide the following overview, which many readers are probably well versed enough in to skim, or skip, over.

6.2.2 Brief History of Einstein's Insight into Symmetry

Newton's Laws: Invariant under Galilean Transformations

F=*m***a** *invariant under rotation and translation*

Newton's laws of motion are trivially symmetric under a rotation transformation because **F**=*m***a** (from which the other two laws can be derived) is a vector equation, and as we showed above, any vector maintains its magnitude and direction unchanged in physical space under rotation. So the **F** and **a** vectors will still be aligned with each other in any new coordinate system and still have the same proportionality constant of *m*. There will be new coordinates for each, but the same equation relating those coordinates will hold. If $F^i = ma^i$, then $F'^i = ma'^i$. Similar results hold for a translation transformation, which you can show by doing Prob. 8.

The (3D) *Galilean transformation is to coordinate system having constant velocity relative to original system*

Another type of transformation involves changing from one coordinate system to another where the transformed system has a constant velocity relative to the first of **v**. Any fixed coordinate value in the original system appears to move in the – **v** direction relative to the second. This, as most readers should know, is called a Galilean transformation for the 3D plus time of classical mechanics and is

$$\begin{bmatrix} x^1 \\ x^2 \\ x^3 \end{bmatrix} \to \begin{bmatrix} x'^1 \\ x'^2 \\ x'^3 \end{bmatrix} = \begin{bmatrix} x^1 - v^1 t \\ x^2 - v^2 t \\ x^3 - v^3 t \end{bmatrix} \quad \text{or} \quad \mathbf{x} \to \mathbf{x} - \mathbf{v}t\,. \tag{6-11}$$

In Newtonian/Galilean mechanics, time does not change from one system to the other. It is invariant and thus labeled by *t* in both systems above.

F=*m***a** *invariant under Galilean transformation*

Newton's second law is invariant under this transformation because of the second order time derivative in x^i. That is,

$$\begin{bmatrix} F^1 \\ F^2 \\ F^3 \end{bmatrix} = m\begin{bmatrix} \ddot{x}^1 \\ \ddot{x}^2 \\ \ddot{x}^3 \end{bmatrix} \to = m\begin{bmatrix} \ddot{x}'^1 \\ \ddot{x}'^2 \\ \ddot{x}'^3 \end{bmatrix} = m\frac{d^2}{dt^2}\begin{bmatrix} x^1 - v^1 t \\ x^2 - v^2 t \\ x^3 - v^3 t \end{bmatrix} = m\begin{bmatrix} \ddot{x}^1 \\ \ddot{x}^2 \\ \ddot{x}^3 \end{bmatrix} = \begin{bmatrix} F^1 \\ F^2 \\ F^3 \end{bmatrix} \quad \text{or} \quad \mathbf{F} = m\mathbf{a} \to \mathbf{F} = m\mathbf{a}. \tag{6-12}$$

Prior to Maxwell's appearance on the scene, it was generally assumed that all laws of nature were invariant under Galilean transformations.

Maxwell's Laws: Invariant under a Different Kind of Transformation

Maxwells' eqs NOT invariant under Galilean transformation

However, with the publication of Maxwell's equations in 1864, it was realized that his laws of nature, in contrast, do not transform symmetrically under a Galilean transformation. If one invoked (6-11) in the coordinates of Maxwell's equations, the result was a set of equations of different form, quite unlike the behavior that we saw in (6-12). That exercise is fairly involved and would lead us too far afield, so we won't get into it here.

It was, however, realized that Maxwell's equations were invariant under a different transformation between coordinate systems in relative motion. This transformation is 4D with time

and space transformations, rather than simply 3D spatial, and is called the Lorentz transformation, after its discoverer. It is, where we lose no generality by restricting relative velocity to a single coordinate direction (since Maxwell's laws are symmetric under rotation), and where we write out *c* in non-natural units this one time,

Maxwells' eqs ARE invariant under Lorentz transformation

$$\begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} \rightarrow \begin{bmatrix} x'^0 \\ x'^1 \\ x'^2 \\ x'^3 \end{bmatrix} = \begin{bmatrix} \gamma\left(x^0 - \frac{v}{c}x^1\right) \\ \gamma\left(x^1 - \frac{v}{c}x^0\right) \\ x^2 \\ x^3 \end{bmatrix} = \underbrace{\begin{bmatrix} \gamma & -\gamma\frac{v}{c} & 0 & 0 \\ -\gamma\frac{v}{c} & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\Lambda} \begin{bmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{bmatrix} \qquad \gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \quad v = v^1. \quad (6\text{-}13)$$

The (4D) *Lorentz transformation*

Again, because it would lead us away from our present tasks, we will not go through the lengthy process of showing that Maxwell's equations retain the same form under (6-13). Note that we will designate the Lorentz transformation with the symbol Λ.

Einstein's Resolution

Scientists wanted all laws symmetric under same transformation

Einstein and others wanted all equations of nature to display invariance under the same transformation. There were two major sets of laws at the time, Newton's for mechanics and Maxwell's for electromagnetism. But they didn't transform to moving coordinates in the same way. Something had to give.

Einstein figured c = same for all observers

Einstein intuited that the speed of light must be the same for all observers, whether they are fixed relative to one another or have relative constant velocity. This was quite a radical insight and turned out to be true. Since the wave solution to Maxwell's equations yielded a speed of e/m waves (light) of *c*, that meant those equations must yield the same result in any coordinate system that was not accelerating (nor in a gravitational field). To do this, they must have the same form in all such coordinate systems. The only transformation that did that was Lorentz's.

This meant Maxwell's eqs same for all

So Lorentz transfn must be correct one for e/m and mechanics laws

Einstein took that "to the bank". He knew it meant that in order for the equations of mechanics to also be symmetric under the Lorentz transformation, they must be modified. Newton's laws were not the exact truth, but only a very good approximation under normal human conditions to more accurate and precise laws.

Einstein reformulated mechanics so it obeyed Lorentz transformation

We won't write Einstein's law of mechanics here, but refer interested readers to any textbook on relativity[1]. The point is that his reformulation of mechanics 1) is invariant in form under Lorentz transformations, and 2) reduces to high accuracy, at speeds for objects of much less than *c*, to Newton's 2nd law.

That reformulation is special relativity

We summarize these results in Wholeness Chart 6-2.

Wholeness Chart 6-2. Galilean vs Lorentz Transformations

	Galilean transformation	**Lorentz transformation**
Newton's laws symmetric?	Yes	No
Maxwell's laws symmetric?	No	Yes
Einstein law of mechanics symmetric?	No	Yes
Valid for any speed?	No	Yes
Valid at low speed?	Yes, approximately	Yes

Special relativity mechanics become classical mechanics at $v << c$ ($v << 1$ in natural units)

[1] For example, Hartle, James B., *Gravity*, Pearson (2003), Chap. 5

Einstein carried this idea further in the development of his general theory of relativity. In very general terms, the same concept holds. At any given point in spacetime, the laws of nature, expressed as relationships between physical entities (like scalars, vectors, and tensors) are invariant in form. However, the Lorentz transformation is specifically for differences in velocity in non-accelerating, non-gravitational, systems. All of our work in this text will assume acceleration and gravitational effects are zero or small enough to be ignored.

Our work special relativity, not general relativity

The bottom line: All laws of nature are symmetric (invariant) under Lorentz transformation[1]. They are the same for all observers in relative constant velocity motion.

Laws of nature symmetric under Lorentz transformation

6.2.3 More with Lorentz Transformations

SEE THE TEXT FOR REST OF THIS CHAPTER

6.3 Problems

1. Is the function $F = 2(x^1)^2 + (x^2)^2$ symmetric under rotation in the x^1-x^2 plane? Guess first, then prove (or disprove) your answer by expressing F in terms of a rotated set of coordinates x'^1-x'^2, i.e, as $F'(x'^1, x'^2)$, where θ is the angle of rotation between the two coordinate systems.

2. In Prob. 1, at the point $(x^1, x^2) = (1,2)$, F has the value 6. If we transform to the rotated coordinate system x'^1-x'^2 with $\theta = 45^o$, what are the coordinates of that same physical point in space in that coordinate system? Using your expression F' for F in terms of x'^1 and x'^2, show that $F'(x'^1, x'^2)$ at that physical point equals 6, as well.

3. Without doing any calculations, is the function $G = (x^1)^2 + (x^2)^2 + (x^3)^2$ symmetric under rotation in 3D space? Is $H = (x^1)^2 + 3(x^2)^2 + (x^3)^2$?

4. Is the function $J = (x^1)^2 + (x^3)^2$ symmetric under the translation $x^2 \rightarrow x'^2 = x^2 + a$, where a is a constant? Is it symmetric under $x^3 \rightarrow x'^3 = x^3 + a$?

5. Is the differential equation $\partial_i x^i = 3$ symmetric under the translation $x^2 \rightarrow x'^2 = x^2 + a$, where a is a constant? Is it symmetric under $x^2 \rightarrow x'^2 = x^2 + (x^2)^2$?

6. Consider the position vector $(x^1, x^2) = (3,4)$. This vector's length is 5, and for the x^1 axis horizontal, its angle with the horizontal is 53^o. What are this vector's position coordinates in the x'^1-x'^2 coordinate system of Prob. 1? What is its length? Calculate it. What is its angle with the horizontal? What is its angle with respect to the x'^1 axis? Express your answer in terms of θ.

7. On page 166 we briefly discussed the spherical symmetry of the electric field around a point charge. It is easier mathematically to consider the symmetry of the simpler case of an infinitely long line of uniformly distributed charge. This radiates an electric field **E** in a coordinate system with x^3 axis aligned with the line of charge of components (where ϕ below is the relevant cylindrical coordinate system angle)

---

[1] I don't want to confuse readers, but most specialists in relativity would bring up caveats here. For one, in systems with clock synchronization done under a different convention than Einstein's, the laws of nature actually do take different form. The Lorentz transformation assumes Einstein synchronization. We will stick with that, the simplest, most efficient, synchronization and with the most widely used transformation of Lorentz. Our statements with regard to symmetries thereunder will hold true in general for our work and are widely accepted as valid criteria, which good theories should meet.

However, a second caveat involves research being done at the time of this writing that questions whether Lorentz symmetry needs to be upheld in certain very advanced theories of elementary particles.

Please do not worry about these things now. You can do so, when and if your work leads you into these other areas.

$$\begin{bmatrix} E^1 \\ E^2 \\ E^3 \end{bmatrix} = \frac{E_0}{r}\begin{bmatrix} \cos\phi \\ \sin\phi \\ 0 \end{bmatrix} = \frac{E_0}{r}\begin{bmatrix} x^1/r \\ x^2/r \\ 0 \end{bmatrix} = \frac{E_0}{\left(x^1\right)^2+\left(x^2\right)^2}\begin{bmatrix} x^1 \\ x^2 \\ 0 \end{bmatrix}.$$

Express E^i and x^i above in the primed coordinate system of Fig. 6-2 on page 164 using (6-3) to show that

$$\begin{bmatrix} E'^1 \\ E'^2 \\ E'^3 \end{bmatrix} = \frac{E_0}{\left(x'^1\right)^2+\left(x'^2\right)^2}\begin{bmatrix} x'^1 \\ x'^2 \\ 0 \end{bmatrix},$$

and thus, that the vector field components E^i in this case are symmetric under the rotation transformation of Fig. 6-2. If you feel ambitious, repeat the analysis for the **E** field around a point charge.

8. Show that $F^i = m\ddot{x}^i$ is symmetric under the transformation $x^i \rightarrow x^i + a^i$, where a^i is a constant for each i.

QED/FIELD THEORY OVERVIEW: PART 1

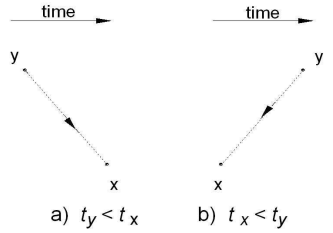
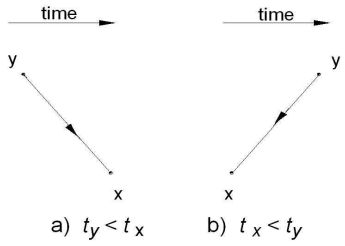
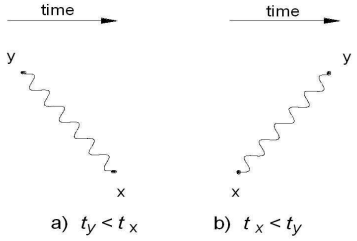
Wholeness Chart 5-4. From Field Equations to Propagators and Observables

Heisenberg Picture, Free Fields

	Spin 0	**Spin ½**	**Spin 1**
Classical Lagrangian density, free	$\mathcal{L}_0^0 = K\left(\partial_\alpha\phi\partial^\alpha\phi - \mu^2\phi\phi\right)$	None. Macroscopic spinor fields not observed.	$\mathcal{L}_0^1 = \underbrace{\frac{\mu^2}{2}A^\mu A_\mu}_{\substack{\mu=0 \\ \text{for photons}}} - \frac{1}{2}\left(\partial_\nu A_\mu\right)\left(\partial^\nu A^\mu\right)$
2^{nd} quantization, Postulate #1	Bosons: Quantum field $\mathcal{L}$ (or equivalently, $\mathcal{H}$) same as classical, fields are complex, and K =1. Spinors: Dirac eq from RQM with states → fields. Deduce $\mathcal{L}$ from Dirac eq; $\mathcal{H}$ from Legendre transf.		
QFT Lagrangian density, free	$\mathcal{L}_0^0 = \left(\partial_\alpha\phi^\dagger\partial^\alpha\phi - \mu^2\phi^\dagger\phi\right)$	$\mathcal{L}_0^{1/2} = \bar{\psi}\left(i\partial - m\right)\psi \qquad \partial = \gamma^\alpha\partial_\alpha$	As above for classical.
	$\mathcal{L}$ ↑ into the Euler-Lagrange equation yields ↓		
Free field eqs	$\left(\partial_\alpha\partial^\alpha + \mu^2\right)\phi = 0$ $\left(\partial_\alpha\partial^\alpha + \mu^2\right)\phi^\dagger = 0$	$(i\gamma^\alpha\partial_\alpha - m)\psi = 0$ $(i\partial_\alpha\bar{\psi}\gamma^\alpha + m\bar{\psi}) = 0 \qquad \bar{\psi} = \psi^\dagger\gamma^0$	$\left(\partial_\alpha\partial^\alpha + \mu^2\right)A^\mu = 0 \quad$ photon $\mu = 0$ $A^{\mu\dagger} = A^\mu$ for chargeless (photon)
Conjugate momenta	$\pi_0^0 = \frac{\partial\mathcal{L}_0^0}{\partial\dot{\phi}} = \dot{\phi}^\dagger;\ \pi_0^{0\dagger} = \frac{\partial\mathcal{L}_0^0}{\partial\dot{\phi}^\dagger} = \dot{\phi}$	$\pi^{1/2} = i\psi^\dagger;\ \ \bar{\pi}^{1/2} = 0$	$\pi_\mu^1 = -\dot{A}_\mu$
Hamiltonian density	$\mathcal{H}_0^0 = \pi_0^0\dot{\phi} + \pi_0^{0\dagger}\dot{\phi}^\dagger - \mathcal{L}_0^0$ $= \left(\dot{\phi}\dot{\phi}^\dagger + \nabla\phi^\dagger\cdot\nabla\phi + \mu^2\phi^\dagger\phi\right)$	$\mathcal{H}_0^{1/2} = \pi^{1/2}\dot{\psi} - \mathcal{L}_0^{1/2}$	$\mathcal{H}_0^1 = \pi_\mu^1\dot{A}^\mu - \mathcal{L}_0^1$
Free field solutions	$\phi = \phi^+ + \phi^-$ $\phi^\dagger = \phi^{\dagger+} + \phi^{\dagger-}$	$\psi = \psi^+ + \psi^-$ $\bar{\psi} = \bar{\psi}^+ + \bar{\psi}^-$	$A^\mu = A^{\mu+} + A^{\mu-}$ (photon)
Discrete eigenstates (Plane waves, constrained to volume V)	$\phi(x) = \sum_{\mathbf{k}}\frac{1}{\sqrt{2V\omega_{\mathbf{k}}}}\left(a(\mathbf{k})e^{-ikx} + b^\dagger(\mathbf{k})e^{ikx}\right)$ $\phi^\dagger(x) = \sum_{\mathbf{k}}\frac{1}{\sqrt{2V\omega_{\mathbf{k}}}}\left(b(\mathbf{k})e^{-ikx} + a^\dagger(\mathbf{k})e^{ikx}\right)$	$\psi = \sum_{r,\mathbf{p}}\sqrt{\frac{m}{VE_{\mathbf{p}}}}\left(c_r(\mathbf{p})u_r(\mathbf{p})e^{-ipx} + d_r^\dagger(\mathbf{p})v_r(\mathbf{p})e^{ipx}\right)$ $\bar{\psi} = \sum_{r,\mathbf{p}}\sqrt{\frac{m}{VE_{\mathbf{p}}}}\left(d_r(\mathbf{p})\bar{v}_r(\mathbf{p})e^{-ipx} + c_r^\dagger(\mathbf{p})\bar{u}_r(\mathbf{p})e^{ipx}\right)$	$A^\mu = \sum_{r,\mathbf{k}}\frac{1}{\sqrt{2V\omega_{\mathbf{k}}}}\left(\varepsilon_r^\mu(\mathbf{k})a_r(\mathbf{k})e^{-ikx} + \varepsilon_r^\mu(\mathbf{k})a_r^\dagger(\mathbf{k})e^{ikx}\right)$
Continuous eigenstates (Plane waves, no volume constraint)	$\phi(x) = \int\frac{d\boldsymbol{k}}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}}}}\left(a(\mathbf{k})e^{-ikx} + b^\dagger(\mathbf{k})e^{ikx}\right)$ $\phi^\dagger(x) = \int\frac{d\boldsymbol{k}}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}}}}\left(b(\mathbf{k})e^{-ikx} + a^\dagger(\mathbf{k})e^{ikx}\right)$	$\psi = \sum_r\sqrt{\frac{m}{(2\pi)^3}}\int\frac{d^3\mathbf{p}}{\sqrt{E_{\mathbf{p}}}}\left(c_r(\mathbf{p})u_r(\mathbf{p})e^{-ipx} + d_r^\dagger(\mathbf{p})v_r(\mathbf{p})e^{ipx}\right)$ $\bar{\psi} = \sum_r\sqrt{\frac{m}{(2\pi)^3}}\int\frac{d^3\mathbf{p}}{\sqrt{E_{\mathbf{p}}}}\left(d_r(\mathbf{p})\bar{v}_r(\mathbf{p})e^{-ipx} + c_r^\dagger(\mathbf{p})\bar{u}_r(\mathbf{p})e^{ipx}\right)$ spinor indices on u_r, v_r, and ψ suppressed. r = 1,2.	$A^\mu = \sum_r\frac{1}{\sqrt{2(2\pi)^3}}\int\frac{d\mathbf{k}}{\sqrt{\omega_{\mathbf{k}}}}\left(\varepsilon_r^\mu(\mathbf{k})a_r(\mathbf{k})e^{-ikx} + \varepsilon_r^\mu(\mathbf{k})a_r^\dagger(\mathbf{k})e^{ikx}\right)$ r = 0,1,2,3 (4 polarization vectors)

2nd quantization Postulate #2	Bosons: $\left[\phi^r(\mathbf{x},t),\pi_s(\mathbf{y},t)\right] = \left[\phi^r\pi_s \mp \pi_s\phi^r\right] = i\delta^r{}_s\delta(\mathbf{x}-\mathbf{y})$, ϕ^r = any field, other commutators = 0. Spinors: Coefficient anti-commutation relations parallel coefficient commutation relations for bosons.		
	Bosons: using conjugate momenta expressions in ↑ yields ↓		
Equal time commutators (intermediate step only)	$\left[\phi(\mathbf{x},t),\dot{\phi}^\dagger(\mathbf{y},t)\right] = i\delta(\mathbf{x}-\mathbf{y})$	Not needed for spinor derivation.	$\left[A^\mu(\mathbf{x},t),\dot{A}^\nu(\mathbf{y},t)\right] = -ig^{\mu\nu}\delta(\mathbf{x}-\mathbf{y})$
	Bosons: Using free field solutions in ↑ with 3D Dirac delta function (e.g., for discrete solutions, $\delta(\mathbf{x}-\mathbf{y}) = \frac{1}{V}\sum_{n=-\infty}^{+\infty} e^{-i\mathbf{k}_n\cdot(\mathbf{x}-\mathbf{y})}$), and matching terms, yields the coefficient commutators ↓.		
Coefficient commutators	$\left[a(\mathbf{k}),a^\dagger(\mathbf{k}')\right] = \left[b(\mathbf{k}),b^\dagger(\mathbf{k}')\right]$	$\left[c_r(\mathbf{p}),c_s^\dagger(\mathbf{p}')\right]_+ = \left[d_r(\mathbf{p}),d_s^\dagger(\mathbf{p}')\right]_+$	$\left[a_r(\mathbf{k}),a_s^\dagger(\mathbf{k}')\right]$
discrete	$= \delta_{\mathbf{kk}'}$	$= \delta_{rs}\delta_{\mathbf{pp}'}$	$= \zeta_r\delta_{rs}\delta_{\mathbf{kk}'}$ $\quad \zeta_0 = -1,\ \zeta_{1,2,3} = 1$
continuous	$= \delta(\mathbf{k}-\mathbf{k}')$	$= \delta_{rs}\delta(\mathbf{p}-\mathbf{p}')$	$= \zeta_r\delta_{rs}\delta(\mathbf{k}-\mathbf{k}')$
Other coeffs	All other commutators = 0	All other anti-commutators = 0	All other commutators = 0
The Hamiltonian Operator			
	Substituting the free field solutions into the free Hamiltonian density $\mathcal{H}_0$, integrating $H_0 = \int\mathcal{H}_0\, d^3x$, and using the coefficient commutators ↑ in the result, yields ↓. Acting on states with H_0 yields number operators.		
H_0	$\sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(N_a(\mathbf{k})+\frac{1}{2}+N_b(\mathbf{k})+\frac{1}{2}\right)$	$\sum_{\mathbf{p},r}E_{\mathbf{p}}\left(N_r(\mathbf{p})-\frac{1}{2}+\bar{N}_r(\mathbf{p})-\frac{1}{2}\right)$	$\sum_{\mathbf{k},r}\omega_{\mathbf{k}}\left(N_r(\mathbf{k})+\frac{1}{2}\right)$
Number operators	$N_a(\mathbf{k}) = a^\dagger(\mathbf{k})a(\mathbf{k})$ $N_b(\mathbf{k}) = b^\dagger(\mathbf{k})b(\mathbf{k})$	$N_r(\mathbf{p}) = c_r{}^\dagger(\mathbf{p})c_r(\mathbf{p})$ $\bar{N}_r(\mathbf{p}) = d_r{}^\dagger(\mathbf{p})d_r(\mathbf{p})$	$N_r(\mathbf{k}) = \zeta_r\, a_r{}^\dagger(\mathbf{k})a_r(\mathbf{k})$
Creation and Destruction Operators			
	Evaluating $N_a(\mathbf{k})\, a(\mathbf{k})\, \lvert n_{\mathbf{k}}\rangle$ (similar for other particle types) with ↑ and the coefficient commutators yields ↓		
creation	$a^\dagger(\mathbf{k}),\ b^\dagger(\mathbf{k})$	$c_r{}^\dagger(\mathbf{p}),\ d_r{}^\dagger(\mathbf{p})$	$a_r{}^\dagger(\mathbf{k})$
destruction	$a(\mathbf{k}),\ b(\mathbf{k})$	$c_r(\mathbf{p}),\ d_r(\mathbf{p})$	$a_r(\mathbf{k})$
Normaliz factors lowering	$a(\mathbf{k})\lvert n_k\rangle = \sqrt{n_k}\ \lvert n_k-1\rangle$	$c_r(\mathbf{p})\lvert\psi_{r,\mathbf{p}}\rangle = \lvert 0\rangle$	as with scalars
raising	$a^\dagger(\mathbf{k})\lvert n_k\rangle = \sqrt{n_k+1}\,\lvert n_k+1\rangle$	$c_r^\dagger(\mathbf{p})\lvert 0\rangle = \lvert\psi_{r,\mathbf{p}}\rangle$	as with scalars
tot particle num	$N(\phi) = \sum_{\mathbf{k}}\left(N_a(\mathbf{k})-N_b(\mathbf{k})\right)$	$N(\psi) = \sum_{\mathbf{p},r}\left(N_r(\mathbf{p})-\bar{N}_r(\mathbf{p})\right)$	$N\left(A^\mu\right) = \sum_{\mathbf{k},r}N_r(\mathbf{k})$
tot particle num: lowering	$\phi = \phi^+ + \phi^-$	$\psi = \psi^+ + \psi^-$	$A^{\mu+}$
raising	$\phi^\dagger = \phi^{\dagger+} + \phi^{\dagger-}$	$\bar{\psi} = \bar{\psi}^+ + \bar{\psi}^-$	$A^{\mu-}$

Four Currents and Probability			
Four currents (operators) $j^{\mu},_{\mu} = 0$	$j^{\mu} = (\rho, \mathbf{j}) = -i\left(\phi^{\dagger},^{\mu}\phi - \phi,^{\mu}\phi^{\dagger}\right)$	$j^{\mu} = (\rho, \mathbf{j}) = \bar{\psi}\gamma^{\mu}\psi$	$j^{\mu} = -i\left(A_{\alpha}{}^{\dagger},^{\mu}A^{\alpha} - A_{\alpha},^{\mu}A^{\alpha\dagger}\right)$ = 0 for photons $\left(A_{\alpha}{}^{\dagger} = A_{\alpha}\right)$
	Emphasis in field theory is usually on the number of particles ($N(\mathbf{k})$ operator), and particle probability densities are rarely used. For completeness, however, and to make the connection with quantum mechanics, they are included below. (Antiparticles would have negative values of those below!)		
Single particle probability density (not operator)	$\rho(\mathbf{x},t) =$ $\langle\phi(\mathbf{x}',t)\vert\, j^0(\mathbf{x},t)\,\vert\phi(\mathbf{x}',t)\rangle$ Note integration over $\mathbf{x}'$, not $\mathbf{x}$ For type a plane wave, $\rho = \frac{1}{V}$	As at left, but with Dirac j^0 above.	= 0 for chargeless particles.
Charge, not probability	Scalar type b particle → negative ρ. Photons → $\rho = 0$. Led to conclusion that j^0 is really proportional to *charge* probability density.		
Observables			
	Observable operators like total energy, three momentum, and charge are found by integrating corresponding density operators over all 3-space. (For spin ½, electrons assumed below with q = - e)		
H	$P_0 = \sum_{\mathbf{k}} \omega_{\mathbf{k}} \left(N_a(\mathbf{k}) + N_b(\mathbf{k})\right)$	$P_0 = \sum_{\mathbf{p},r} E_{\mathbf{p}} \left(N_r(\mathbf{p}) + \bar{N}_r(\mathbf{p})\right)$	$P_0 = \sum_{\mathbf{k},r} \omega_{\mathbf{k}} N_r(\mathbf{k})$
P_i = 3-momentum	$\mathbf{P} = \sum_{\mathbf{k}} \mathbf{k}\left(N_a(\mathbf{k}) + N_b(\mathbf{k})\right)$	$\mathbf{P} = \sum_{\mathbf{p},r} \mathbf{p}\left(N_r(\mathbf{p}) + \bar{N}_r(\mathbf{p})\right)$	$\mathbf{P} = \sum_{\mathbf{k},r} \mathbf{k}\, N_r(\mathbf{k})$
s^{μ}	$q\, j^{\mu} = q(\rho, \mathbf{j})$	$q\left(j^{\mu} - (\text{constant})\right) \rightarrow \partial_{\mu} s^{\mu} = 0$	0 for photons
Q	$\int s^0 d^3x =$ $q\sum_{\mathbf{k}}\left(N_a(\mathbf{k}) - N_b(\mathbf{k})\right)$	$\int s^0 d^3x =$ $-e\sum_{\mathbf{p},r}\left(N_r(\mathbf{p}) - \bar{N}_r(\mathbf{p})\right)$	0 for photons
Spin operator for RQM states and QFT fields	N/A	$\mathbf{\Sigma} = \Sigma_i = \frac{1}{2}\begin{bmatrix}\sigma_i & 0\\ 0 & \sigma_i\end{bmatrix}$ $i = 1,2,3$ σ_i = 2D Pauli matrices	magnitude = 1 for photons,
Helicity operator for RQM states and QFT fields	N/A	$\frac{\mathbf{\Sigma}\cdot\mathbf{p}}{\vert\mathbf{p}\vert}$	helicity eigenstates
Spin operator for QFT states	N/A	$\int \psi^{\dagger}\mathbf{\Sigma}\psi d^3x$	magnitude = 1 for photons,
Helicity operator for QFT states	N/A	$\int \psi^{\dagger}\left(\frac{\mathbf{\Sigma}\cdot\mathbf{p}}{\vert\mathbf{p}\vert}\right)\psi d^3x$	helicity eigenstates

Bosons, Fermions, and Commutators			
	Operations on states with creation, destruction, and number operators above yield the properties below.		
Properties of states:	$n_a(\mathbf{k}) = 0,1,2,\ldots,\infty$ So spin 0 states bosonic.	$n_r(\mathbf{p}) = 0,1$ only So spin ½ states fermionic.	$n_r(\mathbf{k}) = 0,1,2,\ldots,\infty$ So spin 1 states bosonic.
Bosons can only employ commutators Fermions can only employ anti-commutators	If anti-commutators used instead of commutators with Klein-Gordon equation solutions, then observable (not counting vacuum energy) Hamiltonian operator would have form $H_0^0 = 0$ and $H_0^0 \lvert\phi_{\mathbf{k}}\rangle = 0$,i.e., all scalar particles would have zero energy. Hence, we cannot use anticommutators with spin 0 bosons.	Commutators lead to 2 or more identical particle states co-existing in same multiparticle state. Anti-commutators lead to only one given single particle state per multi-particle state. Therefore, commutators cannot be used with spin ½ fermions. This is further proof that we need commutators with bosons.	Same as spin 0.
The Feynman Propagator			
	Creation and destruction of free particles (& antiparticles) and their propagation visualized below.		
Feynman diagrams			
Step 1 Time ordered operator T	If $t_y < t_x$, $T\{\phi(x)\phi^\dagger(y)\} = \phi(x)\phi^\dagger(y)$, i.e., the $\phi^\dagger(y)$ operates first, and should be placed on the right. If $t_x < t_y$, $T\{\phi(x)\phi^\dagger(y)\} = \phi^\dagger(y)\phi(x)$, i.e, the $\phi(x)$ operates first, and should be placed on the right. Note that $\phi(x)$ commutes with $\phi^\dagger(y)$ for $x \neq y$. [Fermion fields anti-commute.]		
Transition amplitude (double density in x and y)	$\langle 0 \lvert T\{\phi(x)\phi^\dagger(y)\} \rvert 0\rangle = i\Delta_F$	$\langle 0 \lvert T\{\psi_\alpha(x)\bar{\psi}_\beta(y)\} \rvert 0\rangle = iS_{F\alpha\beta}$	$\langle 0 \lvert T\{A^\mu(x)A^\nu(y)\} \rvert 0\rangle = iD_F{}^{\mu\nu}$
The above vacuum expectation values (transition amplitudes) represent both 1) creation of a particle at y, destruction at x, and 2) creation of an antiparticle at x, destruction at y } transition amplitude = Feynman propagator			

Step 2 Propagator in terms of two commutators	By adding a term equal to zero to the Feynman propagator above, it can be expressed as vacuum expectation values (VEVs) of two commutators		
	$i\Delta_F(x-y)=$ $\langle 0\vert\left[\phi^+(x),\phi^{\dagger-}(y)\right]\vert 0\rangle \;\; t_y<t_x$ $\langle 0\vert\left[\phi^{\dagger+}(y),\phi^-(x)\right]\vert 0\rangle \;\; t_x<t_y$	$iS_{F\alpha\beta}(x-y)=$ $\langle 0\vert\left[\psi^+_\alpha(x),\bar{\psi}^-_\beta(y)\right]_+\vert 0\rangle \;\; t_y<t_x$ $-\langle 0\vert\left[\bar{\psi}^+_\beta(y),\psi^-_\alpha(x)\right]_+\vert 0\rangle \;\; t_x<t_y$	$iD_F{}^{\mu\nu}(x-y)=$ $\langle 0\vert\left[A^{\mu+}(x),A^{\nu-}(y)\right]\vert 0\rangle \;\; t_y<t_x$ $\langle 0\vert\left[A^{\nu+}(y),A^{\mu-}(x)\right]\vert 0\rangle \;\; t_x<t_y$
Step 3 As 3-momentum integrals	With the coefficient commutation relations, the above two commutators (for each spin type) can be expressed as two integrals over 3-momentum space		
Definition of symbols for commutators	$\left[\phi^+(x),\phi^{\dagger-}(y)\right]=i\Delta^+(x-y)$ $\left[\phi^{\dagger+}(y),\phi^-(x)\right]=i\Delta^-(x-y)$	$\left[\psi^+_\alpha(x),\bar{\psi}^-_\beta(y)\right]_+=iS^+_{\alpha\beta}(x-y)$ $-\left[\bar{\psi}^+_\beta(y),\psi^-_\alpha(x)\right]_+=iS^-_{\alpha\beta}(x-y)$	$A^{\mu+}(x),A^{\nu-}(y)\big]=iD^{\mu\nu+}(x-y)$ $A^{\nu+}(y),A^{\mu-}(x)\big]=iD^{\mu\nu-}(x-y)$
	$i\Delta^\pm=\frac{1}{2(2\pi)^3}\int\frac{e^{\mp ik(x-y)}}{\omega_k}d^3\mathbf{k}$	$iS^\pm=\frac{\pm 1}{2(2\pi)^3}\int\frac{(\not{p}\pm m)e^{\mp ip(x-y)}}{E_p}d^3\mathbf{p}$	$iD^{\mu\nu\pm}=-g^{\mu\nu}i\Delta^\pm$
	Δ^+, S^+, $D^{\mu\nu+}$ represent particles; Δ^-, S^-, $D^{\mu\nu-}$ represent anti-particles. Symbols $S^\pm=S^\pm{}_{\alpha\beta}$ Although fields such as ϕ are operators, because of their coefficient commutation relations, each integral above is a number, not an operator. The expectation value of a number X is simply the same number X. ($\langle 0\vert X\vert 0\rangle=X\langle 0\vert 0\rangle=X$. So, the Feynman propagator will also be simply a number (no brackets needed.)		
Step 4 As contour integrals	Contour integral theory (integration in the complex plane) permits the above two integrals (for each spin type) over real 3-momentum space to be expressed as contour integrals.		
	$i\Delta^\pm=$ $\frac{\mp i}{(2\pi)^4}\int_{C^\pm}\frac{e^{-ik(x-y)}}{k^2-\mu^2}d^4k$	$iS^\pm=$ $\frac{\mp i}{(2\pi)^4}\int_{C^\pm}\frac{(\not{p}+m)e^{-ip(x-y)}}{p^2-m^2}d^4p$	$iD^{\mu\nu\pm}=$ $\frac{\mp ig^{\mu\nu}}{(2\pi)^4}\int_{C^\pm}\frac{e^{-ik(x-y)}}{k^2\underbrace{-\mu^2}_{\text{photon}=0}}d^4k$
Step 5 As one integral	Taking certain limits with contour integrals in the complex plane yields a single form for the Feynman propagator that works for any time ordering and will prove more convenient.		
in physical space	$\Delta_F(x-y)=$ $\frac{1}{(2\pi)^4}\int\frac{e^{-ik(x-y)}}{k^2-\mu^2+i\varepsilon}d^4k$	$S_{F\alpha\beta}(x-y)=$ $\frac{1}{(2\pi)^4}\int\frac{(\not{p}+m)e^{-ip(x-y)}}{p^2-m^2+i\varepsilon}d^4p$	$D_F^{\mu\nu}(x-y)=$ $\frac{-g^{\mu\nu}}{(2\pi)^4}\int\frac{e^{-ik(x-y)}}{k^2+i\varepsilon}d^4k$
in momentum space	$\Delta_F(k)=\frac{1}{k^2-\mu^2+i\varepsilon}$	$S_{F\alpha\beta}(p)=\frac{\not{p}+m}{p^2-m^2+i\varepsilon}$	$D_F^{\mu\nu}(k)=\frac{-g^{\mu\nu}}{k^2+i\varepsilon}$

Prob 17-7) Prove xxx(17-12)

$$\frac{1}{4}\sum_{\text{spins}}|\mathcal{M}_{B2}|^2 = \frac{1}{4}\sum_{\text{spins}} \mathcal{M}_{B2}\mathcal{M}^{*}_{B2} = \frac{1}{4}\Big(-ie^2\,\bar{u}_{s_1'}\gamma^{\alpha}u_{s_1}\frac{1}{(p_2-p_2')^2}\bar{v}_{s_2}\gamma^{\alpha}v_{s_2'}\Big)\times$$
$$\Big(\text{“ }\gamma_\beta\text{ ” “ “ “ }\gamma^\beta\text{ ”}\Big)^{*}$$

$$\frac{1}{4}\sum_{\text{spins}}\frac{\overbrace{-i^2e^4}^{1}}{(p_2-p_2')^4}\Big(\bar{u}_{s_1'}\gamma_\alpha u_{s_1}\Big)\Big(\bar{v}_{s_2}\gamma^\alpha v_{s_2'}\Big)\Big(\bar{u}_{s_1}\gamma_\beta u_{s_1'}\Big)\Big(\bar{v}_{s_2'}\gamma^\beta v_{s_2}\Big) \qquad (1)$$

$\frac{1}{|\Pi_2|^2}$; part of $A_{\alpha\beta}$; part of $B^{\alpha\beta}$

$$A_{\alpha\beta} = \sum_{s_1'}\sum_{s_1}\Big(\bar{u}_{s_1'\delta}(\gamma_\alpha)_{\delta\eta}u_{s_1\eta}\Big)\Big(\bar{u}_{s_1\rho}(\gamma_\beta)_{\rho\sigma}u_{s_1'\sigma}\Big)$$

$$= \Big(\sum_{s_1'}u_{s_1'\sigma}\bar{u}_{s_1'\delta}\Big)(\gamma_\alpha)_{\delta\eta}\Big(\sum_{s_1}u_{s_1\eta}\bar{u}_{s_1\rho}\Big)(\gamma_\beta)_{\rho\sigma}$$

$$\Big(\frac{\not{p}_1'+m}{2m}\Big)_{\sigma\delta} \qquad \Big(\frac{\not{p}_1+m}{2m}\Big)_{\eta\rho}$$

$$= \mathrm{Tr}\Big\{\Big(\frac{\not{p}_1'+m}{2m}\Big)\gamma_\alpha\Big(\frac{\not{p}_1+m}{2m}\Big)\gamma_\beta\Big\}$$

$$B^{\alpha\beta} = \sum_{s_2'}\sum_{s_2}\Big(\bar{v}_{s_2\delta}(\gamma^\alpha)_{\delta\eta}v_{s_2'\eta}\Big)\Big(\bar{v}_{s_2'\rho}(\gamma^\beta)_{\rho\sigma}v_{s_2\sigma}\Big)$$

$$= \Big(\sum_{s_2'}v_{s_2'\eta}\bar{v}_{s_2'\rho}\Big)(\gamma^\beta)_{\rho\sigma}\Big(\sum_{s_2}v_{s_2\sigma}\bar{v}_{s_2\delta}\Big)(\gamma^\alpha)_{\delta\eta}$$

$$= \mathrm{Tr}\Big\{\Big(\frac{\not{p}_2'-m}{2m}\Big)\gamma^\beta\Big(\frac{\not{p}_2-m}{2m}\Big)\gamma^\alpha\Big\} = \mathrm{Tr}\Big\{\Big(\frac{\not{p}_2-m}{2m}\Big)\gamma^\alpha\Big(\frac{\not{p}_2'-m}{2m}\gamma^\beta\Big)\Big\}$$

agrees with Ch 17-5 last eq

an aside

we can use either form of above

Here, we will use the first form

continued on next page

Prob 17-7 CONT.

All traces of odd num gamma matrices = 0

Assume relativistic, $E \approx |\vec{p}| \gg m$ for all particles

Then

$$A_{\alpha\beta} \approx \frac{1}{4m^2} \mathrm{Tr}\{ \not{p}_1' \gamma_\alpha \not{p}_1 \gamma_\beta \} = \frac{1}{4m^2} (p_1'^\delta p_1^\eta) \mathrm{Tr}\{ \gamma_\delta \gamma_\alpha \gamma_\eta \gamma_\beta \}$$

$$4(g_{\delta\alpha} g_{\eta\beta} - g_{\delta\eta} g_{\alpha\beta} + g_{\delta\beta} g_{\alpha\eta})$$

$$= \frac{1}{m^2} \left(p_{1\alpha}' p_{1\beta} - (p_1' p_1) g_{\alpha\beta} + p_{1\beta}' p_{1\alpha} \right)$$

$$= \frac{1}{m^2} \left(p_{1\alpha}' p_{1\beta} + p_{1\alpha} p_{1\beta}' - (p_1' p_1) g_{\alpha\beta} \right)$$

$$B^{\alpha\beta} \approx \frac{1}{4m^2} \mathrm{Tr}\{ \not{p}_2' \gamma^\beta \not{p}_2 \gamma^\alpha \}$$ like — except $1 \leftrightarrow 2$, $\alpha \leftrightarrow \beta$ & minus sign incl.

$$= \frac{1}{m^2} \left(p_2'^\beta p_2^\alpha + p_2^\beta p_2'^\alpha - (p_2' p_2) g^{\alpha\beta} \right)$$

$$A_{\alpha\beta} B^{\alpha\beta} = \frac{1}{m^4} \Big((p_1' p_2)(p_1 p_2') + (p_1' p_2')(p_1 p_2) - (p_1' p_1)(p_2' p_2)$$

$$+ (p_1 p_2)(p_1' p_2') + (p_1 p_2')(p_1' p_2) - (p_1 p_1')(p_2' p_2)$$

$$- (p_1' p_1)(p_2' p_2) - (p_1' p_1)(p_2 p_2') + 4 (p_1' p_1)(p_2' p_2) \Big)$$

$$= \frac{1}{m^4} \left(2 (p_1' p_2)(p_1 p_2') + 2 (p_1 p_2)(p_1' p_2') \right) = \frac{2E^4}{m^4} \left((1 + \cos\theta)^2 + 4 \right)$$

$(p_1' p_2)(p_1 p_2') \approx (E^2 + E^2\cos\theta)^2$ from xxx (17-116)

$(p_1 p_2) \approx 2E^2$ Also From xxx

$(1 + \cos\theta)^2 = (2\cos^2\theta/2)^2$

$$= \frac{2E^4}{m^4} \left(4\cos^4\theta/2 + 4 \right)$$

Back to (1)

$$\frac{1}{4} \sum_{\text{spins}} |\mathcal{M}_{B2}|^2 \approx \frac{1}{4} |\Gamma_2|^2 A_{\alpha\beta} B^{\alpha\beta} = \frac{e^2}{(p_2 - p_2')^4} \frac{E^4}{m^4} \left(\cos^4\frac{\theta}{2} + 1 \right)$$

From xxx (7-139) $(p_2 - p_2')^2 = -4E^2 \sin^2(\theta/2)$

$$= \frac{1}{4} \frac{e^2}{m^4 \, 16 \sin^4(\theta/2)} 8 \left(\cos^4(\theta/2) + 1 \right) = \frac{e}{8m^4} \frac{\left(1 + \cos^4(\theta/2)\right)}{\sin^4(\theta/2)}$$

Addenda

"My father says that almost everyone is asleep. Those few who are awake go around being amazed all the time."

Meg Ryan's character in
movie *Joe Versus the Volcano*

Chapter 18 Path Integrals in Quantum Theories: A Pedagogic First Step

Chapter 19 Looking Backward and Looking Forward: Book Summary and What's Next

Chapter 18 *print vers 3/6/13Copyright of Robert D. Klauber*

Path Integrals in Quantum Theories: A Pedagogic First Step

The universe in each dimension
is vast beyond all comprehension.
A myriad of mysteries,
a multitude of histories …
From *Divine Intentions*
by R. Klauber

18.0 Preliminaries

As I mentioned on the first page of the book, I strongly believe it far easier, and more meaningful, for students to learn quantum field theory (QFT) first by the canonical quantization method, and once that has been digested, move on to the path integral (functional integral, many paths, or sum over histories) approach (functional quantization). The rest of the book is devoted to the first of these; the present chapter, to a brief introduction to the second.

Two approaches to (*ways to quantize*) *QFT:*
1) *canonical*
2) *path integral*

18.0.1 Chapter Overview

This chapter was composed so it can be read independently of (without reading) the rest of the book. So, some things may be defined/discussed again herein that are covered elsewhere in the text.

#1 simpler, rest of book;
#2 introduced in this chapter

In this chapter, we will discuss

- the mathematics of functionals, functional derivatives, and functional integrals, and
- the different kinds of integration with functionals,

then, with regard to non relativistic quantum mechanicis (NRQM),

We'll examine path integrals:
- *math behind*
- *NRQM*
- *QFT*

- transition amplitudes for position eigenstates,
- the role of the Lagrangian and the wave function peak,
- the central idea in Feynman's path integral approach,
- expressing that idea mathematically, including Feynmans' three postulates,
- comparing the path integral approach in NRQM to Schrödinger and Heisenberg's,
- determining the transition amplitude from the path integral, and
- applying the theory to an example.

Then, with regard to QFT, we will investigate

- comparing particle theory (NRQM) to field theory (QFT)
- "derivation" of the many paths approach to QFT, and
- deducing the form of the transition amplitude for QFT

18.1 Background Math: Examples and Definitions

18.1.1 Functionals

Defining the "functional"

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:

Symbol for a functional

$$F\left[x(t)\right] \text{ or } F\left[x(t),\dot{x}(t)\right] \tag{18-1}$$

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 (18-1). 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 (18-1)) function of t, i.e.

$$F\left[x(t),\dot{x}(t),t\right]. \tag{18-2}$$

18.1.2 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:

Functional derivative

$$\frac{\delta F}{\delta x(t)} \text{ or } \frac{\delta F}{\delta x} \tag{18-3}$$

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

18.1.3 Functional Integral

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

Symbolism:

Functional integral

$$\int_{x_a}^{x_b} F\,\delta x \text{ or } \int_{x_a}^{x_b} F\,\delta x(t) \tag{18-4}$$

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

18.2 Different Kinds of Integration with Functionals

Different kinds of integration with functional integrands

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 (18-1) 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 Wholeness Chart 18-1 below. The first three kinds of integration shown below are fairly self explanatory.

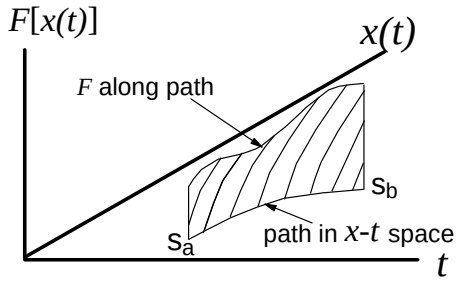
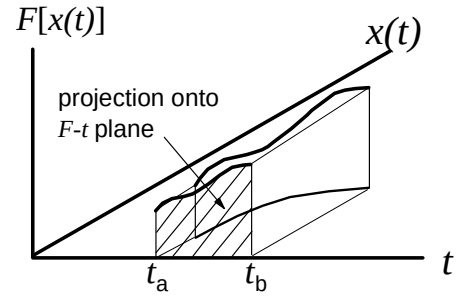
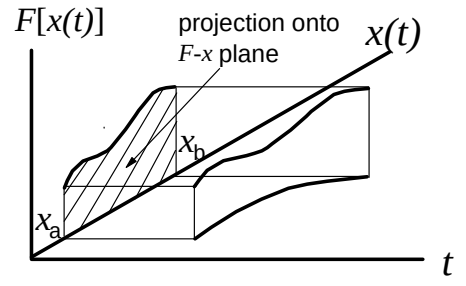
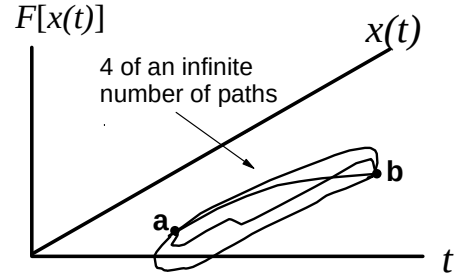
Functional integral defined as #3 in Chart 18-1

The fourth way to integrate 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.

Feynman use of functional integral is #4

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.

Wholeness Chart 18-1. 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		$\int_{s_a}^{s_b} F ds$ where s is spacetime distance along path	No real physical application.
2.	Projection of the area in 1 onto the 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		$\int_{x_a}^{x_b} F \delta x(t)$	This is the usual definition of "functional integral" This is starting point for 4, below
4.	Simultaneous integration as in #3 above over all possible paths in $x(t)$ vs. t space		$\int_{x_a}^{x_b} F \mathcal{D}x(t)$	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

18.3 The Transition Amplitude

18.3.1 General Wave Functions (States)

Review of states, amplitudes, & probability

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),

$$\psi = A_1\psi_1 + A_2\psi_2 + A_3\psi_3 , \tag{18-5}$$

$\psi \to \psi_1$ transition amplitude = A_1

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

$$A_1^* A_1 = |A_1|^2 . \tag{18-6}$$

Probability of transition = $|A_1|^2$

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. (18-6) 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. (As discussed in Chaps. 1, 7, 8, etc.)

Square of absolute value of transition amplitude = probability of transition

Symbolism: The transition amplitude for time of interaction approaching infinity, as in the canonical quantization approach, is typically written as S_{fi} (see chapters cited above). However, in the path integral approach, where elapsed time T between measurements of the initial state ψ_i and final state ψ_f is commonly finite, it is more typical to write

$$U\left(\psi_i,\psi_f;T\right) \qquad \left(=S_{fi} \text{ for } T\to\infty \text{ of canonical quantization}\right). \tag{18-7}$$

Symbol U for path integrals with T finite; S_{fi} for $T\to\infty$ in canonical case

This terminology carries over to inelastic cases (where 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 (2 γ) states. (Almost all of QFT, as seen in the rest of this book, is devoted ultimately to determining the transition amplitudes for the different possible interactions between particles.)

Schrödinger Approach Amplitudes

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

$$U(\psi_i,\psi_f;T)=\underbrace{\langle\psi_f|}_{\substack{\text{final state}\\ \text{measured}\\ \text{at } T+t_a}}\underbrace{e^{-iHT/\hbar}\underbrace{|\psi_i\rangle}_{\substack{\text{inital state}\\ \text{at } t_a}}}_{\substack{\text{evolved state}\\ \text{at } T+t_a}}, \tag{18-8}$$

U for NRQM Schrödinger wave mechanics approach

where H is the Hamiltonian operator, and we retain the symbol $\hbar$ even though $\hbar = 1$ in natural units.

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.

18.3.2 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

$$U(x_i,x_f;T)=\langle x_f|\underbrace{e^{-iHT/\hbar}|x_i\rangle}_{\text{evolved state }\psi}. \tag{18-9}$$

$|x_i\rangle$, eigenstate of position, is a delta function

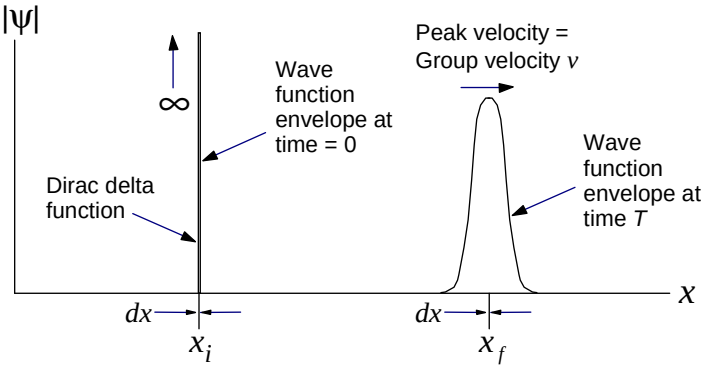


Figure 18-1. Propagation of a Position Eigenstate Quantum Wave

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 Fig. 18-1. 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 Fig. 18-1.) 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 (18-9).

It spreads as it evolves

When measured at peak x_f, wave packet collapses to $|x_f\rangle$, eigenstate of position, i.e., delta function

So U for position eigenstate at x_f, $\to |U|^2$ = probability density at packet peak x_f

We can re-write (18-9), in wave mechanics notation as

$$U\left(x_i,x_f;T\right)=\int_{-\infty}^{+\infty}\delta\left(x-x_f\right)\psi\left(x,T\right)dx=\psi\left(x_f,T\right). \tag{18-10}$$

Thus,

$$\left|U(x_i, x_f; T)\right|^2 = \psi^*(x_f, T)\psi(x_f, T) = \text{probability density at } x_f \,. \tag{18-11}$$

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 (18-5).

As we will see, the value found using the RHS of (18-9), i.e., that of the Schrödinger approach, is the same as the value found using Feynman's many paths approach.

18.4 Expressing the Wave Function Peak in Terms of the Lagrangian

18.4.1 Background

Path integral approach expresses wave function peak in terms of Lagrangian

One of Feynman's assumptions for his path integral approach to NRQM, RQM (relativistic quantum mechanics), and QFT was to express the wave function value at the peak of a wave packet (see Fig. 18-1) in terms of the Lagrangian (exact relation shown at the end of this section 18.4). 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 rationale 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.

18.4.2 Deducing Feynman's Phase Peak Relationship

Heuristic way to deduce $\psi_{pezk} = Ae^{i\int \frac{L}{\hbar} dt}$

The Simplified, Heuristic Argument

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

$$\psi = Ae^{-i(Et - \mathbf{p}\cdot\mathbf{x})/\hbar} \,, \tag{18-12}$$

means the phase angle, at any given **x** and *t*, is

$$\phi = -(Et - \mathbf{p}\cdot\mathbf{x})/\hbar \quad . \tag{18-13}$$

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 that 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 **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

$$\frac{d\phi}{dt} = \frac{-(E - \mathbf{p}\cdot\mathbf{v})}{\hbar} = \frac{-T - V + \mathbf{p}\cdot\mathbf{v}}{\hbar} \,, \tag{18-14}$$

where **v** is the velocity of the wave peak, T is kinetic energy, and V is potential energy. Non-relativistically,

$$T = \tfrac{1}{2}mv^2 \qquad \mathbf{p} = m\mathbf{v} \quad \rightarrow \quad \mathbf{p}\bullet\mathbf{v} = 2T \,, \tag{18-15}$$

so, in terms of the classical Lagrangian L, (18-14) becomes

$$\frac{d\phi}{dt} = \frac{T - V}{\hbar} = \frac{L}{\hbar} \,. \tag{18-16}$$

More formally, using the Legendre transformation

$$H = p_i\dot{q}_i - L \quad (E = \mathbf{p}\cdot\mathbf{v} - L \quad \text{here}) \,, \tag{18-17}$$

directly in (18-14), we get (18-16).

Thus, from (18-16), the phase difference between two events the particle traverses is

$$\phi = \int \frac{L}{\hbar} dt = \frac{S}{\hbar} \ , \tag{18-18}$$

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 (18-18) is an integral of type 2 in Wholeness Chart 18-1.

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

$$\psi_{peak} = Ae^{i\int \frac{L}{\hbar} dt} = Ae^{i\frac{S}{\hbar}} \tag{18-19}$$

This is the typical starting point assumption when teaching the Feynman path integral approach (still to be developed beginning in Section 18.5.)

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

More Precise Argument

The precise expression for a QM particle wave packet[1], where overbars designate expectation (classical) values; v_g, the group (peak, classical) velocity; and $g(p)$, the momentum space distribution is

More formal way to deduce $\psi_{peak} = Ae^{i\int \frac{L}{\hbar} dt}$

$$\psi(x,t) = e^{-\frac{i}{\hbar}\left(\bar{E}t - \bar{p}x\right)} \underbrace{\frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty} \underbrace{g(p)}_{\text{real}} \underbrace{e^{-\frac{i}{\hbar}(v_g t - x)(p-\bar{p})}}_{\substack{=1 \text{ for } x=\text{ peak,} \\ \text{i.e., for } x=v_g t}} \underbrace{e^{-\frac{i}{2\hbar}\frac{t}{m}(p-\bar{p})^2}}_{\substack{\text{time depend} \\ \& \text{complex}}} dp}_{\text{call this } A(t) \text{ for } x=x_{peak}} . \tag{18-20}$$

We are interested in the value of (18-20) at the peak, $\psi(x_{peak},t)$, where $x_{peak} = v_g t$. To begin, note that with $x=x_{\text{peak}}$ 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 (18-20) 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

$$\psi(x_{peak},t) = A(t)e^{-\frac{i}{\hbar}\left(\bar{E}t - \bar{p}x_{peak}\right)} . \tag{18-21}$$

Except for the time dependence in $A(t)$, this is equivalent to (18-12), as the expectation values for E and p equal the classical values for the particle. So, with regard to the exponent factor in (18-21), all of the logic from (18-13) through (18-19) 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 (18-10).

$$\boxed{\psi\left(x_{peak},T\right) = A(T)e^{i\int_0^T \frac{L}{\hbar} dt} = A(T)e^{i\frac{S}{\hbar}} = U(x_i,x_f,T)} \tag{18-22}$$

$\psi_{peak}(T) = U(x_i, x_f, T)$

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.

18.5 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.

Feynman's idea: particle travels all paths in spacetime simultaneously

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 Schrödinger with a single wave. The two different approaches are equivalent.

---

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

Superimposing all $e^{i\int \frac{L}{\hbar}dt}$ from all those paths, gives same |U| as other methods

Definition: Feynman's method is called the "path integral", "functional 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, **F** = m**a**, 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.

18.6 Superimposing a Finite Number of Paths[1]

18.6.1 The Rotating Phasor

The phasor of (18-22) 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 $L/\hbar$, 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).

Visualizing Feynman's rotating phasor

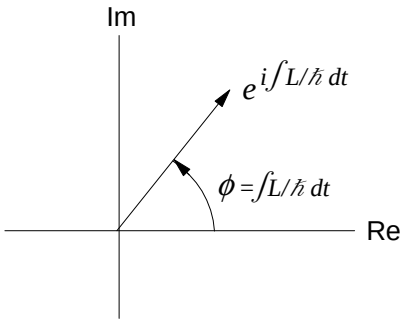


Figure 18-2. Rotating Phasor

For the purposes of Feynman's approach, we can consider the particle as a wave packet with phase at the peak determined by (18-22), 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.

18.6.2 Several Paths Graphically

Fig. 24 in Feynman's book *QED: The Strange Theory of Light and Matter*[2], 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. 18-3 for an electron rather than a photon. In Fig. 18-3 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.

Graphical visualization of one example of many paths superposition

We look at a representative 15 different paths for the electron, out of 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 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.

---

[1] 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 Eduction, Universiteit van Amsterdam.

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

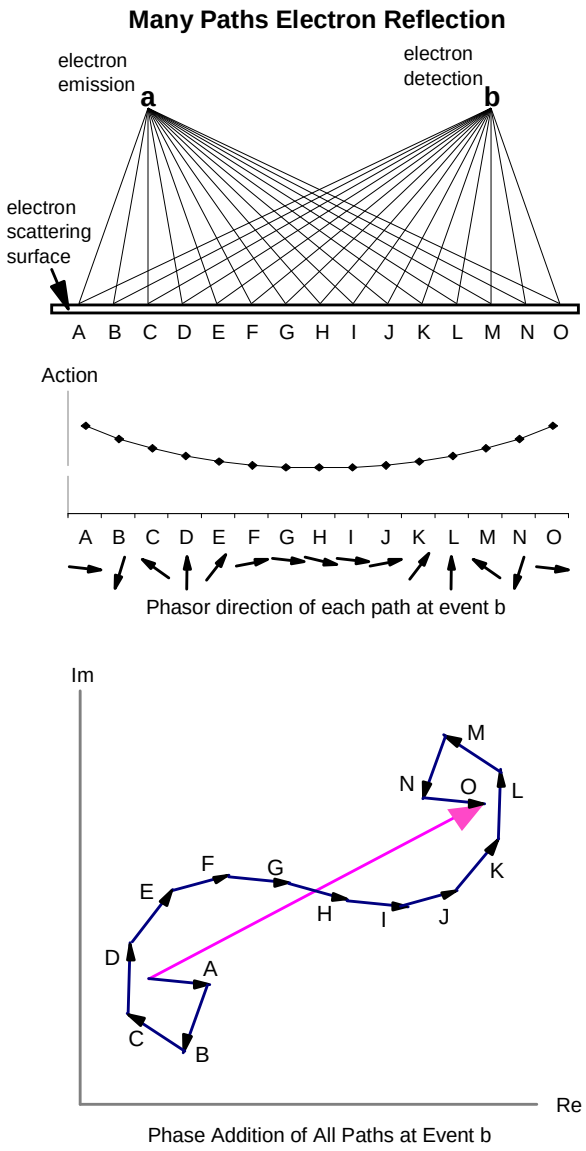


Figure 18-3. Graphical Justification for Many Paths Approach

From (18-22) and Fig. 18-2, we can determine the phasor $e^{iS/\hbar}$ of (18-22) for the particle/wave arriving at event b, for each path, where $S = LT = \frac{1}{2}\, mv^2T$. The phasor direction in complex space for each path at the detection event b is depicted in the middle of Fig. 18-3.

The bottom part of Fig. 18-3 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 greater contributions to the final sum. A similar effect would occur if the value for Planck's constant were smaller. As $\hbar \to 0$, all paths but H would tend to cancel out.

Clarification

I once thought 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 (18-22).) 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.

Magnitude of final sum over all paths is important thing. Total phase not relevant to probability

Note that in order to get a graphically significant Fig. 18-3, 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 illumined, 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 from the website for this book[1].

End of Clarification

[1] www.quantumfieldtheory.info/manypathsgraphicelectronreflection.xls

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.

18.6.3 Many Paths Mathematically

Math expression of our example of many paths superposition

Consider particle paths similar to those of Fig. 18-3, where the wave function peak for path number 1, with $A_1(T)$ as in (18-22), as

$$\psi_1^{peak} = A_1(T)e^{iS_1/\hbar}. \tag{18-23}$$

In the spirit of the prior section, one considers the phasor of (18-23) *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

$$e^{iS_1/\hbar} + e^{iS_2/\hbar} + e^{iS_3/\hbar} + ... = A_b e^{i\phi_{\text{sum}}} \tag{18-24}$$

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,

$$\lim_{N\to\infty}\sum_{j=1}^{N} e^{iS_j/\hbar} = A_b e^{i\phi_{\text{sum}}} \propto U(x_i, x_f, T) \qquad \left|A_b e^{i\phi_{\text{sum}}}\right|^2 = \left|A_b\right|^2 \propto \left|U\right|^2 \text{ (probability density)} \tag{18-25}$$

We will learn how to evaluate the limit in (18-25).

18.6.4 Another Example

Another example: double slit experiment

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.

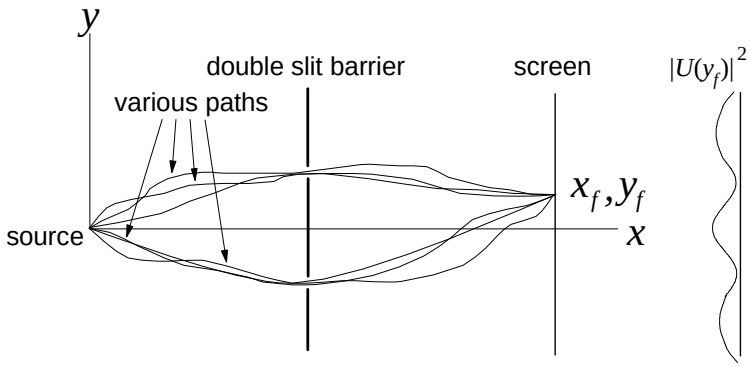


Figure 18-4. Double Slit Experiment in Many Paths Approach

By the Schrödinger 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.

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 Fig. 18-4. The result would be proportional to the amplitude at the spot found in the Schrödinger approach. That is, the sum of all phasors at x_f,y_f (see (18-25)) yields

$$C\lim_{N\to\infty}\sum_{j=1}^{N} e^{iS_j/\hbar} = U(x_i, y_i; x_f, y_f; T), \tag{18-26}$$

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 $U(y_f)$.

18.6.5 Finding the Proportionality Constant: By Example

Feynman result of summation is only proportional to $|U|^2$. Need to find proportionality constant another way.

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.,

$$\int_{y_f=-\infty}^{y_f=+\infty} \left| C \lim_{N\to\infty} \sum_{j=1}^{N} e^{iS_j/\hbar} \right|^2 dy_f = \int_{y_f=-\infty}^{y_f=+\infty} \left| U(y_f) \right|^2 dy_f = 1, \tag{18-27}$$

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

18.7 Summary of Approaches

18.7.1 Feynman's Postulates

Path integral three starting 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.

Phasors are not weighted when summing them

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 (18-24) and Fig. 18-3). 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 (18-19), 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[1].

18.7.2 Comparison of Approaches to QM

Comparing 3 equivalent approaches to NRQM

Wholeness Chart 18-2 summarizes the major similarities and differences between alternative approaches to NRQM.

[1] 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 18.6) 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 have not personally investigated this and do not know for sure.

Wholeness Chart 18-2. Equivalent Approaches to Non-relativistic Quantum Mechanics

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

18.8 Finite Sums to Functional Integrals

18.8.1 Time Slicing: The Concept

Slicing time into "pieces" for discrete time analysis

A simple example

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. 18-3 and (18-24), 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. 18-5 where, for convenience, we plot time vertically and space horizontally. As opposed to our spatially 2D example in Fig. 18-3, different paths in Fig. 18-5 actually refer to the particle traveling along the *x* axis only between *i* and *f*, though at varying (both positive and negative) 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

$$\text{phasor at } \mathbf{f} = \underbrace{e^{i\int_{t_i}^{t_f} \frac{L}{\hbar} dt} = e^{iS/\hbar}}_{\text{one path}} , \tag{18-28}$$

The amplitude *U* for the transition from **i** to **f** is proportional to the sum of (18-28) for all paths,

$$\text{sum of } \infty \text{ phasors at } \mathbf{f} = \lim_{N\to\infty} \sum_{j=1}^{N} e^{iS_j/\hbar} \quad . \tag{18-29}$$

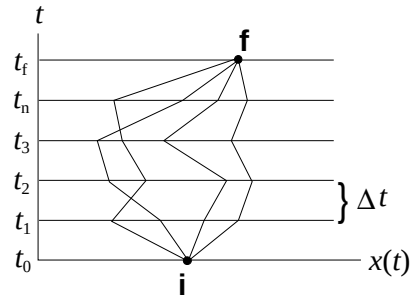


Figure 18-5. Time Slicing for Finite Number of Paths

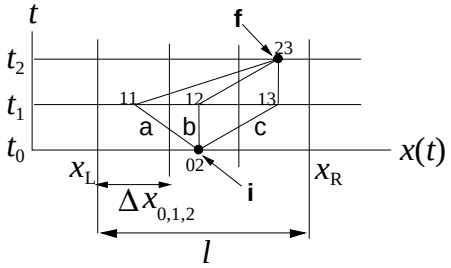


Figure 18-6. Space Slicing for Three Discrete Paths

18.8.2 Space Slicing: Simple Paths with Discrete Approximation

SEE THE TEXT FOR THE REMAINDER OF THIS CHAPTER.

H Operator on Non Eigen States & Wave Packets

Version date November 21, 2011

Robert D. Klauber
www.quantumfieldtheory.info

Special note: The material (a bit later into this document) on non-eigenstates and wave packets was developed solely by the author, who was unable to find such material in the literature (though it probably exists somewhere.) Its accuracy has not been checked by others.

The following Wholeness Chart was completed on November 21, 2011. The material following it was done as of March 13, 2010, and was never fully completed (although almost). That material may have some errors in it, i.e., it may not exactly parallel the chart below (although it should be close.)

Wholeness Chart 10-3. Discrete vs Continuous Versions of QFT

(Only Scalars Shown)

	Discrete	**Continuous**
Field Equations Solutions	$\phi(x) = \sum_{\mathbf{k}} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}} \left(a(\mathbf{k})e^{-ikx} + b^\dagger(\mathbf{k})e^{ikx}\right)$ $\phi^\dagger(x) = \sum_{\mathbf{k}} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}} \left(b(\mathbf{k})e^{-ikx} + a^\dagger(\mathbf{k})e^{ikx}\right)$	$\phi(x) = \int \frac{d^3k}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}}}} \left(a(\mathbf{k})e^{-ikx} + b^\dagger(\mathbf{k})e^{ikx}\right)$ $\phi^\dagger(x) = \int \frac{d^3k}{\sqrt{2(2\pi)^3\omega_{\mathbf{k}}}} \left(b(\mathbf{k})e^{-ikx} + a^\dagger(\mathbf{k})e^{ikx}\right)$
Coefficient commutators	$\left[a(\mathbf{k}), a^\dagger(\mathbf{k}')\right] = \left[b(\mathbf{k}), b^\dagger(\mathbf{k}')\right] = \delta_{\mathbf{kk}'}$	$\left[a(\mathbf{k}), a^\dagger(\mathbf{k}')\right] = \left[b(\mathbf{k}), b^\dagger(\mathbf{k}')\right] = \delta(\mathbf{k}-\mathbf{k}')$
$\mathcal{H}_0^0$	$\pi_0^0\dot{\phi} + \pi_0^{0\dagger}\dot{\phi}^\dagger - \mathcal{L}_0^0 = \left(\dot{\phi}\dot{\phi}^\dagger + \nabla\phi^\dagger \cdot \nabla\phi + \mu^2\phi^\dagger\phi\right)$	as at left
H_0^0	$\sum_{\mathbf{k}} \omega_{\mathbf{k}} \left(N_a(\mathbf{k}) + \frac{1}{2} + N_b(\mathbf{k}) + \frac{1}{2}\right)$ $N_a(\mathbf{k}) = a^\dagger(\mathbf{k})\, a(\mathbf{k})$, $N_b(\mathbf{k}) = b^\dagger(\mathbf{k})\, b(\mathbf{k})$	$\int \omega_{\mathbf{k}} \left(N_a(\mathbf{k}) + \frac{1}{2} + N_b(\mathbf{k}) + \frac{1}{2}\right) d^3k$ $N_a(\mathbf{k}) = a^\dagger(\mathbf{k})\, a(\mathbf{k})$, $N_b(\mathbf{k}) = b^\dagger(\mathbf{k})\, b(\mathbf{k})$
Operator Units	$N_a(\mathbf{k})$, number of real particles, unitless, M^0 ½ , number of vacuum particles, unitless $a(\mathbf{k})$, $a^\dagger(\mathbf{k})$, unitless Similar for $N_b(\mathbf{k})$, $b(\mathbf{k})$, $b^\dagger(\mathbf{k})$,	$N_a(\mathbf{k})$, (num real particles)/(**k** space vol), M^{-3} ½ , (num vacuum particles)/(**k** space vol), M^{-3} $a(\mathbf{k})$, $a^\dagger(\mathbf{k})$, $M^{-3/2}$ Similar for $N_b(\mathbf{k})$, $b(\mathbf{k})$, $b^\dagger(\mathbf{k})$,
Single Particle State Relations (Only particles, not anti-particles shown)		
Eigenstate Creation	$a^\dagger(\mathbf{k})\lvert 0\rangle = \lvert\phi_{\mathbf{k}}\rangle = \left\lvert \frac{e^{-ikx}}{\sqrt{V}} \right\rangle$ Eigenstate at one point in **k** space, spread over volume V in **x** space.	$a^\dagger(\mathbf{k})\lvert 0\rangle = \lvert\phi(\mathbf{k})\rangle = \left\lvert \frac{e^{-ikx}}{\sqrt{(2\pi)^3}} \right\rangle$ Eigenstate at one point in **k** space, spread over universe in **x** space. $V \to \infty$

General State Creation, ***C* is General State Creation Operator**	$\|\phi\rangle = C\|0\rangle = \sum_{\mathbf{k}} A_{\mathbf{k}} a^{\dagger}(\mathbf{k})\|0\rangle$ $= \left\| \frac{1}{\sqrt{V}} \left(A_{\mathbf{k}_1} e^{-ik_1 x} + A_{\mathbf{k}_2} e^{-ik_2 x} + .. \right) \right\rangle = \left\| \sum_{\mathbf{k}} A_{\mathbf{k}} \frac{e^{-ikx}}{\sqrt{V}} \right\rangle$ Coefficient $A_{\mathbf{k}}$ unitless. State units $l^{-3/2} = M^{3/2}$	$\|\phi\rangle = C\|0\rangle = \int A(\mathbf{k}) a^{\dagger}(\mathbf{k}) d^3k \|0\rangle$ $= \left\| \int A(\mathbf{k}) \frac{e^{-ikx}}{\sqrt{(2\pi)^3}} d^3k \right\rangle$ Coefficient $A(\mathbf{k})$ units $l^{3/2} = M^{-3/2}$. State units $l^{-3/2} = M^{3/2}$
Creation Operator *C*	$C = \sum_{\mathbf{k}} A_{\mathbf{k}} a^{\dagger}(\mathbf{k})$	$C = \int A(\mathbf{k}) a^{\dagger}(\mathbf{k}) d^3k$
State Norms	$\langle \phi \| \phi \rangle = 1$	$\langle \phi \| \phi \rangle = 1$
Coefficient Properties	$\sum_{\mathbf{k}} \|A_{\mathbf{k}}\|^2 = 1$	$\int \|A(\mathbf{k})\|^2 d^3k = 1$
For an Eigenstate	Only one $A_{\mathbf{k}}$, with $\|A_{\mathbf{k}}\| = 1$. $\|A_{\mathbf{k}}\| / \sqrt{V} \to 1/\sqrt{V}$	Not very meaningful.
***N_a* (k) Acting on General State**	$N_a(\mathbf{k}) \left\| \sum_{\mathbf{k}'} A_{\mathbf{k}'} \frac{e^{-ik'x}}{\sqrt{V}} \right\rangle$ $= \|A_{\mathbf{k}'}\|^2 \delta_{\mathbf{k}\mathbf{k}'} \left\| \sum_{\mathbf{k}'} A_{\mathbf{k}'} \frac{e^{-ik'x}}{\sqrt{V}} \right\rangle = \|A_{\mathbf{k}}\|^2 \left\| \sum_{\mathbf{k}'} A_{\mathbf{k}'} \frac{e^{-ik'x}}{\sqrt{V}} \right\rangle$	$N_a(\mathbf{k}) \left\| \int A(\mathbf{k}') \frac{e^{-ik'x}}{\sqrt{(2\pi)^3}} d^3k' \right\rangle$ $= \|A(\mathbf{k}')\|^2 \delta(\mathbf{k} - \mathbf{k}') \left\| \int A(\mathbf{k}') \frac{e^{-ik'x}}{\sqrt{(2\pi)^3}} d^3k' \right\rangle$
Eigenstate Energy Expectation Value	$\bar{E} = \langle \phi_{\mathbf{k}'} \| \sum_{\mathbf{k}} \omega_{\mathbf{k}} \left(N_a(\mathbf{k}) + \frac{1}{2} + N_b(\mathbf{k}) + \frac{1}{2} \right) \| \phi_{\mathbf{k}'} \rangle$ $= \left(\omega_{\mathbf{k}'} + \frac{1}{2} \sum_{\mathbf{k}} \omega_{\mathbf{k}} + \frac{1}{2} \sum_{\mathbf{k}} \omega_{\mathbf{k}} \right) \langle \phi_{\mathbf{k}'} \| \phi_{\mathbf{k}'} \rangle$ $= \omega_{\mathbf{k}'} + \sum_{\mathbf{k}} \omega_{\mathbf{k}}$	$\bar{E} = \langle \phi(\mathbf{k}') \| \int \omega_{\mathbf{k}} (N_a(\mathbf{k}) + \frac{1}{2} + N_b(\mathbf{k}) + \frac{1}{2}) d^3k \| \phi(\mathbf{k}') \rangle$ $= \int \omega_{\mathbf{k}} \left(\delta(\mathbf{k} - \mathbf{k}') + \frac{1}{2} + \frac{1}{2} \right) d^3k \langle \phi \| \phi \rangle$ $= \omega_{\mathbf{k}'} + \int \omega_{\mathbf{k}} d^3k$
General State Energy Expectation Value	$\bar{E} = \left\langle \sum_{\mathbf{k}'} A^{\dagger}_{\mathbf{k}'} \frac{e^{ik'x}}{\sqrt{V}} \right\| \sum_{\mathbf{k}} \omega_{\mathbf{k}} (N_a(\mathbf{k}) + \frac{1}{2} + N_b(\mathbf{k}) + \frac{1}{2}) \left\| \sum_{\mathbf{k}'} A_{\mathbf{k}'} \frac{e^{-ik'x}}{\sqrt{V}} \right\rangle$ $= \left(\sum_{\mathbf{k}'} \omega_{\mathbf{k}'} A^{\dagger}_{\mathbf{k}'} A_{\mathbf{k}'} \delta_{\mathbf{k}\mathbf{k}'} + \frac{1}{2} \sum_{\mathbf{k}} \omega_{\mathbf{k}} + \frac{1}{2} \sum_{\mathbf{k}} \omega_{\mathbf{k}} \right) \langle \phi \| \phi \rangle$ $= \sum_{\mathbf{k}} \|A_{\mathbf{k}}\|^2 \omega_{\mathbf{k}} + \sum_{\mathbf{k}} \omega_{\mathbf{k}} = \bar{\omega} + \sum_{\mathbf{k}} \omega_{\mathbf{k}}$	$\bar{E} = \left\langle \int A(\mathbf{k}') \frac{e^{-ik'x}}{\sqrt{(2\pi)^3}} d^3k' \right\| \int \omega_{\mathbf{k}} (N_a(\mathbf{k}) + \frac{1}{2} + N_b(\mathbf{k}) + \frac{1}{2}) d^3k \left\| \int A(\mathbf{k}') \frac{e^{-ik'x}}{\sqrt{(2\pi)^3}} d^3k' \right\rangle$ $= \int \omega_{\mathbf{k}} \left(\|A(\mathbf{k}')\|^2 \delta(\mathbf{k} - \mathbf{k}') + \frac{1}{2} + \frac{1}{2} \right) d^3k \langle \phi \| \phi \rangle$ $= \int \|A(\mathbf{k})\|^2 \omega_{\mathbf{k}} d^3k + \int \omega_{\mathbf{k}} d^3k = \bar{\omega} + \int \omega_{\mathbf{k}} d^3k$
Multi-particle State Relations (Only particles, not anti-particles shown)		
Multi Eigen Particles Creation	$a^{\dagger}(\mathbf{k}_1) a^{\dagger}(\mathbf{k}_2) ... \|0\rangle = \left\| \phi_{\mathbf{k}_1}, \phi_{\mathbf{k}_2}, .. \right\rangle$	$a^{\dagger}(\mathbf{k}_1) a^{\dagger}(\mathbf{k}_2) ... \|0\rangle = \left\| \phi(\mathbf{k}_1), \phi(\mathbf{k}_2), .. \right\rangle$

Multi General Particles Creation, C_q is qth Particle Creation Operator	$\left\|\phi_q,\phi_r,...\right\rangle = \left(C_q C_r..\right)\left\|0\right\rangle$ $= \left(\sum_{\mathbf{k}} A_{q\mathbf{k}} a^\dagger(\mathbf{k})\right)\left(\sum_{\mathbf{k}} A_{r\mathbf{k}} a^\dagger(\mathbf{k})\right)(...)\left\|0\right\rangle$ $= \left\|\sum_{\mathbf{k}} A_{q\mathbf{k}} \frac{e^{-ikx_q}}{\sqrt{V}}, \sum_{\mathbf{k}} A_{r\mathbf{k}} \frac{e^{-ikx_r}}{\sqrt{V}}, ..\right\rangle$	$\left\|\phi_q,\phi_r,...\right\rangle = \left(C_q C_r..\right)\left\|0\right\rangle$ $= \left(\int A_q(\mathbf{k}) a^\dagger(\mathbf{k}) d^3k\right)\left(\int A_r(\mathbf{k}) a^\dagger(\mathbf{k}) d^3k\right)(..)\left\|0\right\rangle$ $= \left\|\int A_q(\mathbf{k}) \frac{e^{-ikx_q}}{\sqrt{(2\pi)^3}} d^3k, \int A_r(\mathbf{k}) \frac{e^{-ikx_r}}{\sqrt{(2\pi)^3}} d^3k, ..\right\rangle$
Normalized Operator C_q	$C_q = \sum_{\mathbf{k}} A_{q\mathbf{k}} a^\dagger(\mathbf{k})$	$C_q = \int A_q(\mathbf{k}) a^\dagger(\mathbf{k}) d^3k$
State Norms	$\left\langle\phi\middle|\phi\right\rangle = \left\langle\phi_q,\phi_r,...\middle|\phi_q,\phi_r,...\right\rangle = 1$	$\left\langle\phi\middle|\phi\right\rangle = \left\langle\phi_q,\phi_r,...\middle|\phi_q,\phi_r,...\right\rangle = 1$
Coefficient Properties	$\sum_{\mathbf{k}} \left\|A_{q\mathbf{k}}\right\|^2 = 1, \quad \sum_{\mathbf{k}} \left\|A_{r\mathbf{k}}\right\|^2 = 1,$ etc.	$\int \left\|A_q(\mathbf{k})\right\|^2 d^3k = 1, \quad \int \left\|A_r(\mathbf{k})\right\|^2 d^3k = 1,$ etc.
N_a (k) Acting on Multi General Particles State	$N_a(\mathbf{k})\left\|\phi_q, 2\phi_r, ..\right\rangle$ $= N_a(\mathbf{k})\left\|\sum_{\mathbf{k}'} A_{q\mathbf{k}'} \frac{e^{-ik'x_q}}{\sqrt{V}}, 2\sum_{\mathbf{k}'} A_{r\mathbf{k}'} \frac{e^{-ik'x_r}}{\sqrt{V}}, ..\right\rangle$ $= \left(\left\|A_{q\mathbf{k}'}\right\|^2 \delta_{\mathbf{k}\mathbf{k}'} + 2\left\|A_{r\mathbf{k}'}\right\|^2 \delta_{\mathbf{k}\mathbf{k}'} + ..\right)\left\|\phi_r, 2\phi_r, ..\right\rangle$ $= \left(\left\|A_{q\mathbf{k}}\right\|^2 + 2\left\|A_{r\mathbf{k}}\right\|^2 + ..\right)\left\|\phi_r, 2\phi_r, ..\right\rangle$	$N_a(\mathbf{k})\left\|\phi_q, 2\phi_r, ..\right\rangle = N_a(\mathbf{k})\times$ $\left\|\int A_q(\mathbf{k}') \frac{e^{-ik'x_q}}{\sqrt{(2\pi)^3}} d^3k', 2\int A_r(\mathbf{k}') \frac{e^{-ik'x_r}}{\sqrt{(2\pi)^3}} d^3k', ..\right\rangle$ $= (\left\|A_q(\mathbf{k}')\right\|^2 \delta(\mathbf{k}-\mathbf{k}')$ $+ 2\left\|A_r(\mathbf{k}')\right\|^2 \delta(\mathbf{k}-\mathbf{k}') + ..)\left\|\phi_r, 2\phi_r, ..\right\rangle$
Multi Eigen Particles Energy Expectation Value	$\bar{E} = \left\langle\phi_{\mathbf{k}_1}, 2\phi_{\mathbf{k}_2}, ..\middle|\sum_{\mathbf{k}} \omega_{\mathbf{k}} (N_a(\mathbf{k}) + \frac{1}{2} + N_b(\mathbf{k}) + \frac{1}{2})\middle|\phi_{\mathbf{k}_1}, 2\phi_{\mathbf{k}_2}, ..\right\rangle$ $= (\omega_{\mathbf{k}_1} + 2\omega_{\mathbf{k}_2} + .. + \frac{1}{2}\sum_{\mathbf{k}} \omega_{\mathbf{k}} + \frac{1}{2}\sum_{\mathbf{k}} \omega_{\mathbf{k}})\left\langle\phi_{\mathbf{k}_1}, 2\phi_{\mathbf{k}_2} ..\middle|\phi_{\mathbf{k}_1}, 2\phi_{\mathbf{k}_2}, ..\right\rangle$ $= \omega_{\mathbf{k}_1} + 2\omega_{\mathbf{k}_2} + ... + \sum_{\mathbf{k}} \omega_{\mathbf{k}}$	$\bar{E} = \left\langle\phi(\mathbf{k}_1), 2\phi(\mathbf{k}_2), ..\middle|\int \omega_{\mathbf{k}} (N_a(\mathbf{k}) + \frac{1}{2} + N_b(\mathbf{k}) + \frac{1}{2}) d^3k\middle|\phi(\mathbf{k}_1), 2\phi(\mathbf{k}_2), ..\right\rangle$ $= (\omega_{\mathbf{k}_1} + 2\omega_{\mathbf{k}_2} + ... + \frac{1}{2}\int \omega_{\mathbf{k}} d^3k + \frac{1}{2}\int \omega_{\mathbf{k}} d^3k)\left\langle\phi(\mathbf{k}_1), 2\phi(\mathbf{k}_2), ..\middle|\phi(\mathbf{k}_1), 2\phi(\mathbf{k}_2), ..\right\rangle$ $= \omega_{\mathbf{k}_1} + 2\omega_{\mathbf{k}_2} + ... + \int \omega_{\mathbf{k}} d^3k$
Multi General Particles Energy Expectation Value	$\bar{E} = \left\langle\phi_q, 2\phi_r, ..\middle|\sum_{\mathbf{k}} \omega_{\mathbf{k}} (N_a(\mathbf{k}) + \frac{1}{2} + N_b(\mathbf{k}) + \frac{1}{2})\middle|\phi_q, 2\phi_r, ..\right\rangle$ $= (\sum_{\mathbf{k}} \left\|A_{q\mathbf{k}}\right\|^2 \omega_{\mathbf{k}} + 2\sum_{\mathbf{k}} \left\|A_{r\mathbf{k}}\right\|^2 \omega_{\mathbf{k}} + ... + \frac{1}{2}\sum_{\mathbf{k}} \omega_{\mathbf{k}} + \frac{1}{2}\sum_{\mathbf{k}} \omega_{\mathbf{k}})\left\langle\phi_q, 2\phi_r, ..\middle|\phi_q, 2\phi_r, ..\right\rangle$ $= \bar{\omega}_q + \bar{\omega}_r + ... + \sum_{\mathbf{k}} \omega_{\mathbf{k}}$	$\bar{E} = \left\langle\phi_q, 2\phi_r, ..\middle|\int \omega_{\mathbf{k}} (N_a(\mathbf{k}) + \frac{1}{2} + N_b(\mathbf{k}) + \frac{1}{2}) d^3k\middle|\phi_q, 2\phi_r, ..\right\rangle$ $= (\int \omega_{\mathbf{k}} \left\|A_q(\mathbf{k})\right\|^2 d^3k + 2\int \omega_{\mathbf{k}} \left\|A_r(\mathbf{k})\right\|^2 d^3k + ... + \frac{1}{2}\int \omega_{\mathbf{k}} d^3k + \frac{1}{2}\int \omega_{\mathbf{k}} d^3k)\left\langle\phi_q, 2\phi_r, ..\middle|\phi_q, 2\phi_r, ..\right\rangle$ $= \bar{\omega}_q + 2\bar{\omega}_r + ... + \int \omega_{\mathbf{k}} d^3k$
Note	In the energy expectation derivation for the continuous case, one finds a delta function squared in the vacuum energy part. This is undefined mathematically. By some perspectives, its evaluation leaves a vacuum term of energy $\omega(\mathbf{k}=0)$ which equals μ (one particle mass). An alternative perspective is shown above.	

1. Solutions to Free Field Equations

For a scalar, we have

1.1 Free field solutions

$$\begin{aligned}\phi &= \phi^{+} + \phi^{-} \\ \phi^{\dagger} &= \phi^{\dagger +} + \phi^{\dagger -}\end{aligned} \tag{1}$$

1.1.1 Discrete eigenstates (finite volume B.C.'s or periodic B.C.'s)

$$\begin{aligned}\phi(x) &= \sum_{\mathbf{k}} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}}\left[a(\mathbf{k})e^{-ikx} + b^{\dagger}(\mathbf{k})e^{ikx}\right] \\ \phi^{\dagger}(x) &= \sum_{\mathbf{k}} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}}\left[b(\mathbf{k})e^{-ikx} + a^{\dagger}(\mathbf{k})e^{ikx}\right]\end{aligned} \tag{2}$$

where the summation is from infinite **k** in the negative x direction to infinite **k** in the positive x direction plus similar summations for the y and z directions.

1.1.2 Continuous eigenstates (no B.C.'s over all space)

$$\begin{aligned}\phi(x) &= \frac{1}{\sqrt{(2\pi)^3}}\int \frac{d\mathbf{k}}{\sqrt{2\omega_{\mathbf{k}}}}\left[a(\mathbf{k})e^{-ikx} + b^{\dagger}(\mathbf{k})e^{ikx}\right] \\ \phi^{\dagger}(x) &= \frac{1}{\sqrt{(2\pi)^3}}\int \frac{d\mathbf{k}}{\sqrt{2\omega_{\mathbf{k}}}}\left[b(\mathbf{k})e^{-ikx} + a^{\dagger}(\mathbf{k})e^{ikx}\right]\end{aligned} \tag{3}$$

where the integration ranges as the summation in (2) except that here **k** is a continuous variable.

2. Relativistic Quantum Mechanics (RQM)

In relativistic quantum mechanics (RQM) $\phi(x)$ of (2) represents a single particle general state that is a sum of discrete momentum eigenstates of that single particle. The coefficients $a(\mathbf{k})$ and $b^{\dagger}(\mathbf{k})$ are numbers, amplitudes which, when squared, equal the probability of finding the single particle in that discrete eigenstate.

In RQM, $\phi(x)$ of (3) represents a single particle wave packet comprising an integral over momentum eigenstates that are continuous. The coefficients $a(\mathbf{k})$ and $b^{\dagger}(\mathbf{k})$ are numbers which represent the Fourier transform amplitudes of the eigenstates in the continuous momentum space.

In quantum field theory (QFT) these coefficients are not numbers but operators that each create or destroy single particle eigenstates. Commonly in QFT one employs one term in (2) to create or destroy a single particle discrete momentum eigenstate (having no uncertainty in its momentum, but infinite uncertainty in its spatial location.)

2.1 Discrete Solutions

The solutions (2) in RQM are single particle general (sum of eigenstates) states, not operators, of form (where we substitute numerical $A_{\mathbf{k}}$ in RQM for operator $a(\mathbf{k})$ in QFT, etc.)

$$|\phi\rangle = \sum_{\mathbf{k}} \frac{A_{\mathbf{k}}}{\sqrt{2\omega_{\mathbf{k}}}}\frac{e^{-ikx}}{\sqrt{V}} = \sum_{\mathbf{k}} \frac{A_{\mathbf{k}}}{\sqrt{2\omega_{\mathbf{k}}}}|\phi_{\mathbf{k}}\rangle \tag{4}$$

where $|\phi_{\mathbf{k}}\rangle$ has unit norm. That is,

$$|\phi_{\mathbf{k}}\rangle = \frac{e^{-ikx}}{\sqrt{V}}, \tag{5}$$

so that

$$\langle \phi_{\mathbf{k}} | \phi_{\mathbf{k}} \rangle = \frac{1}{V} \int e^{ikx} e^{-ikx} d^3x = 1 , \tag{6}$$

or more generally,

$$\langle \phi_{\mathbf{k}} | \phi_{\mathbf{k}'} \rangle = \frac{1}{V} \int e^{ikx} e^{-ikx'} d^3x = \delta_{\mathbf{k}\mathbf{k}'} . \tag{7}$$

2.1.1 Probability for Discrete Solutions

For a single particle state in RQM the probability density[1] is

$$\rho = i\left(\langle \phi | \phi_{x0} \rangle_{n.i.} - \langle \phi_{x0} | \phi \rangle_{n.i.} \right) = \left(2\sum_{\mathbf{k}} \frac{A_{\mathbf{k}}^{\dagger}}{\sqrt{2\omega_{\mathbf{k}}}} \frac{e^{ikx}}{\sqrt{V}} \right) \left(\sum_{\mathbf{k}'} \frac{\omega_{\mathbf{k}'} A_{\mathbf{k}'}}{\sqrt{2\omega_{\mathbf{k}'}}} \frac{e^{-ik'x}}{\sqrt{V}} \right) \tag{8}$$

where the subscript "*n.i.*" implies we are not integrating over space inside the bracket. When we do integrate, using the Kronecker delta function relation of (7), we get

$$\int \rho d^3x = \sum_{\mathbf{k}} | A_{\mathbf{k}} |^2 = 1 , \tag{9}$$

where $|A_{\mathbf{k}}|^2$ is the probability of measuring the **k**th eigenstate.

Note that this is the reason for the normalization factors $1/\sqrt{2\omega_{\mathbf{k}} V}$ used in (2). Those factors result in a total probability of one for a single particle and $|A_{\mathbf{k}}|^2$ as the probability for measuring the **k**th state. That is, the form of the relativistic field equation gave us the form of the probability density in the middle of (8). (See footnote 1.) The time derivatives in (8) gave us a factor of $\omega_{\mathbf{k}}$, and the two terms a factor of 2. These cancel in (9) with the $2\omega_{\mathbf{k}}$ in the denominators of the terms in (2). The *V* term in the denominator cancels in the integration over volume in (9) and the result is a total probability of 1.

This probability value of unity is a relativistic invariant. If we change our frame, the energy spectrum (i.e., the $\omega_{\mathbf{k}}$ values) will change (K.E. looks different for a given energy-momentum eigenstate). But these factors cancel out in the probability calculation and always result in one for any frame. Further, the $A_{\mathbf{k}}$ here are constants that do not vary with frame, so the probability of finding any particular state is also independent of what frame the measurements are taken in.

As an aside, note that

$$\langle \phi | \phi \rangle = \sum_{\mathbf{k}} \frac{(A_{\mathbf{k}})^2}{2\omega_{\mathbf{k}}} \neq 1 \tag{10}$$

because (unlike in NRQM) the LHS of (10) does not represent the integral of the probability density over space in RQM.

2.1.2 Expectation Values for Discrete Solutions

An expectation value, for energy in this example, is found using the probability density (8) in parallel fashion to that of NRQM. That is, we "sandwich" the energy operator $i\frac{\partial}{\partial t}$ inside the probability density and integrate over the volume, i.e.,

$$\bar{E} = i\left(\langle \phi | i\tfrac{\partial}{\partial t} | \phi_{x0} \rangle - \langle \phi_{x0} | i\tfrac{\partial}{\partial t} | \phi \rangle \right) = \sum_{\mathbf{k}} \langle \phi_{\mathbf{k}} | (A_{\mathbf{k}})^2 i\tfrac{\partial}{\partial t} | \phi_{\mathbf{k}} \rangle = \sum_{\mathbf{k}} (A_{\mathbf{k}})^2 \omega_{\mathbf{k}} . \tag{11}$$

This is for a single particle state and equals the statistically weighted average of the single particle eigenstate energies, as it must.

---

[1] Similar to non-relativistic quantum mechanics (NRQM), take the field equation (Klein-Gordon rather than Schroedinger) and post multiply by $|\phi\rangle$, then subtract from it the same equation pre-multiplied by $\langle\phi|$, and note the result has the form of the continuity equation (conservation of probability not mass or charge in this case.) The $\frac{\partial \rho}{\partial t}$ term in this equation has ρ of the form of the middle relation in (8).

For a multiparticle state $\left|\phi_p\phi_q\phi_r....\right\rangle$ where the particles are all general rather than eigenstates, the expectation value for the total energy of all states is found by

$$\bar{E} = i\left(\left\langle\phi_p\phi_q\phi_r....\right|i\tfrac{\partial}{\partial t}\left|\phi_p\phi_q\phi_r....\right\rangle_{,0} - \left\langle(\phi_p\phi_q\phi_r....)_{,0}\right|i\tfrac{\partial}{\partial t}\left|\phi_p\phi_q\phi_r....\right\rangle\right) \tag{12}$$

and turns out to be

$$\bar{E}_{total} = \bar{E}_p + \bar{E}_q + \bar{E}_r + \ldots \tag{13}$$

i.e, the sum of expectation energies of all individual (general state) particles.

For a multiparticle state $\left|\phi_p\phi_q\phi_r....\right\rangle$ where the particles are all in energy eigenstates, the expectation value for the total energy of all states is found by

$$E = i\left(\left\langle\phi_p\phi_q\phi_r....\right|H\left|\phi_p\phi_q\phi_r....\right\rangle_{,0} - \left\langle(\phi_p\phi_q\phi_r....)_{,0}\right|H\left|\phi_p\phi_q\phi_r....\right\rangle\right) \tag{14}$$

and turns out to be

$$E_{total} = E_p + E_q + E_r + \ldots \tag{15}$$

i.e, the sum of the energies of all individual (energy eigenstate) particles.

2.2 Continuous Solutions

The continuous solutions (3) in RQM are single particle wave packet states of form (where we substitute the Fourier amplitude $A(\mathbf{k})$ [a continuous numerical function of **k]** for operator $a(\mathbf{k})$, etc.)

$$\left|\phi(x)\right\rangle = \frac{1}{\sqrt{(2\pi)^3}}\int\frac{d\mathbf{k}}{\sqrt{2\omega_{\mathbf{k}}}}A(\mathbf{k})e^{-ikx}\,. \tag{16}$$

2.2.1 Probability for Continuous Solutions

For a single particle wave packet in RQM the probability density is

$$\begin{aligned}\rho &= i\left(\left\langle\phi\right|\phi_{,0}\right\rangle_{n.i.} - \left\langle\phi_{,0}\right|\phi\left\rangle_{n.i.}\right)\\ &= 2\left(\frac{1}{\sqrt{(2\pi)^3}}\int\frac{d\mathbf{k}}{\sqrt{2\omega_{\mathbf{k}}}}A(\mathbf{k})e^{-ikx}\right)\left(\frac{1}{\sqrt{(2\pi)^3}}\int\frac{d\mathbf{k}'}{\sqrt{2\omega_{\mathbf{k}'}}}\omega_{\mathbf{k}'}A(\mathbf{k}')e^{-ik'x}\right).\end{aligned} \tag{17}$$

We integrate this over space, using the Dirac delta function relation (which is the continuous solution case analog of (7))

$$\delta^{(3)}(\mathbf{k}-\mathbf{k}') = \frac{1}{(2\pi)^3}\int e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}}d^3x\,. \tag{18}$$

We thus find the total probability

$$\int\rho d^3x = \int|A(\mathbf{k})|^2\,d\mathbf{k} = 1\,, \tag{19}$$

which is the correct result if $A(\mathbf{k})$ is the properly normalized Fourier amplitude[2].

2.2.2 Expectation Values for Continuous Solutions

The energy expectation value (with $H = i\frac{\partial}{\partial t}$) for the wave packet is

$$\bar{E} = i\left(\left\langle\phi\right|H\left|\phi_{,0}\right\rangle - \left\langle\phi_{,0}\right|H\left|\phi\right\rangle\right) = \int|A(\mathbf{k})|^2\,\omega_{\mathbf{k}}d\mathbf{k}\,, \tag{20}$$

which again equals the statistically weighted average.

For a multiparticle state $\left|\phi_p\phi_q\phi_r....\right\rangle$ where the particles are wave packets, the expectation value for the total energy of all states turns out to be

---

[2] We note as another aside, that for wave packets in RQM $\left\langle\phi\right|\left.\phi\right\rangle = \int\frac{|A(\mathbf{k})|^2}{2\omega_{\mathbf{k}}}d\mathbf{k} \neq 1$, unlike the corresponding result for NRQM wave packets.

$$\bar{E}_{total} = \bar{E}_p + \bar{E}_q + \bar{E}_r + \ldots, \tag{21}$$

the sum of expectation energies of all individual wave packets.

3. Quantum Field Theory (QFT)

3.1 The Hamiltonian

3.1.1 Hamiltonian Density Operator

$$\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - \mu^2 \phi^\dagger \phi \tag{22}$$

$$\begin{aligned} \mathcal{H} &= \sum \pi_i \dot{\phi}^i - \mathcal{L} = \dot{\phi}^\dagger \dot{\phi} + \dot{\phi}\dot{\phi}^\dagger - \partial_\mu \phi^\dagger \partial^\mu \phi + \mu^2 \phi^\dagger \phi \\ &= \dot{\phi}\dot{\phi}^\dagger + \nabla \phi^\dagger \cdot \nabla \phi + \mu^2 \phi^\dagger \phi \end{aligned} \tag{23}$$

3.1.2 Hamiltonian Operator

Discrete solutions

$$\begin{aligned} H &= \int \mathcal{H} d^3x = \int \left(\dot{\phi}\dot{\phi}^\dagger + \nabla \phi^\dagger \cdot \nabla \phi + \mu^2 \phi^\dagger \phi \right) d^3x \\ &= \int \left(\sum_{\mathbf{k}} \frac{\partial}{\partial t} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}} \left[a(\mathbf{k}) e^{-ikx} + b^\dagger(\mathbf{k}) e^{ikx} \right] \right) \left(\sum_{\mathbf{k}'} \frac{\partial}{\partial t} \frac{1}{\sqrt{2V\omega_{\mathbf{k}'}}} \left[b(\mathbf{k}') e^{-ik'x} + a^\dagger(\mathbf{k}') e^{ik'x} \right] \right) d^3x \\ &\quad + \int \left(-\partial_i \phi^\dagger \partial^i \phi + \mu^2 \phi^\dagger \phi \right) d^3x \end{aligned} \tag{24}$$

The middle line of (24), i.e., the $\int \dot{\phi}\dot{\phi}^\dagger d^3x$ part, becomes

$$\int \left(\sum_{\mathbf{k}} \frac{i\omega_{\mathbf{k}}}{\sqrt{2V\omega_{\mathbf{k}}}} \left[-a(\mathbf{k}) e^{-ikx} + b^\dagger(\mathbf{k}) e^{ikx} \right] \right) \left(\sum_{\mathbf{k}'} \frac{i\omega_{\mathbf{k}'}}{\sqrt{2V\omega_{\mathbf{k}'}}} \left[-b(\mathbf{k}') e^{-ik'x} + a^\dagger(\mathbf{k}') e^{ik'x} \right] \right) d^3x \tag{25}$$

or

$$\sum_{\mathbf{k}} \sum_{\mathbf{k}'} \left(\frac{-\sqrt{\omega_{\mathbf{k}}}\sqrt{\omega_{\mathbf{k}'}}}{2V} \int \begin{bmatrix} a(\mathbf{k}) b(\mathbf{k}') e^{-ikx} e^{-ik'x} - a(\mathbf{k}) a^\dagger(\mathbf{k}') e^{-ikx} e^{ik'x} \\ -b^\dagger(\mathbf{k}) b(\mathbf{k}') e^{ikx} e^{-ik'x} + b^\dagger(\mathbf{k}) a^\dagger(\mathbf{k}') e^{ikx} e^{ik'x} \end{bmatrix} d^3x \right). \tag{26}$$

Note: One can shortcut the steps from here to (30) if the concern is only with finding expectation values for energy, i.e., $\bar{E} = \langle \phi | H | \phi \rangle$, and not considering eigenvalue determination, i.e., $H|\phi\rangle = E|\phi\rangle$. For the former case, all terms except those of form $\mathbf{a}^\dagger(\mathbf{k})\mathbf{a}(\mathbf{k}) + \mathbf{b}^\dagger(\mathbf{k})\mathbf{b}(\mathbf{k})$ will drop out as kets will not match bras in **k** otherwise. For the latter case, terms will survive in *H* that, for example, raise the ket by one particle, and hence there will be no eigenstate solution.

All terms in the integration in (26) result in zero except when **k=k′** or **k= - k′**. (Note that the sum over **k** and **k′** is from negative infinity to positive infinity in the x, y, and z directions.) Since the volume of integration in (26) is finite and equal to *V*, we end up with

$$\int \dot{\phi}\dot{\phi}^\dagger d^3x = \sum_{\mathbf{k}} \frac{\omega_{\mathbf{k}}}{2} \left(-a(\mathbf{k}) b(-\mathbf{k}) + a(\mathbf{k}) a^\dagger(\mathbf{k}) + b^\dagger(\mathbf{k}) b(\mathbf{k}) - b^\dagger(\mathbf{k}) a^\dagger(-\mathbf{k}) \right). \tag{27}$$

Following similar steps for the next term in (24) we get

$$\begin{aligned} -\int \partial_i \phi^\dagger \partial^i \phi d^3x &= \int \partial_i \phi^\dagger \partial_i \phi d^3x \\ &= \int \left(\sum_{\mathbf{k}} \frac{ik_i}{\sqrt{2V\omega_{\mathbf{k}}}} \left[b(\mathbf{k}) e^{-ikx} - a^\dagger(\mathbf{k}) e^{ikx} \right] \right) \left(\sum_{\mathbf{k}'} \frac{ik_i'}{\sqrt{2V\omega_{\mathbf{k}'}}} \left[a(\mathbf{k}') e^{-ik'x} - b^\dagger(\mathbf{k}') e^{ik'x} \right] \right) d^3x \\ &= \sum_{\mathbf{k}} \frac{\mathbf{k}^2}{2\omega_{\mathbf{k}}} \left(b(\mathbf{k}) a(-\mathbf{k}) + a^\dagger(\mathbf{k}) a(\mathbf{k}) + b(\mathbf{k}) b^\dagger(\mathbf{k}) + a^\dagger(\mathbf{k}) b^\dagger(-\mathbf{k}) \right) \end{aligned} \tag{28}$$

where we note that terms in the summation with both **k** and - **k** have a extra sign change since $k_i = -k_i'$ in the multiplication in the second line of (28).

Similarly, for the mass term in (24) we get

$$\begin{aligned}
&\int \mu^2 \phi^\dagger \phi d^3x = \\
&\int \mu^2 \left(\sum_{\mathbf{k}} \frac{1}{\sqrt{2V\omega_{\mathbf{k}}}} \left[b(\mathbf{k}) e^{-ikx} + a^\dagger(\mathbf{k}) e^{ikx} \right] \right) \left(\sum_{\mathbf{k}'} \frac{1}{\sqrt{2V\omega_{\mathbf{k}'}}} \left[a(\mathbf{k}') e^{-ik'x} + b^\dagger(\mathbf{k}') e^{ik'x} \right] \right) d^3x \qquad (29) \\
&= \sum_{\mathbf{k}} \frac{\mu^2}{2\omega_{\mathbf{k}}} \left(b(\mathbf{k}) a(-\mathbf{k}) + b(\mathbf{k}) b^\dagger(\mathbf{k}) + a^\dagger(\mathbf{k}) a(\mathbf{k}) + a^\dagger(\mathbf{k}) b^\dagger(-\mathbf{k}) \right)
\end{aligned}$$

Adding the last lines of (27),(28), and (29), and using $\mathbf{k}^2 + \mu^2 = (\omega_{\mathbf{k}})^2$ along with the coefficient commutation relations, we end up with

$$\begin{aligned}
H &= \sum_{\mathbf{k}} \frac{\omega_{\mathbf{k}}}{2} \left(a(\mathbf{k}) a^\dagger(\mathbf{k}) + a^\dagger(\mathbf{k}) a(\mathbf{k}) + b^\dagger(\mathbf{k}) b(\mathbf{k}) + b(\mathbf{k}) b^\dagger(\mathbf{k}) \right) \\
&= \sum_{\mathbf{k}} \omega_{\mathbf{k}} \left(a^\dagger(\mathbf{k}) a(\mathbf{k}) + \tfrac{1}{2} + b^\dagger(\mathbf{k}) b(\mathbf{k}) + \tfrac{1}{2} \right)
\end{aligned} \qquad (30)$$

or simply

$$H = \sum_{\mathbf{k}} \omega_{\mathbf{k}} \left(N_a(\mathbf{k}) + \tfrac{1}{2} + N_b(\mathbf{k}) + \tfrac{1}{2} \right). \qquad (31)$$

This is the Hamiltonian operator that acts on discrete solution states. If it is correct, to be consistent, its eigen value for a state must be the total energy of the state. For example, for a multiparticle state with 1 particle having energy eigenvalue ω_p, 2 particles having ω_q, 1 particle having ω_r, we have

$$\begin{aligned}
E_{Tot} \left| \phi_p 2\phi_q \phi_r \right\rangle &= \left(\omega_p + 2\omega_q + \omega_r \right) \left| \phi_p 2\phi_q \phi_r \right\rangle \\
&= H \left| \phi_p 2\phi_q \phi_r \right\rangle = \sum_{\mathbf{k}} \omega_{\mathbf{k}} \left(N_a(\mathbf{k}) + \tfrac{1}{2} + N_b(\mathbf{k}) + \tfrac{1}{2} \right) \left| \phi_p 2\phi_q \phi_r \right\rangle \qquad (32) \\
&= \left(n_p \omega_p + n_q \omega_q + n_r \omega_r + (\text{half integer energy states of vacuum}) \right) \left| \phi_p 2\phi_q \phi_r \right\rangle.
\end{aligned}$$

This leads us to conclude that the N_a and N_b operators must be number operators, which are unitless.

Note that if normal ordering (where one simply assumes non-commuting operators commute) were used in the first line of (30), then we would have no ½ factors in the above.

Continuous solutions

$$\begin{aligned}
H &= \int \mathcal{H} d^3x = \int \left(\dot{\phi} \dot{\phi}^\dagger + \nabla \phi^\dagger \nabla \phi + \mu^2 \phi^\dagger \phi \right) d^3x \\
&= \int \left(\frac{1}{2(2\pi)^3} \int \frac{d\boldsymbol{k}}{\sqrt{\omega_{\mathbf{k}}}} i\omega_{\mathbf{k}} \left[-a(\mathbf{k}) e^{-ikx} + b^\dagger(\mathbf{k}) e^{ikx} \right] \right) \left(\int \frac{d\boldsymbol{k}'}{\sqrt{\omega_{\mathbf{k}'}}} i\omega_{\mathbf{k}'} \left[-b(\mathbf{k}') e^{-ik'x} + a^\dagger(\mathbf{k}') e^{ik'x} \right] \right) d^3x \qquad (33) \\
&\quad + \int \left(\nabla \phi^\dagger \nabla \phi + \mu^2 \phi^\dagger \phi \right) d^3x
\end{aligned}$$

The middle line of (33), i.e., the $\int \dot{\phi} \dot{\phi}^\dagger d^3x$ part, becomes

$$\int \left(\int \left(\frac{1}{2(2\pi)^3} \int \frac{(-\omega_{\mathbf{k}} \omega_{\mathbf{k}'})}{\sqrt{\omega_{\mathbf{k}}} \sqrt{\omega_{\mathbf{k}'}}} \left[-a(\mathbf{k}) e^{-ikx} + b^\dagger(\mathbf{k}) e^{ikx} \right] \left[-b(\mathbf{k}') e^{-ik'x} + a^\dagger(\mathbf{k}') e^{ik'x} \right] d\boldsymbol{k}' \right) d\boldsymbol{k} \right) d^3x \qquad (34)$$

or

$$\int\left(\int\left(\frac{1}{2(2\pi)^3}\frac{-\omega_{\mathbf{k}}\omega_{\mathbf{k}'}}{\sqrt{\omega_{\mathbf{k}}}\sqrt{\omega_{\mathbf{k}'}}}\int\left[\begin{array}{l}a(\mathbf{k})b(\mathbf{k}')e^{-ikx}e^{-ik'x}-a(\mathbf{k})a^{\dagger}(\mathbf{k}')e^{-ikx}e^{ik'x}\\-b^{\dagger}(\mathbf{k})b(\mathbf{k}')e^{ikx}e^{-ik'x}+b^{\dagger}(\mathbf{k})a^{\dagger}(\mathbf{k}')e^{ikx}e^{ik'x}\end{array}\right]d^3x\right)d\mathbf{k}'\right)d\mathbf{k}\,. \qquad (35)$$

Using the Dirac delta function relation (18) for integration over all (infinite) space in the integral over $d\mathbf{k}'$ in (35) results in a relation parallel to (27), i.e.,

$$\int\dot{\phi}\dot{\phi}^{\dagger}d^3x=\int\frac{\omega_{\mathbf{k}}}{2}\left(-a(\mathbf{k})b(-\mathbf{k})+a(\mathbf{k})a^{\dagger}(\mathbf{k})+b^{\dagger}(\mathbf{k})b(\mathbf{k})-b^{\dagger}(\mathbf{k})a^{\dagger}(-\mathbf{k})\right)d\mathbf{k}\,. \qquad (36)$$

Evaluating the other two terms in (33) (last line) in similar fashion to that of (28) through (30), one ends up with the parallel relation to (31), i.e.,

$$H=\int\omega_{\mathbf{k}}\left(N_a(\mathbf{k})+\tfrac{1}{2}+N_b(\mathbf{k})+\tfrac{1}{2}\right)d\mathbf{k}\,. \qquad (37)$$

This is the form of the Hamiltonian for continuous solution states, i.e., to be used with wave packets. For a single particle state wave packet, there is no eigenstate of energy as the packet, by definition is a superposition of eigenstates (of infinite number and infinitesimal width in **k** space). We discuss energy expectation values for general (non-eigenstate) single and multi particle states in Section 5.

Note that the integral in **k** vector space of (37) is a 3D integral (in that space, not physical space), so it can also be expressed as a scalar integral

$$H=\int\omega_{\mathbf{k}}\left(N_a(\mathbf{k})+\tfrac{1}{2}+N_b(\mathbf{k})+\tfrac{1}{2}\right)d^3k\,, \qquad (38)$$

where d^3k here is an infinitesimal volume in **k** space. **k** has magnitude $2\pi/\lambda$ and thus units of 1/length, so d^3k has units of 1/volume. From (38), the units of $N_a(\mathbf{k})$ and $N_b(\mathbf{k})$ must be then be volume, or length^3. And thus, from $N_a(\mathbf{k}) = a^{\dagger}(\mathbf{k})a(\mathbf{k})$ and $N_b(\mathbf{k}) = b^{\dagger}(\mathbf{k})b(\mathbf{k})$, $a(\mathbf{k})$ and $b(\mathbf{k})$ must have units of $\sqrt{volume}$ = length$^{3/2}$. This differs from the discrete case where all these operators were unitless.

4. Creating and Destroying General (non-eigenstate) States

Questions arise in QFT as to what is created or destroyed by the general solution $\phi(x)$ (or $\phi^{\dagger}(x)$), which for discrete eigenstates, is a summation of terms, each containing a single particle eigenstate creation/destruction operator. Does operation of $\phi^{\dagger}(x)$ on the vacuum, for instance, create an infinite number of single particles, or a single particle comprising an infinite number of momentum eigenstates? If the latter, what amplitudes (whose squares are probabilities) are assigned to each such eigenstate?

Similar questions also arise regarding the general continuous eigenstate solution of (3). These are compounded by the continuous nature of $\phi(x)$. Does $\phi^{\dagger}(x)$ acting on the vacuum create a single particle wave packet state? If so, what (continuous) Fourier amplitude spectrum does the wave packet have? That is, how do we determine how "spread out" or how "tight" the created wave packet is?

Answer:

We do not use $\phi^{\dagger}(x)$ to create particles, so we should not be worried about the sum of terms in $\phi^{\dagger}(x)$ for creating states. We use $\mathbf{a}^{\dagger}(\mathbf{k})$ to create a unit normed state. Field operators like $\phi^{\dagger}(x)$ appear in bi-linear form (such as $\phi^{\dagger}\phi$) in all observable operators like H, and it is only these operators that have expectation values, e.g., $\bar{E}=\langle\phi|H|\phi\rangle$. In these cases all factors like $e^{-ikx}/\sqrt{2\omega_{\mathbf{k}}V}$ drop out and we are left with just number operators (and things like $\omega_{\mathbf{k}}$).

Confusion can arise here when one considers the heuristic treatment for finding the propagator in which the relation

$$\langle 0|\left[\phi(x),\phi^{\dagger}(y)\right]|0\rangle=i\Delta(x-y) \qquad (39)$$

was used to describe a particle created out of the vacuum at *y* and annihilated at *x*. This led to the Feynman propagator, i.e., the amplitude for a virtual particle traveling from *y* to *x*. One can then begin to think in terms of $\phi^\dagger$ as the operator to use to create and destroy states.

In reality, the propagator comes out of the mathematics in finding the S operator between initial and final states. In the Interaction Picture, the equation of motion for the states involved the Hamiltonian operator *H*. Integrating this equation involved the Dyson-Wicks expansion in which terms therein ended up containing factors of the form of (39). The bi-linear operator form of *H* led to such factors. These factors are the propagators for the virtual particles between *y* and *x*.

One can think of these factors (i.e., of (39)) roughly as $\phi^\dagger$ creating a state at *y* that ϕ destroys at *x*, but that is not completely accurate, and as noted, can lead to confusion. In reality the $\mathbf{a}^\dagger$ operator does the creating, and the other factors in a particular term in $\phi^\dagger$ lead to the correct form for the propagator in $e^{-ik(x-y)}$ etc.

4.1 Discrete Eigenstates

4.1.1 Creating a Single Particle State (Discrete Solutions Form)

Single Eigenstate for Single Particle (Discrete Solution Form)

We know that

$$a^\dagger(\mathbf{k})|0\rangle = |\phi_{\mathbf{k}}\rangle = \left|\frac{e^{-ik\tilde{x}}}{\sqrt{V}}\right\rangle, \tag{40}$$

which has unit norm, and for which we employ "~" over *x* to distinguish it from the *x* dependence in field operators such as $\phi^\dagger$. The reason for this follows.

Aside

Suppose we wish to evaluate an expression similar to (39), such as (41) below. Using (2) on half of the commutator, we would have

$$\langle 0|\phi\phi^\dagger|0\rangle = \sum_{\mathbf{k}}\sum_{\mathbf{k}'}\left(\langle\phi_{\mathbf{k}'}(\tilde{y})|\frac{e^{ik'y}}{\sqrt{2\omega_{\mathbf{k}'}V}}\frac{e^{-ikx}}{\sqrt{2\omega_{\mathbf{k}}V}}|\phi_{\mathbf{k}}(\tilde{x})\rangle\right). \tag{41}$$

The point is that the integration implied by the inner product of the bracket is over the "~" coordinates, not the *xy* coordinates. The result of (41) is a function (without operators involved) of *y*-*x*.

The state created by $a^\dagger(\mathbf{k})$, i.e., $|\phi_{\mathbf{k}}(\tilde{x})\rangle$, is a function of $\tilde{x}$, which is a different position variable than *x* in the e^{-ikx} factor shown in (41). That is, the created state has its own particular function of position and time that is unrelated to that of the other position dependent term in $\phi^\dagger$.

In QFT the accustomed manner of treating states is simply to use the ket form $|\phi_{\mathbf{k}}\rangle$ without showing, or dealing with, the inherent spacetime dependence explicitly. We will gravitate towards this usage as well, but for some of the derivations that follow directly, it can help if that spacetime dependence is shown explicitly. The interested reader can verify for him/herself that if the created state depended on the same spacetime coordinates *x* as the e^{ikx} factor in $\phi^\dagger$, then incorrect results arise.

Each $|\phi_{\mathbf{k}}\rangle$ in (41)) has unit norm and is orthogonal to every other such eigenstate. That is,

$$|\phi_{\mathbf{k}}\rangle = \frac{e^{-ik\tilde{x}}}{\sqrt{V}}, \tag{42}$$

so that[3]

---

[3] Actually, though it is a subtle point at this stage, $\left|e^{-ikx'}\right\rangle$ here has no energy (time) dependence in the exponent as long as we are working in with the Heisenberg picture (which we do with free fields in the usual development of QFT.) In the Schroedinger picture, the state would have time dependence. When

$$\langle \phi_{\mathbf{k}} | \phi_{\mathbf{k}} \rangle = \frac{1}{V} \int e^{ik\tilde{x}} e^{-ik\tilde{x}} d^3x' = 1 \tag{43}$$

or more generally,

$$\langle \phi_{\mathbf{k}} | \phi_{\mathbf{k}'} \rangle = \delta_{\mathbf{k}\mathbf{k}'} . \tag{44}$$

In QFT the middle part of (43) is rarely expressed and one simply uses (44).

General non-Eigen State for Single Particle (Discrete Solution Form)

To create a general particle state, which is a sum of eigenstates, we would need an operator of form

$$C = \sum_{\mathbf{k}} A_{\mathbf{k}} \mathbf{a}_{\mathbf{k}}^{\dagger} , \tag{45}$$

so that

$$\begin{aligned} C|0\rangle &= \sum_{\mathbf{k}} A_{\mathbf{k}} \mathbf{a}_{\mathbf{k}}^{\dagger} |0\rangle = A_1 \left| \frac{e^{-ik_1 x}}{\sqrt{V}} \right\rangle + A_2 \left| \frac{e^{-ik_2 x}}{\sqrt{V}} \right\rangle + A_3 \left| \frac{e^{-ik_3 x}}{\sqrt{V}} \right\rangle + ... \\ &= A_1 |\phi_1\rangle + A_2 |\phi_2\rangle + A_3 |\phi_3\rangle + ... = |\phi\rangle . \end{aligned} \tag{46}$$

In (45) and (46) $A_{\mathbf{k}}$ is a numerical coefficient, the square of which (for proper normalization) equals the probability of finding the **k** eigenstate. (See (4), (8), and (9).)

If only one term in C is used, then only one eigenstate with $|A_{\mathbf{k}}| = 1$ is created. If a more general state, comprising a sum of eigenstates, is created, then we are free to select the $A_{\mathbf{k}}$ as we please in order to create the particular general state we like, provided (for conservation of probability and correct normalization so total probability is unity)

$$\sum_{\mathbf{k}} |A_{\mathbf{k}}|^2 = 1 . \tag{47}$$

Important point: Note in QFT we have

$$\langle \phi | \phi \rangle = \sum_{\mathbf{k}} |A_{\mathbf{k}}|^2 = 1 \tag{48}$$

whereas in NRQM, we had (see (10), repeated below)

$$\langle \phi | \phi \rangle = \sum_{\mathbf{k}} \frac{(A_{\mathbf{k}})^2}{2\omega_{\mathbf{k}}} \neq 1 . \tag{49}$$

This is because the kets created in QFT via $\mathbf{a}_{\mathbf{k}}^{\dagger}$ have unit norm, whereas the ket solutions to the field equations in NRQM do not.

4.1.2 Destroying a Single Particle State (Discrete)

Note that the operator

$$D = \sum_{\mathbf{k}} \mathbf{a}_{\mathbf{k}} \tag{50}$$

acting on any single particle general state will lower that state to the vacuum. If acting on the vacuum, each term in (50) will destroy it. That is,

---

going to the interaction picture (for interacting fields in QFT) the time dependence for states will be on the interaction part of the Hamiltonian (but not the free part.)

$$\left(\sum_{\mathbf{k}}\mathbf{a}_{\mathbf{k}}\right)|\phi\rangle=\left(\sum_{\mathbf{k}}\mathbf{a}_{\mathbf{k}}\right)\left|A_1|\phi_1\rangle+A_2|\phi_2\rangle+A_3|\phi_3\rangle+...\right\rangle$$
$$=\left(\left(\sum_{\mathbf{k}}\mathbf{a}_{\mathbf{k}}\right)A_1|\phi_1\rangle+\left(\sum_{\mathbf{k}}\mathbf{a}_{\mathbf{k}}\right)A_2|\phi_2\rangle+\left(\sum_{\mathbf{k}}\mathbf{a}_{\mathbf{k}}\right)A_3|\phi_3\rangle+\ldots\right) \quad (51)$$
$$=A_1|0\rangle+0+0+...+0+A_2|0\rangle+0+...=|0\rangle$$

(We re-normalized the vacuum on the RHS above.)

4.1.3 Creating a Multi-particle State (Discrete Solution Form)

Applying operators similar in form to (45) (with typically different values for $A_{\mathbf{k}}$ in each operator) twice in succession creates a two particle state where each particle is a single particle general state (i.e., each is a summation of momentum eigenstates.) Any number of such operators may be applied to create a state of any number of particles, each in a general (not eigen) state.

Multiparticle states have unit norms, e.g.,

$$\langle\phi_p 2\phi_q\phi_r||\phi_p 2\phi_q\phi_r\rangle=1. \quad (52)$$

4.1.4 Destroying a Multi-particle State (Discrete Solution Form)

Application of (50) repeatedly will destroy one general state single particle upon each application.

4.2 Continuous Eigenstates

4.2.1 Creating a Wave Packet (Single Particle State of Continuous Solution Form)

For continuous solution form states, we parallel our use in 4.1 above of the creation operators $\mathbf{a}_{\mathbf{k}}^{\dagger}$ in (2) to create a general creation operator. (Note there is no such thing as an eigenstate of continuous solution form.) We use the operators in (3) to create a single particle wave packet composed of an integral of continuous momenta eigenstates. That is, by analogy,

$$C=\int d\mathbf{k}\,A(\mathbf{k})\mathbf{a}^{\dagger}(\mathbf{k}), \quad (53)$$

which can be seen with the aid of the table below.

	Discrete Solution Form	Continuous Solution Form
Eigenstate creation operator	$\mathbf{a}^{\dagger}_{\mathbf{k}}\|0\rangle=\|\phi_{\mathbf{k}}\rangle=\left\|\frac{e^{-ikx}}{\sqrt{V}}\right\rangle$	$\mathbf{a}^{\dagger}(\mathbf{k})\|0\rangle=\|\phi_{\mathbf{k}}\rangle=\left\|\frac{e^{-ikx}}{\sqrt{(2\pi)^3}}\right\rangle$
General state creation operator	$C=\sum_{\mathbf{k}}A_{\mathbf{k}}\mathbf{a}^{\dagger}_{\mathbf{k}}$	$C=\int d\mathbf{k}\,A(\mathbf{k})\mathbf{a}^{\dagger}(\mathbf{k})$
General state	$\|\phi\rangle=C\|0\rangle=A_1\left\|\frac{e^{-ik_1x}}{\sqrt{V}}\right\rangle+A_2\left\|\frac{e^{-ik_2x}}{\sqrt{V}}\right\rangle+...$ $=A_1\|\phi_1\rangle+A_2\|\phi_2\rangle+...$ $=\sum_{\mathbf{k}}A_{\mathbf{k}}\left\|\frac{e^{-ikx}}{\sqrt{V}}\right\rangle$	$\|\phi\rangle=C\|0\rangle$ $=\int d\mathbf{k}\,A(\mathbf{k})\left\|\frac{e^{-ikx}}{\sqrt{(2\pi)^3}}\right\rangle$

$A(\mathbf{k})$ is the Fourier amplitude, which is a numerical continuous function of **k**, and which we can choose as we like to create the wave packet shape desired.

4.2.2 Destroying a Wave Packet (Single Particle State of Continuous Solution Form)

Once again, by analogy, we have a wave packet destruction operator

$$D = \int d\boldsymbol{k}\, \mathbf{a}(\mathbf{k}), \tag{54}$$

which can be seen with the aid of the table below.

	Discrete Solution Form	Continuous Solution Form
Eigenstate destruction operator	$\mathbf{a}_\mathbf{k}\left\|\phi_\mathbf{k}\right\rangle = \mathbf{a}_\mathbf{k}\left\|\frac{e^{-ikx}}{\sqrt{V}}\right\rangle = \left\|0\right\rangle$	$\mathbf{a}(\mathbf{k})\left\|\phi_\mathbf{k}\right\rangle = \mathbf{a}_\mathbf{k}\left\|\frac{e^{-ikx}}{\sqrt{(2\pi)^3}}\right\rangle = \left\|0\right\rangle$
General state destruction operator	$D = \sum_\mathbf{k} \mathbf{a}_\mathbf{k}$	$D = \int d\boldsymbol{k}\, \mathbf{a}(\mathbf{k})$

4.2.3 Creating a Multi-Wave Packet State (Multi-particle Continuous Solution Form)

Applying operators similar in form to C of (53) (with typically different Fourier spectra $A(\mathbf{k})$ in each operator) twice in succession creates a two particle state where each particle is a wave packet.

4.2.4 Destroying a Multi-Wave Packet State (Multi-particle Continuous Solution Form)

Application of (54) repeatedly will destroy one wave packet particle upon each application.

5. Probability and Expectation Values in QFT

As described in footnote 1 on page 2, we can determine a continuity (conservation) equation, i.e., a 4D current divergence equal to zero, in similar fashion to what was done in NRQM to find the probability conservation relation. Take the field equation (Klein-Gordon rather than Schroedinger) and pre multiply by $\phi^\dagger$, then subtract it from the complex conjugate Klein-Gordon equation post multiplied by ϕ, and note the result has the form of the continuity equation

$$\frac{\partial \rho'_{oper}}{\partial t} + \nabla \cdot \mathbf{j}'_{\mathbf{oper}} = j'_{oper}{}^{\mu}{}_{,\mu} = 0 \tag{55}$$

where

$$\rho'_{oper} = j'_{oper}{}^{0} = i\left(\phi^\dagger \phi_{,0} - \phi^\dagger{}_{,0}\,\phi\right) \qquad \mathbf{j}'_{oper} = j'_{oper}{}^{i} = -i\left(\phi^\dagger \phi_{,i} - \phi^\dagger{}_{,i}\phi\right). \tag{56}$$

We use the subscript "oper" to distinguish between the ρ of RQM, which represented the numerical particle density, and ρ'_{oper} of QFT, which is composed of operators and is therefore itself an operator. We also use a prime on ρ'_{oper} for reasons which will be seen. We derive below a single particle density *operator* ρ_{oper} that is closely related to ρ'_{oper}.

5.1 Probability Density for Discrete Solutions in QFT

Note that from (2) ρ'_{oper} has the form

$$\begin{aligned}\rho'_{oper} = &\left(2\sum_\mathbf{k} \frac{\mathbf{a}^\dagger_\mathbf{k}}{\sqrt{2\omega_\mathbf{k}}}\frac{e^{ikx}}{\sqrt{V}}\right)\left(\sum_{\mathbf{k}'} \frac{\omega_{\mathbf{k}'}\mathbf{a}_{\mathbf{k}'}}{\sqrt{2\omega_{\mathbf{k}'}}}\frac{e^{-ik'x}}{\sqrt{V}}\right) - \left(2\sum_\mathbf{k} \frac{\omega_\mathbf{k}\mathbf{b}^\dagger_\mathbf{k}}{\sqrt{2\omega_\mathbf{k}}}\frac{e^{ikx}}{\sqrt{V}}\right)\left(\sum_{\mathbf{k}'} \frac{\mathbf{b}_{\mathbf{k}'}}{\sqrt{2\omega_{\mathbf{k}'}}}\frac{e^{-ik'x}}{\sqrt{V}}\right) \\ &- \left(2\sum_\mathbf{k} \frac{\omega_\mathbf{k}}{\sqrt{2\omega_\mathbf{k}}}\frac{e^{ikx}}{\sqrt{V}}\right)\left(\sum_{\mathbf{k}'} \frac{1}{\sqrt{2\omega_{\mathbf{k}'}}}\frac{e^{-ik'x}}{\sqrt{V}}\right) + (\text{terms in } \mathbf{a}^\dagger{}_\mathbf{k}\mathbf{b}^\dagger{}_{\mathbf{k}'} \text{ and } \mathbf{b}_\mathbf{k}\mathbf{a}_\mathbf{k})\end{aligned}. \tag{57}$$

The third term in (57) will cause us problems later on. We can circumvent those problems by noting that ρ'_{oper} is not the only entity that satisfies the continuity equation (55).

Take the field equation (Klein-Gordon here) and *post* multiply it by $\phi^\dagger$, then subtract that from the complex conjugate field equation *pre* multiplied by ϕ. You obtain a ρ''_{oper} and $\mathbf{j}''_{oper}$ that satisfy the continuity equation (55), for which

$$\rho''_{oper} = i\left(\phi\phi^\dagger{}_{,0} - \phi_{,0}\phi^\dagger\right) \qquad j''^{\,i} = -i\left(\phi\phi^\dagger{}_{,i} - \phi_{,i}\phi^\dagger\right). \tag{58}$$

XXX Need to check signs here XXX The complete form for the probability density operator should then be a linear combination of ρ'_{oper} and ρ''_{oper}. We take this to be the average of the two, i.e.,

$$\rho_{oper} = \frac{\rho'_{oper} + \rho''_{oper}}{2}. \tag{59}$$

In (59), not only do terms like the third one in (57) cancel out, but the entire bottom row in (57) does as well. Thus, we find

$$\rho_{oper} = \left(2\sum_{\mathbf{k}} \frac{\mathbf{a}^\dagger_{\mathbf{k}}}{\sqrt{2\omega_{\mathbf{k}}}} \frac{e^{ikx}}{\sqrt{V}}\right)\left(\sum_{\mathbf{k}'} \frac{\omega_{\mathbf{k}'}\mathbf{a}_{\mathbf{k}'}}{\sqrt{2\omega_{\mathbf{k}'}}} \frac{e^{-ik'x}}{\sqrt{V}}\right) - \left(2\sum_{\mathbf{k}} \frac{\omega_{\mathbf{k}}\mathbf{b}^\dagger_{\mathbf{k}}}{\sqrt{2\omega_{\mathbf{k}}}} \frac{e^{ikx}}{\sqrt{V}}\right)\left(\sum_{\mathbf{k}'} \frac{\mathbf{b}_{\mathbf{k}'}}{\sqrt{2\omega_{\mathbf{k}'}}} \frac{e^{-ik'x}}{\sqrt{V}}\right). \tag{60}$$

5.1.1 Single Particle Probability Density (QFT, Discrete)

Single Particle Eigenstate

Hence, for a single particle in an eigenstate $|\phi_{\mathbf{k}''}\rangle$, the numerical probability density is the expectation value of the corresponding operator,

$$\rho = \langle\phi_{\mathbf{k}''}|\rho_{oper}|\phi_{\mathbf{k}''}\rangle = \langle\phi_{\mathbf{k}''}(\tilde{x})|\rho_{oper}(x)|\phi_{\mathbf{k}''}(\tilde{x})\rangle. \tag{61}$$

All terms in (60) with $\mathbf{k} \neq \mathbf{k}'$ will result in different particles (in orthogonal states) in the bra and ket, and drop out, leaving

$$\begin{aligned}\rho &= \langle\phi_{\mathbf{k}''}|\frac{\sum_{\mathbf{k}}(\mathbf{a}^\dagger{}_{\mathbf{k}}\mathbf{a}_{\mathbf{k}} - \mathbf{b}^\dagger{}_{\mathbf{k}}\mathbf{b}_{\mathbf{k}})}{V}|\phi_{\mathbf{k}''}\rangle = \langle\phi_{\mathbf{k}''}|\frac{\sum_{\mathbf{k}}(N_{\mathbf{a}}(\mathbf{k}) - N_{\mathbf{b}}(\mathbf{k}))}{V}|\phi_{\mathbf{k}''}\rangle, \\ &= \frac{1}{V}.\end{aligned} \tag{62}$$

Note that the bracket integration over $\tilde{x}$ causes only terms where $\mathbf{k} = \mathbf{k}'$ to survive. This also results in the cancellation of factors in the numerators and denominators, and the severing of dependence on *x*.

The final result in (62) is what we would expect. The total probability is the integral of ρ over the volume and equals unity. Note that an antiparticle state would have a negative probability density (from the $N_{\mathbf{b}}(\mathbf{k})$ operator) and a total probability of negative one. This led to the interpretation of ρ as charge density (probability of finding the given charge at any particular location) and its integral over volume as the particle/antiparticle charge.

Note further, that probability density in QFT for a unit norm ket is not invariant, due to the relativistic change in volume *V* for a different frame. But total probability is invariant (and always equals one), since in the integration over volume, the *V* factors cancel.

Single Particle General State

For a general single particle state, composed of a superposition of eigenstates, where

$$|\phi\rangle = A_1|\phi_1\rangle + A_2|\phi_2\rangle + A_3|\phi_3\rangle + \ldots, \tag{63}$$

we have, ignoring anti-particles for simplicity,

$$\begin{aligned}&\rho = \langle\phi|\rho_{oper}|\phi\rangle = \langle\phi(\tilde{x})|\rho_{oper}(x)|\phi(\tilde{x})\rangle = \\ &\left\langle\sum_{\mathbf{k}} A^\dagger_{\mathbf{k}} \frac{e^{ik\tilde{x}}}{\sqrt{V}}\right|\left(2\sum_{\mathbf{k}'} \frac{\mathbf{a}^\dagger_{\mathbf{k}'}}{\sqrt{2\omega_{\mathbf{k}'}}} \frac{e^{ik'x}}{\sqrt{V}}\right)\left(\sum_{\mathbf{k}''} \frac{\omega_{\mathbf{k}''}\mathbf{a}_{\mathbf{k}''}}{\sqrt{2\omega_{\mathbf{k}''}}} \frac{e^{-ik''x}}{\sqrt{V}}\right)\left|\sum_{\mathbf{k}'''} A_{\mathbf{k}'''} \frac{e^{ik'''\tilde{x}}}{\sqrt{V}}\right\rangle.\end{aligned} \tag{64}$$

To help in evaluating (64), look initially at only the first two terms in each of the ket and the right hand operator summations.

$$\left(\sum_{\mathbf{k}''}\frac{\omega_{\mathbf{k}''}\mathbf{a}_{\mathbf{k}''}}{\sqrt{2\omega_{\mathbf{k}''}}}\frac{e^{-ik''x}}{\sqrt{V}}\right)\left|\sum_{\mathbf{k}'''}A_{\mathbf{k}'''}\frac{e^{-ik'''\tilde{x}}}{\sqrt{V}}\right\rangle \rightarrow \text{1st two terms} =$$

$$\left(\frac{\omega_{\mathbf{k}_1}\mathbf{a}_{\mathbf{k}_1}}{\sqrt{2\omega_{\mathbf{k}_1}}}\frac{e^{-ik_1x}}{\sqrt{V}}+\frac{\omega_{\mathbf{k}_2}\mathbf{a}_{\mathbf{k}_2}}{\sqrt{2\omega_{\mathbf{k}_2}}}\frac{e^{-ik_2x}}{\sqrt{V}}\right)\left|A_{\mathbf{k}_1}\frac{e^{-ik_1\tilde{x}}}{\sqrt{V}}+A_{\mathbf{k}_2}\frac{e^{-ik_2\tilde{x}}}{\sqrt{V}}\right\rangle = \tag{65}$$

$$\frac{\omega_{\mathbf{k}_1}}{\sqrt{2\omega_{\mathbf{k}_1}}}\frac{e^{-ik_1x}}{\sqrt{V}}A_{\mathbf{k}_1}\left|0\right\rangle+\frac{\omega_{\mathbf{k}_2}}{\sqrt{2\omega_{\mathbf{k}_2}}}\frac{e^{-ik_2x}}{\sqrt{V}}A_{\mathbf{k}_2}\left|\frac{e^{-ik_2\tilde{x}}}{\sqrt{V}}\right\rangle$$

The two summations (in $\mathbf{k}$ and $\mathbf{k}'$) on the left side of (64) look like

$$2\left\langle 0\right|A^{\dagger}_{\mathbf{k}_1}\frac{1}{\sqrt{2\omega_{\mathbf{k}_1}}}\frac{e^{ik_1x}}{\sqrt{V}}+2\left\langle\frac{e^{ik_1\tilde{x}}}{\sqrt{V}}\right|A^{\dagger}_{\mathbf{k}_2}\frac{1}{\sqrt{2\omega_{\mathbf{k}_2}}}\frac{e^{ik_2x}}{\sqrt{V}}. \tag{66}$$

If we take the bra-ket, i.e., the integral over $\tilde{\mathbf{x}}$, of (64), using (66) and the last line of (65), we get (including all terms, not just the first two)

$$\rho=\frac{1}{V}\sum_{\mathbf{k}}A^{\dagger}_{\mathbf{k}}\sum_{\mathbf{k}''}A_{\mathbf{k}}\,, \tag{67}$$

which is the probability density. If we integrate it over all space to get total probability, we find

$$\int\rho dV=\sum_{\mathbf{k}}\left|A_{\mathbf{k}}\right|^2=1\,, \tag{68}$$

which is what it should be, and also equals the total number of particles.

Parallel remarks to those made above with regard to single particle eigenstates for total probability, antiparticle charge/probability density, and invariance apply to general single particle states, as well.

5.1.2 Multiple Particle State Probability Density (QFT, Discrete)

All Particles in Eigenstates

For a multi particle state in which all particles are in eigenstates, such as

$$\left|\phi\right\rangle=\left|\phi_p 2\phi_q\phi_r\right\rangle \tag{69}$$

we have

$$\rho=\left\langle\phi_p 2\phi_q\phi_r\right|\rho_{oper}\left|\phi_p 2\phi_q\phi_r\right\rangle. \tag{70}$$

Any operator acts on a multi particle ket one particle at a time, much like a derivative on a product of fucntions. Hence, for a destruction operator $\mathbf{a}_{\mathbf{k}'}$, one would have

$$\begin{aligned}\sum_{\mathbf{k}'}\mathbf{a}_{\mathbf{k}'}\left|\phi_p 2\phi_q\phi_r\right\rangle=&\left|\left(\sum_{\mathbf{k}'}\mathbf{a}_{\mathbf{k}'}\phi_p\right)2\phi_q\phi_r\right\rangle+\left|\phi_p\left(\sum_{\mathbf{k}'}\mathbf{a}_{\mathbf{k}'}\phi_q\right)\phi_q\phi_r\right\rangle\\&+\left|\phi_p\phi_q\left(\sum_{\mathbf{k}'}\mathbf{a}_{\mathbf{k}'}\phi_q\right)\phi_r\right\rangle+\left|\phi_p 2\phi_q\left(\sum_{\mathbf{k}'}\mathbf{a}_{\mathbf{k}'}\phi_r\right)\right\rangle\end{aligned}. \tag{71}$$

with a parallel relation for the action of $\mathbf{a}^{\dagger}_{\mathbf{k}}$ on the bra. When (71) is used with (60) in (70), all kets are destroyed (become equal to zero) except those for which $\mathbf{k}'$ equals the eigen momentum of one of the particles. This leaves only those eigen momentum terms inside (70), and thus we have a relationship similar to that for a single particle eigenstate (62), i.e.,

$$\rho = \langle\phi| \frac{\sum_{\mathbf{k}} (\mathbf{a}^{\dagger}{}_{\mathbf{k}}\mathbf{a}_{\mathbf{k}} - \mathbf{b}^{\dagger}{}_{\mathbf{k}}\mathbf{b}_{\mathbf{k}})}{V} |\phi\rangle = \langle\phi_p 2\phi_q\phi_r| \frac{\sum_{\mathbf{k}} (N_{\mathbf{a}}(\mathbf{k}) - N_{\mathbf{b}}(\mathbf{k}))}{V} |\phi_p 2\phi_q\phi_r\rangle$$
$$= \frac{1+2+1}{V} = \frac{4}{V}. \tag{72}$$

The integral of this over the volume yields a total probability of 4, which for 4 particles, might make sense in some sort of way. Since this integral equals the number of particles, ρ can thus be more properly interpreted as particle number density (or charge density) where antiparticles have negative numbers.

In QFT, which invariably deals with multiparticle states it is more advantageous to focus on the number operators. In fact, we can think of the total number operator as the integral of ρ_{oper} over the volume.

$$N = \int \rho_{oper} dV = \sum_{\mathbf{k}} (N_{\mathbf{a}}(\mathbf{k}) - N_{\mathbf{b}}(\mathbf{k})). \tag{73}$$

Particles in General States

Consider multi particle states where the particles are in general (non eigen) states, i.e., $|\phi_p 2\phi_q\phi_r\rangle$, where, for example,

$$|\phi_p\rangle = \left| \sum_{\mathbf{k}'''} A_{\mathbf{k}'''} \frac{e^{-ik'''\tilde{x}}}{\sqrt{V}} \right\rangle_p, \tag{74}$$

and similar relations hold for the other particles in the multi particle state.

When (71) and its parallel relation for the bra are used in (70), we get a term similar to (67) for each ket term on the right side of (71), i.e.,

$$\rho = \frac{1}{V} \left\{ \begin{array}{l} \left(\sum_{\mathbf{k}'} A^{\dagger}_{\mathbf{k}'} \frac{e^{ik'x}}{\sqrt{2\omega_{\mathbf{k}'}}} \right)_{\phi_p} \left(\sum_{\mathbf{k}''} \omega_{\mathbf{k}''} \frac{e^{-ik''x}}{\sqrt{2\omega_{\mathbf{k}''}}} A_{\mathbf{k}''} \right)_{\phi_p} \\ + \left(\sum_{\mathbf{k}'} A^{\dagger}_{\mathbf{k}'} \frac{e^{ik'x}}{\sqrt{2\omega_{\mathbf{k}'}}} \right)_{\phi_q} \left(\sum_{\mathbf{k}''} \omega_{\mathbf{k}''} \frac{e^{-ik''x}}{\sqrt{2\omega_{\mathbf{k}''}}} A_{\mathbf{k}''} \right)_{\phi_q} + \ldots \end{array} \right\} \tag{75}$$

When we integrate (75) over all space we get

$$\int \rho dV = \left\{ \begin{array}{l} \delta_{\mathbf{k}'\mathbf{k}''} \left(\sum_{\mathbf{k}'} \frac{A^{\dagger}_{\mathbf{k}'}}{\sqrt{2\omega_{\mathbf{k}'}}} \right)_{\phi_p} \left(\sum_{\mathbf{k}''} \omega_{\mathbf{k}''} \frac{A_{\mathbf{k}''}}{\sqrt{2\omega_{\mathbf{k}''}}} \right)_{\phi_p} \\ + 2\delta_{\mathbf{k}'\mathbf{k}''} \left(\sum_{\mathbf{k}'} \frac{A^{\dagger}_{\mathbf{k}'}}{\sqrt{2\omega_{\mathbf{k}'}}} \right)_{\phi_q} \left(\sum_{\mathbf{k}''} \omega_{\mathbf{k}''} \frac{A_{\mathbf{k}''}}{\sqrt{2\omega_{\mathbf{k}''}}} \right)_{\phi_q} \\ + \delta_{\mathbf{k}'\mathbf{k}''} \left(\sum_{\mathbf{k}'} \frac{A^{\dagger}_{\mathbf{k}'}}{\sqrt{2\omega_{\mathbf{k}'}}} \right)_{\phi_r} \left(\sum_{\mathbf{k}''} \omega_{\mathbf{k}''} \frac{A_{\mathbf{k}''}}{\sqrt{2\omega_{\mathbf{k}''}}} \right)_{\phi_r} \end{array} \right\} \tag{76}$$

or

$$\int \rho dV = \sum_{\mathbf{k}'} \left|{}_p A_{\mathbf{k}'}\right|^2 + 2\sum_{\mathbf{k}'} \left|{}_q A_{\mathbf{k}'}\right|^2 + \sum_{\mathbf{k}'} \left|{}_r A_{\mathbf{k}'}\right|^2 = 1 + 2 + 1 = 4. \tag{77}$$

Relation (75) is cumbersome to say the least, whereas (77) is quite simple and equals the total number of particles. In QFT, it turns out to be invariably simpler to focus on the number operators, for which

$$\begin{aligned}&\text{Number of particles} = \int \rho dV = \int \langle\phi|\rho_{oper}(x)|\phi\rangle d^3\mathbf{x}\\ &= \text{Expectation value of number operator}\\ &= \langle\phi|\sum_{\mathbf{k}}(\mathbf{a}^\dagger{}_{\mathbf{k}}\mathbf{a}_{\mathbf{k}} - \mathbf{b}^\dagger{}_{\mathbf{k}}\mathbf{b}_{\mathbf{k}})|\phi\rangle = \langle\phi|\sum_{\mathbf{k}}(N_{\mathbf{a}}(\mathbf{k}) - N_{\mathbf{b}}(\mathbf{k}))|\phi\rangle\\ &= n_a - n_b.\end{aligned} \tag{78}$$

In our example,

$$\begin{aligned}&= \langle\phi_p 2\phi_q\phi_r|\frac{\sum_{\mathbf{k}}(N_{\mathbf{a}}(\mathbf{k}) - N_{\mathbf{b}}(\mathbf{k}))}{V}|\phi_p 2\phi_q\phi_r\rangle\\ &= \frac{1+2+1}{V}\end{aligned} \tag{79}$$

For a multi particle state $|\phi\rangle = |(A_{p1}\phi_{p1} + A_{p2}\phi_{p2} + ...), 2(A_{q1}\phi_{q1} + A_{q2}\phi_{q2} + ...), ...\rangle$ in which the particles are in general, not eigen, states, we have

$$\begin{aligned}\rho = \langle\phi|\rho_{oper}|\phi\rangle &= \langle\phi|\frac{\sum_{\mathbf{k}}(N_{\mathbf{a}}(\mathbf{k}) - N_{\mathbf{b}}(\mathbf{k}))}{V}|\phi\rangle\\ &= \frac{(A_{p1}^2 + A_{p2}^2 + A_{p3}^2 + ...) + 2(A_{q1}^2 + A_{q2}^2 + A_{q3}^2 + ...) + ...}{V}\\ &= \frac{1+2+...}{V} = \frac{\text{Total num of particles}}{V}.\end{aligned} \tag{80}$$

Again, we can consider the total particle number operator as having the form in (73).

5.2 Probability Density for Continuous Solutions in QFT

The continuous solution probability density and number operator relations are developed below in direct parallel with the discrete solutions development above. Note that, in analogy with (60)

$$\begin{aligned}\rho_{oper} = 2&\left(\int \frac{\mathbf{a}^\dagger(\mathbf{k}')}{\sqrt{2\omega_{\mathbf{k}'}}}\frac{e^{ik'x}}{\sqrt{(2\pi)^3}}d\mathbf{k}'\right)\left(\int \frac{\omega_{\mathbf{k}''}\mathbf{a}(\mathbf{k}'')}{\sqrt{2\omega_{\mathbf{k}''}}}\frac{e^{-ik''x}}{\sqrt{(2\pi)^3}}d\mathbf{k}''\right)\\ &-2\left(\int \frac{\mathbf{b}^\dagger(\mathbf{k}')}{\sqrt{2\omega_{\mathbf{k}'}}}\frac{e^{ik'x}}{\sqrt{(2\pi)^3}}d\mathbf{k}'\right)\left(\int \frac{\omega_{\mathbf{k}''}\mathbf{b}(\mathbf{k}'')}{\sqrt{2\omega_{\mathbf{k}''}}}\frac{e^{-ik''x}}{\sqrt{(2\pi)^3}}d\mathbf{k}''\right),\end{aligned} \tag{81}$$

and as before, the numerical probability density is the expectation value of (81),

$$\rho = \langle\phi(\tilde{x})|\rho_{oper}(x)|\phi(\tilde{x})\rangle. \tag{82}$$

5.2.1 Single Particle Wave Packet Probability Density (QFT, Continuous)

Consider $|\phi\rangle$ as a single particle wave packet where, if for simplicity we ignore antiparticles,

$$|\phi(\tilde{x})\rangle = \left|\int A(\mathbf{k}''')\frac{e^{-ik'''\tilde{x}}}{\sqrt{(2\pi)^3}}d\mathbf{k}'''\right\rangle, \tag{83}$$

where proper normalization for $A(\mathbf{k}''')$ is assumed. As an aside, note that (need to have defined **a**, $\mathbf{a}^+$, and $N_{\mathbf{a}}$ operator action on continuous ket before here)

$$\langle\phi(\tilde{x})|\mathbf{a}^\dagger(\mathbf{k}')\mathbf{a}(\mathbf{k}'')|\phi(\tilde{x})\rangle = A^\dagger(\mathbf{k}')A(\mathbf{k}'') \tag{84}$$

and this is only non-zero when $\mathbf{k}' = \mathbf{k}''$.

By analogy to (67), we have probability density

$$\rho = \frac{1}{\sqrt{(2\pi)^3}}\left(\int A^{\dagger}(\mathbf{k}')\frac{e^{ik'x}}{\sqrt{2\omega_{\mathbf{k}'}}}d\mathbf{k}'\right)\left(\int \omega_{\mathbf{k}''}\frac{e^{-ik''x}}{\sqrt{2\omega_{\mathbf{k}''}}}A(\mathbf{k}'')d\mathbf{k}''\right). \tag{85}$$

Integrating (85) over all space (and using the Dirac delta relation (18)) results in

$$\int \rho dV = \int \left|A(\mathbf{k}')\right|^2 d\mathbf{k}' = 1. \tag{86}$$

This equals the total probability of finding a single wave packet somewhere over all space and looks familiar to what we have seen in non-relativistic quantum mechanics. It also equals the number of particles. Thus we may define a number operator as

$$N = \int \rho_{oper} dV = \int \left(N_a(\mathbf{k}) - N_b(\mathbf{k})\right)d\mathbf{k} \tag{87}$$

in analogy with (73) for discrete solution states and consonant with (37).

5.2.2 Multiple Particle General State (QFT, Continuous)

Consider continuous solution multi particle states where the particles are in general (non eigen) states, i.e., $\left|\phi_p 2\phi_q\phi_r\right\rangle$, where, for example,

$$\left|\phi_p\right\rangle = \left|\int {}_p A(\mathbf{k}''')\frac{e^{-ik'''\tilde{x}}}{\sqrt{(2\pi)^3}}d\mathbf{k}'''\right\rangle, \tag{88}$$

and similar relations hold for the other particles in the multi particle state. By analogy to (75)

$$\rho = \frac{1}{(2\pi)^3}\left\{\begin{array}{l}\left(\int {}_p A^{\dagger}(\mathbf{k}')\frac{e^{ik'x}}{\sqrt{2\omega_{\mathbf{k}'}}}d\mathbf{k}'\right)_{\phi_p}\left(\int \omega_{\mathbf{k}''}\frac{e^{-ik''x}}{\sqrt{2\omega_{\mathbf{k}''}}}{}_p A(\mathbf{k}'')d\mathbf{k}''\right)_{\phi_p} \\ +\left(\int {}_q A^{\dagger}(\mathbf{k}')\frac{e^{ik'x}}{\sqrt{2\omega_{\mathbf{k}'}}}d\mathbf{k}'\right)_{\phi_q}\left(\int \omega_{\mathbf{k}''}\frac{e^{-ik''x}}{\sqrt{2\omega_{\mathbf{k}''}}}{}_q A(\mathbf{k}'')d\mathbf{k}''\right)_{\phi_q} + \ldots\end{array}\right\} \tag{89}$$

Thus,

$$\int \rho dV = \int \left|{}_p A_{\mathbf{k}'}\right|^2 d\mathbf{k}' + 2\int \left|{}_q A_{\mathbf{k}'}\right|^2 d\mathbf{k}' + \int \left|{}_r A_{\mathbf{k}'}\right|^2 d\mathbf{k}' = 1+2+1 = 4. \tag{90}$$

This, again, equals the number of particles and hence

$$\begin{aligned}&\text{Number of particles} = \int \rho dV = \int \langle\phi|\rho_{oper}(x)|\phi\rangle d^3\mathbf{x} \\ &= \text{Expectation value of number operator} \\ &= \langle\phi|\int\left(\mathbf{a}^{\dagger}(\mathbf{k})\mathbf{a}(\mathbf{k}) - \mathbf{b}^{\dagger}(\mathbf{k})\mathbf{b}(\mathbf{k})\right)d\mathbf{k}|\phi\rangle = \langle\phi|\int\left(N_{\mathbf{a}}(\mathbf{k}) - N_{\mathbf{b}}(\mathbf{k})\right)d\mathbf{k}|\phi\rangle \\ &= n_a - n_b.\end{aligned} \tag{91}$$

6. Action of Hamiltonian on States (QFT)

6.1 Discrete Eigenstates (QFT)

We treat only the general state particle case, as the eigen state particle case is a special case where all $A_{\mathbf{k}}$ are zero except one, with that one having an absolute value (modulus) of 1.

6.1.1 General Single Particle State (QFT, Discrete)

For the Hamiltonian of (31), and again concentrating for simplicity only on particles (and not anti-particles) kets, the energy expectation value is

$$\begin{aligned}\bar{E} &= \langle\phi|H|\phi\rangle \\ &= \left\langle \sum_{\mathbf{k}} A_{\mathbf{k}}^{\dagger}\frac{e^{ik\tilde{x}}}{\sqrt{V}}\middle|\sum_{\mathbf{k}'}\omega_{\mathbf{k}'}\left(N_a(\mathbf{k}')+\tfrac{1}{2}+N_b(\mathbf{k}')+\tfrac{1}{2}\right)\middle|\sum_{\mathbf{k}''}A_{\mathbf{k}''}\frac{e^{ik''\tilde{x}}}{\sqrt{V}}\right\rangle \\ &= \sum_{\mathbf{k}}A_{\mathbf{k}}^{\dagger}\sum_{\mathbf{k}'}\omega_{\mathbf{k}'}A_{\mathbf{k}'}\delta_{\mathbf{k}\mathbf{k}'} = \sum_{\mathbf{k}}\left|A_{\mathbf{k}}\right|^2\omega_{\mathbf{k}}.\end{aligned} \tag{92}$$

In the last line we ignored the ½$\omega_{\mathbf{k}}$ contributions from the vacuum.

For an eigenstate $|\phi\rangle = |\phi_{\mathbf{k}}\rangle$ all but one coefficient in (92) equals zero, and we have $E = \bar{E} = \omega_{\mathbf{k}}$.

6.1.2 General Multi Particle State (QFT, Discrete)

For multi particle states where at least some of the particles are in general states, we have

$$\bar{E} = \langle\phi|H|\phi\rangle = \left\langle \phi_p 2\phi_q\phi_r\middle|\sum_{\mathbf{k}'}\omega_{\mathbf{k}'}\left(N_a(\mathbf{k}')+\tfrac{1}{2}+N_b(\mathbf{k}')+\tfrac{1}{2}\right)\middle|\phi_p 2\phi_q\phi_r\right\rangle. \tag{93}$$

With the Hamiltonian operator acting on the ket as the $\mathbf{a}_{\mathbf{k}}$ operator did in (71), this results in

$$\bar{E} = \sum_{\mathbf{k}}\omega_{\mathbf{k}}\left(\left|{}_pA_{\mathbf{k}}\right|^2 + 2\left|{}_qA_{\mathbf{k}}\right|^2\left|{}_rA_{\mathbf{k}}\right|^2\right) = \bar{E}_p + 2\bar{E}_q + \bar{E}_r, \tag{94}$$

wherein the total expected energy value equals the sum of the expectation energies for each particle in the state, and we have again ignored the vacuum contribution.

For all particles in eigenstates of energy, this reduces to $E = \omega_p + 2\omega_q + \omega_r$.

6.2 Continuous Eigenstates

6.2.1 Single Particle Wave Packet State (QFT, Continuous)

For a single particle wave packet, the ket has form

$$|\phi(\tilde{x})\rangle = \left|\int A(\mathbf{k}''')\frac{e^{-ik'''\tilde{x}}}{\sqrt{(2\pi)^3}}d\mathbf{k}'''\right\rangle, \tag{95}$$

and the energy expectation value is

$$\begin{aligned}\bar{E} &= \langle\phi|H|\phi\rangle \\ &= \left\langle\int A^{\dagger}(\mathbf{k})\frac{e^{ik\tilde{x}}}{\sqrt{(2\pi)^3}}d\mathbf{k}\middle|\int\omega_{\mathbf{k}'}\left(N_a(\mathbf{k}')+\tfrac{1}{2}+N_b(\mathbf{k}')+\tfrac{1}{2}\right)d\mathbf{k}'\middle|\int A(\mathbf{k}''')\frac{e^{-ik'''\tilde{x}}}{\sqrt{(2\pi)^3}}d\mathbf{k}'''\right\rangle.\end{aligned} \tag{96}$$

Noting that each $N_a(\mathbf{k}')$ operator acting on the ket leaves zero except when $\mathbf{k}' = \mathbf{k}'''$, we have

$$\bar{E} = \left\langle\int A^{\dagger}(\mathbf{k})\frac{e^{ik\tilde{x}}}{\sqrt{(2\pi)^3}}d\mathbf{k}\middle|\middle|\int\omega_{\mathbf{k}'}A(\mathbf{k}')\frac{e^{-ik'\tilde{x}}}{\sqrt{(2\pi)^3}}d\mathbf{k}'\right\rangle. \tag{97}$$

Integrating this over all space and using the Dirac delta function again, we end up with

$$\bar{E} = \int\omega_{\mathbf{k}}A^{\dagger}(\mathbf{k})A(\mathbf{k})d\mathbf{k} = \int\omega_{\mathbf{k}}\left|A(\mathbf{k})\right|^2 d\mathbf{k} \tag{98}$$

in complete analogy with (92).

6.2.2 Multi Particle Wave Packet States (QFT, Continuous)

For a multiparticle state where the particles are wave packets,

$$\begin{aligned}\bar{E} &= \langle\phi|H|\phi\rangle \\ &= \left\langle\phi_p 2\phi_q\phi_r\middle|\int\omega_{\mathbf{k}'}\left(N_a(\mathbf{k}')+\tfrac{1}{2}+N_b(\mathbf{k}')+\tfrac{1}{2}\right)d\mathbf{k}'\middle|\phi_p 2\phi_q\phi_r\right\rangle\end{aligned} \tag{99}$$

and once again we have the operator acting sequentially on each particle in the ket. Ignoring the vacuum contribution, this results in a series of terms like (98), i.e., XXX Think thru XXX

$$\bar{E} = \int \omega_{\mathbf{k}} \left(\left|{}_{p}A(\mathbf{k})\right|^2 + 2\left|{}_{q}A(\mathbf{k})\right|^2 + \left|{}_{r}A(\mathbf{k})\right|^2 \right) d\mathbf{k} = \bar{E}_p + 2\bar{E}_q + \bar{E}_r \,. \tag{100}$$

Thus, the expected energy is the sum of the expected energies for each wave packet particle.

7. Action of Hamiltonian on the Vacuum

Rough thoughts only as of March 13, 2010..