*Chapter 2*               ***vers 8/23/22 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 $\varepsilon_o$ and permeability $\mu_o$ 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 $4\pi$ 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 massive particle is equal to its mass (at rest) $m$ 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 $3 \times 10^{10}$ cm/s.

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 $1.0545 \times 10^{-27}$ 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 ($\mathbf{F}=m\mathbf{a}$ for force, etc.). We only use three laws here for simplicity and brevity.

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

### 2.1.3 Deducing Natural Units

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

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

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

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

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 <u>mega-electron-volts</u>, i.e., <u>MeV</u> = 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 0.511 MeV rest energy also has 0.511 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 $\omega$ 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

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

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

### 2.1.5 Converting from One System to Another

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

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

$t = 10^{16}$ (MeV)$^{-1} \times \hbar = 10^{16}$ (MeV)$^{-1}$  where $\hbar =1$, 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,

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

The same time interval is described as either $10^{16}$ (MeV)$^{-1}$ or $6.58 \times 10^{-6}$ 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.

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

### 2.1.6 Mass and Energy in the Hybrid and Natural Systems

As mentioned, the hybrid system is not the same as the cgs system, even though both use centimeters and seconds. In the cgs system, energy is measured in ergs and mass in grams. In the hybrid system, energy is measured in MeV and mass in unfamiliar, and never used, units. (See Wholeness Chart 2-1 below.) It may be confusing, but when experimentalists talk of mass, energy, length, and time, they like to use the hybrid system, *yet* they commonly refer to mass in MeV. For example, in high energy physics, the mass of the electron is commonly referred to as 0.511 MeV, rather than hybrid (unfamiliar) or cgs (gram) mass units. Hopefully, Wholeness Chart 2-1 will help to keep all of this straight.

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

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

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

| Natural Units | | Hybrid Units | | cgs Units | |
|---|---|---|---|---|---|
| $c = \hbar = 1$ | | $c = 2.99 \times 10^{10}$ cm/s<br>$\hbar = 6.58 \times 10^{-22}$ MeV-s<br>$\hbar\,c = 1.973 \times 10^{-11}$ MeV-cm | | conversion factor<br>$F = 1.602 \times 10^{-6}$ ergs/MeV | |
| **Quantity,**<br>**units of $(\text{MeV})^M$** | **$M$** | **Multiply ← value**<br>**by ↓ to get →** | **in**<br>**MeV-cm-s** | **Multiply ← value**<br>**by ↓ to get →** | **in**<br>**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 | $m\mathbf{a}$ 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) $\times$ ?, where ? must be $c^2$.

Conversion of numeric quantities

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

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

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

## 2.1.8 QFT Approach to Units

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

In <u>summary</u>:

relations in cgs $\rightarrow$ same relations in natural units $\rightarrow$ develop theory in natural units $\rightarrow$ predict experiment in natural units $\rightarrow$ same predictions in hybrid (MeV-cm-s) 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 <u>contravariant components</u> $x^\mu$ of the 4D position vector as 3D Cartesian coordinates $X_i$ plus $ct$ (see Appendix A 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

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

If we define <u>covariant components</u> 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

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

where on the RHS, we have introduce the shorthand <u>Einstein summation convention</u>, 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.,

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

*Getting covariant components from contravariant ones*

where the matrix $\underline{g_{\mu\nu}}$ is known as the <u>metric tensor</u>. Its <u>inverse</u>, $g^{\mu\nu}$, 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 $\underline{\delta_\alpha{}^\nu}$ is the <u>Kronecker delta</u> ( = 0 if row $\alpha \neq$ column $\nu$; = 1 if $\alpha = \nu$) . With the metric tensor and its inverse, we can re-write (2-4) as

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

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

*Contravariant and covariant derivatives*

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

Note the spatial parts of $x^\mu$ and $\partial^\mu$ have opposite signs.

*Raising and lowering indices*

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

For a matrix (a tensor using two indices), rather than a column quantity (vector with one index), we can use $g^{\mu\nu}$ to raise (or $g_{\mu\nu}$ to lower) either index, or use $g^{\mu\nu}$ twice for both indices. For example, for the matrix (tensor) $M_{\alpha\beta}$, we would have $M^{\mu\nu} = g^{\mu\alpha}g^{\nu\beta}M_{\alpha\beta}$ .

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

*Script → density*

Further, as one repeatedly sums $p_\mu$ and $x^\mu$ 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.)

*$p_\mu x^\mu = px$*

## 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.

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.)

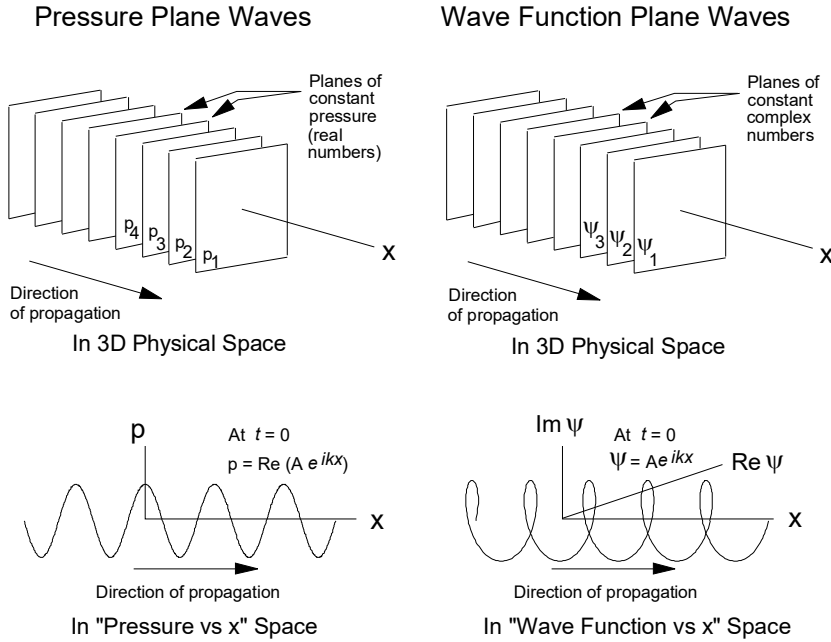*Real vs. complex (quantum) plane waves*

## Pressure Plane Waves          Wave Function Plane Waves



**Figure 2-1.  Classical vs Quantum Plane Waves**

## 2.4 Review of Variational Methods

### 2.4.1 Classical Particle Theory

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

*Definition of classical mechanics Lagrangian L*

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

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

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

*Governing equation = Euler-Lagrange equation*

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.

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

*Legendre transformation $H \leftrightarrow L$*

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

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

It is an <u>important point</u> 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.

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

<u>Equivalent entities</u>

Lagrangian *L*  $\leftrightarrow$  equations of motion  $\leftrightarrow$  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

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.

*Variational math can be applied to many diverse areas in physics and elsewhere*

### 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).]

<div align="center">Particle Theory → Field Theory</div>

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

*Analogous entities in particle and field theories*

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

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

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

*Intuitive deduction of field relations from particle ones*

The <u>Euler-Lagrange equation for fields</u> 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 <u>Legendre transformation for the Hamiltonian density</u>, with $\pi_r$ being the conjugate momentum density of the field $\phi^r$, 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.

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.

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

<u>Derivation of Euler-Lagrange Equation for Fields</u>

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

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

*Formal derivation of Euler-Lagrange equation for fields*

where $d^4x = d^3\mathbf{x} dt$ is an element of 4D volume, is stationary. More precisely, consider a virtual variation in $\phi$ 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 $\Omega$, i.e., $\delta\phi = 0$ on $\Gamma$. 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 $\Omega$ large enough so the field $\phi$ 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 3D "surface" $\Gamma$, as we show under the downward pointing bracket. In that integral, $n_\mu$ is the unit length 4D vector normal to the 3D surface $\Gamma$ at every point on the surface, and it forms an inner product with the quantity in brackets by virtue of the summation over $\mu$. 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),

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

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

End of derivation

### 2.4.4 Real vs. Complex Fields

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

*Classical field real; quantum fields usually complex*

### 2.5 Classical Mechanics: An Overview

Wholeness Chart 2-2 is a summary of the key relations in all of classical physical theory (from the variational viewpoint.) The chart is intended primarily as an overview of past courses and as a lead in to quantum field theory, so a detailed study of it is not really warranted at this time. We have

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

**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 | $L = L(q_i, \dot{q}_i, t)$ | $L = L\left(x^i, \dot{x}^i, t\right) = \sum_i \frac{1}{2}m\left(\dot{x}^i\right)^2 - V\left(x^i, t\right)$ |
| Action | $S = \int L dt$ | as at left |
| Euler- Lagrange equation (From $\delta S = 0$.) | $\dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{q}_i}\right) - \dfrac{\partial L}{\partial q_i} = 0$ | $\dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{x}^i}\right) - \dfrac{\partial L}{\partial x^i} = 0$ |
| Equations of motion for chosen coordinates | use explicit form for $L$ in Euler-Lagrange equation | $m\ddot{x}^i = -\dfrac{\partial V}{\partial x^i}$ usually $V$ not function of $t$ |
| Conjugate momentum density; total | see Fields columns; $p_i = \dfrac{\partial L}{\partial \dot{q}_i}$ | n/a ; $p_i = \dfrac{\partial L}{\partial \dot{x}^i} = m\dot{x}^i$ $\left(= p^i \text{ for Cartesian}\right)$ |
| 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 ; $H = p_i\dot{x}^i - L = p^2/2m + V$ |
| Hamilton's Equations of Motion for conjugate variables | $\dot{p}_i = -\dfrac{\partial H}{\partial q_i}$     $\dot{q}_i = \dfrac{\partial H}{\partial p_i}$ | $\dot{p}_i = -\dfrac{\partial H}{\partial x^i} = -\dfrac{\partial V}{\partial x^i}$    $\dot{x}^i = \dfrac{\partial H}{\partial p_i}$ |
| Poisson Brackets, definition | for $u = u(q_i, p_i, t)$, $v = v(q_i, p_i, t)$ <br> $\{u,v\} = \dfrac{\partial u}{\partial q_i}\dfrac{\partial v}{\partial p_i} - \dfrac{\partial u}{\partial p_i}\dfrac{\partial v}{\partial q_i}$ | for $u = u(x^i, p_i, t)$, $v = v(x^i, p_i, t)$ <br> $\{u,v\} = \dfrac{\partial u}{\partial x^i}\dfrac{\partial v}{\partial p_i} - \dfrac{\partial u}{\partial p_i}\dfrac{\partial v}{\partial x^i}$ |
| Equations of motion in terms of Poisson brackets <br> i) any variable <br> ii) conjugate variables | i) for $v = H$   $\dfrac{du}{dt} = \{u, H\} + \dfrac{\partial u}{\partial t}$ <br><br> ii) for i) plus $u = q_i$ or $p_i$ <br><br> $\dot{p}_i = \{p_i, H\} = -\dfrac{\partial H}{\partial q_i}$; $\dot{q}_i = \{q_i, H\} = \dfrac{\partial H}{\partial p_i}$ | i) for $v = H$   $\dfrac{du}{dt} = \{u, H\} + \dfrac{\partial u}{\partial t}$ <br><br> ii) for i) plus $u = x^i$ or $p_i$ <br><br> $\dot{p}_i = \{p_i, H\} = -\dfrac{\partial H}{\partial x^i}$; $\ddot{x}^i = \{x^i, H\} = \dfrac{\partial H}{\partial p_i}$ |
| 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 |
|---|---|---|
| $x^i, t \qquad i = 1, 2, 3$ | $t$ | $x^\mu \qquad \mu = 0, 1, 2, 3$ |
| $\phi^r(x^i, t) \quad r = \text{field type} = 1, \dots, n$ | $x^i = x^i(t), \quad i = 1, 2, 3$ | $\phi^r(x^\mu) \quad r = \text{field type} = 1, \dots, n$ |
| $\mathcal{L} = \mathcal{L}\left(\phi^r, \dot{\phi}^r, \partial_i \phi^r, x^i, t\right)$ | not applicable for particle | $\mathcal{L} = \mathcal{L}\left(\phi^r, \partial_\mu \phi^r, x^\mu\right)$ |
| $L = \int \mathcal{L} d^3 x$ | $L\left(x^i, v^i, t\right) = -m\sqrt{1 - v^2} - V$ | $L = \int \mathcal{L} d^3 x$ |
| $S = \int L dt = \int \mathcal{L} d^3 x dt$ | $S = \int L dt$ | $S = \int L dt = \int \mathcal{L} d^3 x dt$ |
| $\dfrac{d}{dt}\left(\dfrac{\partial \mathcal{L}}{\partial \dot{\phi}^r}\right) + \dfrac{d}{dx^i}\left(\dfrac{\partial \mathcal{L}}{\partial \phi^r_{,i}}\right) - \dfrac{\partial \mathcal{L}}{\partial \phi^r} = 0$ | $\dfrac{d}{dt}\left(\dfrac{\partial L}{\partial v^i}\right) - \dfrac{\partial L}{\partial x^i} = 0$ | $\dfrac{\partial}{\partial x^\mu}\left(\dfrac{\partial \mathcal{L}}{\partial \phi^r_{,\mu}}\right) - \dfrac{\partial \mathcal{L}}{\partial \phi^r} = 0$ |
| $\mathcal{L}$ above in Euler-Lagrange equation | $\dfrac{d}{dt}\left(\dfrac{\partial L}{\partial v^i}\right) = -\dfrac{\partial V}{\partial x^i} \; ; \; V\left(x^i, v^i\right)$ | $\mathcal{L}$ above in Euler-Lagrange equation |
| $\pi_r = \dfrac{\partial \mathcal{L}}{\partial \dot{\phi}^r} \; ; \qquad \Pi_r = \int \pi_r d^3 x$ | $\text{n/a} \; ; \; p^i = \dfrac{\partial L}{\partial v^i} = \dfrac{m v^i}{\sqrt{1 - v^2}} - \dfrac{dV}{dv^i}$ | $\pi_r = \dfrac{\partial \mathcal{L}}{\partial \dot{\phi}^r} \; ; \qquad \Pi_r = \int \pi_r d^3 x$ |
| $\not{p}_i = \pi_r \dfrac{\partial \phi^r}{\partial x^i} \; ; \quad p_i = \int \not{p}_i d^3 x$ | $\text{n/a} \; ; \; = \text{conjugate momentum}$ | $\not{p}_i = \pi_r \dfrac{\partial \phi^r}{\partial x^i} \; ; \quad p_i = \int \not{p}_i d^3 x$ |
| $\mathcal{L} = \mathcal{L}\left(\phi^r, \pi_r, \partial_i \phi^r, x^i, t\right)$ | $L = L(x^i, p^i, t)$ | $\mathcal{L} = \mathcal{L}\left(\phi^r, \pi_r, \partial_i \phi^r, x^i, t\right)$ |
| $\mathcal{H} = \pi_r \dot{\phi}^r - \mathcal{L} \; ; \; H = \int \mathcal{H} d^3 x$ | $\text{n/a} \; ; \; H = p^i v^i - L = T + V$ | $\mathcal{H} = \pi_r \dot{\phi}^r - \mathcal{L} \; ; \; H = \int \mathcal{H} d^3 x$ |
| same form as Relativistic Fields | $\dot{p}^i = -\dfrac{\partial H}{\partial x^i} = -\dfrac{\partial V}{\partial x^i} \quad \dot{x}^i = \dfrac{\partial H}{\partial p^i}$ | $\dot{\pi}_r = -\dfrac{\delta \mathcal{H}}{\delta \phi^r} \qquad \dot{\phi}^r = \dfrac{\delta \mathcal{H}}{\delta \pi_r}$ <br> where $\dfrac{\delta}{\delta \phi^r} = \dfrac{\partial}{\partial \phi^r} - \dfrac{\partial}{\partial x^i}\left(\dfrac{\partial}{\partial \phi^r_{,i}}\right)$ |
| same form as Relativistic Fields | same form as Non-relativistic Particle, but different meaning for $p^i$ | for $u = u(\phi^r, \pi_r, \partial_i \phi^r, t), \; v = v(\phi^r, \pi_r, \partial_i \phi^r, t)$ <br> $\{u, v\} = \left(\dfrac{\delta u}{\delta \phi^r}\dfrac{\delta v}{\delta \pi_r} - \dfrac{\delta u}{\delta \pi_r}\dfrac{\delta v}{\delta \phi^r}\right)\delta(\mathbf{x} - \mathbf{y})$ |
| same form as Relativistic Fields | same form as Non-relativistic Particle | i) for $U = \int u dV; \; \{U, H\} = \iint \{u, \mathcal{H}\} d^3 y d^3 \mathbf{x}$ <br> $\dot{U} = \dfrac{dU}{dt} = \{U, H\} + \dfrac{\partial U}{\partial t}$ <br> ii) for $u = \pi_r; \quad \dot{\Pi}_r = \{\Pi_r, H\}$ |
| 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, due to subtleties in the theory, non-relativistic chart relationships are most easily, and best considered at this point, expressed in Cartesian coordinates, where $x^i \to X_i$ and $p_i = p^i$.

### 2.5.1 Key Concepts in Field Theory

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.

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

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.

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

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.

   *$r$ label = different field types or different components of field*

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

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

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

5. For fields,

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

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

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

---

**Box 2-1. Time Derivatives and Fields**

Any field, say $\phi$, 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.) 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: $\phi^r$*

---

**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 by assuming our physical 3-momentum density $\not{p}^i$ obeys the classical field variational relation of the RHS of (B2-2.1). (This can be intuited from (2-11), except that there we used a Cartesian system where $p_i = p^i$, and here we use the relativistic Minkowski metric system, where $p_i = -p^i$.) If we divide the particle relation by volume, we get a density relation.

$$p_i = \frac{\partial L}{\partial \dot{x}^i} \qquad \xrightarrow[\text{divide by particle volume}]{\text{for small particle in medium,}} \qquad \not{p}_i = \frac{\partial \mathcal{L}}{\partial \dot{x}^i} \,. \qquad (B2\text{-}2.1)$$

For continuous media like a fluid, $\dot{x}^i$ is the velocity of the medium (field) at the point where $\not{p}_i$ is measured. We note carefully that our $x^i$ here is the position coordinate of a point fixed relative to the field (fluid particle in our example) and thus is time dependent. (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 $\dot{x}^i = dx^i/dt$ 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.

Now take the conjugate momentum density relation for relativistic fields (2-14),

$$\pi_r = \frac{\partial \mathcal{L}}{\partial \dot{\phi}^r} \,, \qquad (B2\text{-}2.2)$$

and divide the RHS of (B2-2.1) by (B2-2.2),

$$\frac{\not{p}_i}{\pi_r} = \frac{\partial \mathcal{L}/\partial \dot{x}^i}{\partial \mathcal{L}/\partial \dot{\phi}^r} = \frac{\partial \dot{\phi}^r}{\partial \dot{x}^i} = \frac{\partial \phi^r/\partial t}{\partial x^i/\partial t} = \frac{\partial \phi^r}{\partial x^i} \quad \rightarrow \quad \not{p}_i = \pi_r \frac{\partial \phi^r}{\partial x^i} \quad \rightarrow \quad \not{p}^i = -\pi_r \frac{\partial \phi^r}{\partial x^i} \,. \qquad (B2\text{-}2.3)$$

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

---

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

*Several ways to formulate variational relativistic theory*

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

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

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

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

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

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

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

*Poisson bracket definition used in equation of motion*

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

Summary of Forms of Differential Equations of Motion

For generalized coordinates (all three below are equivalent)

1. Lagrangian into Euler-Lagrange equation

2. Hamilton's equations of motion

3. Poisson bracket notation for 2 above

*Forms for differential equations of motion*

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

1. Total time derivative expressed as partial derivatives (see (2-23), not shown in Wholeness Chart 2-2.)

2. Total time derivative expressed in terms of Poisson bracket notation (see (2-24), also shown in Wholeness Chart 2-2.)

11. (See Appendix A 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[1], i.e., they are <u>Lorentz invariant</u>. 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 <u>Lorentz covariant</u>. Four scalars (world scalars) are invariant under a Lorentz transformation and look exactly the same to any observer. (e.g., <u>Rest mass *m*</u> [or simply <u>mass *m*</u> as it is more commonly called in relativity] of a free

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

---

[1] 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 $p'_\mu$, but would find the same mass via $p'^\mu p'_\mu = m^2$.)

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, $x^\mu, \phi^r$, and derivatives of $x^\mu$ are Lorentz covariant or invariant. So, in order for the whole equation to be Lorentz invariant, the Lagrangian density $\mathcal{L}$ must be invariant, i.e., a world scalar.

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

Since $d^4x$ is also a <u>Lorentz (world or four-) scalar</u> (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.

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

<u>End of Key Concepts in Field Theory points</u>

## 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.

These different, but equivalent, ways are called different <u>pictures</u> 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 <u>Schrödinger picture</u>, 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 <u>Heisenberg picture</u>, 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.,

*Different pictures in quantum theory*

    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

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 <u>expectation value</u>. 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 $\psi$)

*Operator expectation value = "expected" or mean measurement*

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

*Calculating expectation value*

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 (see Appendix B, Section 2.10.1, if you are concerned about switching total derivatives for partial derivatives below)

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

*Eq of motion of expectation value*

In the Schrödinger picture, the solutions to the Schrödinger equation

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

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

are the states $\psi_S$ (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

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

*An example*

(2-25) is

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

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

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

Equation (2-26) for $p_1$ is then

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

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

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

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

*Eq of motion of momentum expectation value*

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 <u>unitary transformation</u> (see Box 2-3). The particular unitary transformation (where $U$ is a <u>unitary operator</u>) for this is

$$U = e^{-iHt} \qquad \left(= e^{-iHt/\hbar} \text{ in non-natural units}\right), \tag{2-33}$$

where states and operators transform as

*Transforming between Schrödinger and Heisenberg pictures*

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

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,

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

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

Now taking the time derivative of the second relation in the top row of (2-34), we have (see Appendix B, Section 2.10.2, if you are concerned about switching total to partial derivatives)

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

←Note the hat on $\partial$ defintion

(2-37)   *In H.P., eq of motion of operator*

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

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 $\mathbf{A}$ has an inverse equal to the transpose of that matrix, i.e., $\mathbf{A}^{-1}=\mathbf{A}^{T}$. In the complex space of state vectors, a unitary transformation $U$ has an analogous form for its inverse, the complex conjugate transpose, i.e., $U^{-1}=U^{\dagger}$ and so $U^{\dagger}U=1$. The following example may make this clearer.

Consider $U=e^{-iHt}$, where $H$ is the (hermitian) 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 $\mathbf{A}$ = matrix | Unitary $U=e^{iX}$ |
| **Effect on vector** | rotates in real space | "rotates" in complex space |
| **Physical effect** | vector length unchanged | probability unchanged |
| **Inverse** | $\mathbf{A}^{-1}=\mathbf{A}^{T}$ | $U^{-1}=U^{\dagger}$ |

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.,

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

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. 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

$$\frac{d\bar{\mathcal{O}}}{dt} = {}_S\langle\psi|UU^\dagger\left(-i\left[\mathcal{O}^S,H\right]\right)UU^\dagger|\psi\rangle_S + {}_S\langle\psi|UU^\dagger\frac{\partial\mathcal{O}^S}{\partial t}UU^\dagger|\psi\rangle_S$$

$$= {}_H\langle\psi|\left(-i\left[\mathcal{O}^H,H\right]\right)|\psi\rangle_H + {}_H\langle\psi|\frac{\hat{\partial}\mathcal{O}^H}{\partial t}|\psi\rangle_H\ .\tag{2-39}$$

From which we see that the equation of motion for the expectation value of an operator has 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 $i\dfrac{d}{dt}|\psi\rangle_S = H|\psi\rangle_S$ (Schrödinger eq) | Usually time independent $\dfrac{d\mathcal{O}^S}{dt} = \underbrace{\dfrac{\partial\mathcal{O}^S}{\partial t} = 0}_{\text{usually}}$ | $\dfrac{d\bar{\mathcal{O}}}{dt} = {}_S\langle\psi|\left(-i\left[\mathcal{O}^S,H\right]+\dfrac{\partial\mathcal{O}^S}{\partial t}\right)|\psi\rangle_S$ $\|\psi\rangle_S$ changes in time; $\mathcal{O}^S$ usually const in time |
| Transform via $U = e^{-iHt/\hbar}$ $\Downarrow$ | $U^\dagger|\psi\rangle_S = |\psi\rangle_H$ | $U^\dagger\mathcal{O}^S U = \mathcal{O}^H$ | $\dfrac{d\bar{\mathcal{O}}}{dt}$ invariant under the transformation |
| **Heisenberg Picture** | Time independent $\dfrac{d|\psi\rangle_H}{dt} = 0$ | Often time dependent $\dfrac{d\mathcal{O}^H}{dt} = -i\left[\mathcal{O}^H,H\right]+\underbrace{\dfrac{\hat{\partial}\mathcal{O}^H}{\partial t}}_{\substack{\text{usually}\\=0}}$ | Same as Schrödinger picture above with sub and superscript $S \to H$ and $\partial\mathcal{O}^H = \hat{\partial}\mathcal{O}^H$ $\|\psi\rangle_H$ 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 what move and evolve. 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.

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

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

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 summarizes QM*

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.

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

### 2.7.1 Classical vs. Quantum: Much is the Same

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.)

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

### 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

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

in NRQM in the Heisenberg picture. 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 fundamental 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 significant difference. All relations remain effectively the same except that the commutators of quantum theory correspond to Poisson brackets of classical theory (times a factor of $-i/\hbar$.)

### 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

| | **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<br>  i) any dynamical variable<br>  ii) conjugate variables | Correspondence principle:<br>Classical $\rightarrow$ Quantum $$\{u,v\} \rightarrow \frac{-i}{\hbar}[u,v]$$ | i) for $v = H$   $\dfrac{du}{dt} = \dfrac{-i}{\hbar}[u,H] + \dfrac{\partial u}{\partial t}$<br><br>ii) for i) plus $u = x^i$ or $p_i$<br><br>$\dot{p}_i = \dfrac{-i}{\hbar}[p_i, H] = -\dfrac{\partial H}{\partial x^i}$;   $\dot{x}^i = \dfrac{-i}{\hbar}[x^i, H] = \dfrac{\partial H}{\partial p_i}$ |
| Uncertainty principle | | $\left[x^i, p_j\right] = i\hbar\delta^i{}_j$    $\left[x^i, x^j\right] = \left[p_i, p_j\right] = 0$ |

$$\underbrace{\left\{x^i, p_j\right\}}_{\substack{\text{classical} \\ \text{dynamic} \\ \text{variables}}} = \delta^i{}_j = \underbrace{\frac{-i}{\hbar}\left[x^i, p_j\right]}_{\substack{\text{quantum} \\ \text{operators}}}. \qquad \begin{pmatrix} \text{Cartesian system, where} \\ p_j = p^j = \text{3-momentum} \end{pmatrix} \qquad (2\text{-}42)$$

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

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

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

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*

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

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 <u>canonical quantization</u>. 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 $u = u\,(\phi^r, \pi_r, \partial_i\phi^r, t\,)$, $v = v\,(\phi^r, \pi_r, \partial_i\phi^r, t\,)$ <br><br> $[u,v] = uv - vu$ |
| | See Non-relativistic Quantum Mechanics section | i) for $U = \int u\,dV$; $[U,H] = UH - HU$ <br><br> $\dot{U} = \dfrac{dU}{dt} = [U,H] + \dfrac{\hat{\partial}U}{\partial t}$ <br><br> ii) for $u = \pi_r$; $\dot{\Pi}_r = [\Pi_r, H]$ |
| | 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 righthand 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 A: 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 <u>Minkowski space</u> (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.,

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

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

where, with a future purpose in mind, we insert an identity matrix, represented in index notation by the Kronecker delta $\delta_{ij}$ ( = 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 $\tau$ on an object passing by us. The $0^{th}$ coordinate is now time instead of a spatial $X_0$ coordinate. From SRT, we know

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

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

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, $x^\mu$ was designated as the <u>contravariant components</u> form, and $x_\mu$ as the <u>covariant components</u> 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 $x^\mu$. For example, the <u>four-velocity of relativity $u^\mu$</u> 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 $\tau$, $v^i$ is that with respect to coordinate time $t$, $\gamma$ 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

$$(u)^2 = |u^\mu|^2 = u^\mu u_\mu = \frac{dx^\mu}{d\tau}\frac{dx_\mu}{d\tau} = \begin{bmatrix} u^0 & u^1 & u^2 & u^3 \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} u^0 & u^1 & u^2 & u^3 \end{bmatrix} \begin{bmatrix} u^0 \\ -u^1 \\ -u^2 \\ -u^3 \end{bmatrix}$$

(2-54)

$$= \left(u^0\right)^2 - \left(u^1\right)^2 - \left(u^2\right)^2 - \left(u^3\right)^2 = \gamma^2\left(c^2 - \left(v^1\right)^2 - \left(v^2\right)^2 - \left(v^3\right)^2\right) = c^2,$$

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 <u>4-momentum $p^\mu = mu^\mu$</u> is then found from

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

(2-55) tells us that for (massless) photons $(p)^2 = 0$, even though $p^\mu \neq 0$. (See Prob. 13.) Note from (2-55) that $p^0 = \gamma mc = E/c$, where $E$ is relativistic energy, and $p^i$ = relativistic 3-momentum.

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

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

In addition, we will often use differential elements of 4 vectors, such as $dx^\mu$, 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 contravariant 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 <u>metric</u> (of the coordinate space, which in this case is <u>Minkowski coordinate space</u>.) It lowers a raised index. It has an inverse that turns out to have the same form as it does.

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

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 <u>covariant metric</u>, and $g^{\mu\nu}$, the <u>contravariant metric</u>. The term *metric* used alone usually means $g_{\mu\nu}$.

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

$$\left(c\tau\right)^2 = x^\mu x_\mu = g_{\mu\nu}x^\mu x^\nu. \qquad (2\text{-}60)$$

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 $g_{\mu\nu}$. Note that in general relativity, you will find the Minkowski metric, which is commonly designated by $g_{\mu\nu}$ in QFT, to be designated by the symbol $\eta_{\mu\nu}$. In relativity, $g_{\mu\nu}$ usually refers to any general metric, not necessarily of form shown in (2-57). But in this book, the symbol $g_{\mu\nu}$ always equals $\eta_{\mu\nu}$, the Minkowski metric.

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, so for any vector **v**, $v^i = v_i$. (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 <u>4D scalar</u> (or <u>world scalar</u> or <u>Lorentz scalar</u>.) 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 <u>invariant</u> under transformation to a new coordinate system.

The quantity $c\tau$, or simply the proper time $\tau$ passed on an object, is the same for all observers, is invariant in 4D spacetime, and hence is a scalar. $\left(c\tau\right)^2 = x^\mu x_\mu = x'^\mu x_\mu'$, even though $x'^\mu \neq x^\mu; x_\mu' \neq x_\mu$. The term <u>Lorentz invariance</u> 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 $x^\mu$, $u^\mu$, and $p_\mu$) are said to be <u>covariant</u> under transformation to a new coordinate system. For spacetime, the term <u>Lorentz covariance</u> 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 <u>Lorentz invariance of the laws of nature</u>. 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 <u>differential geometry</u> (or <u>tensor analysis</u>, or in the old days, <u>Riemannian geometry</u>.) 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 A provides some justification for their use, which is widespread in QFT.

## 2.10 Appendix B: Partial vs Total Derivatives

### 2.10.1 For Relations Like (2-26)

In equation (2-26), one might think that, according to the product differentiation rule, the factor $\dfrac{\partial \mathcal{O}^S}{\partial t}$ should be a total derivative, as in $\dfrac{d\mathcal{O}^S}{dt}$, rather than a partial derivative. That is, we would expect the equation to look like (2-61), where the second line comes from Box 2-1, pg. 22.

$$\frac{d\bar{\mathcal{O}}}{dt} = \frac{d}{dt}\langle \psi|\mathcal{O}|\psi\rangle = \left\langle \frac{d\psi}{dt}\middle|\mathcal{O}\middle|\psi\right\rangle + \left\langle \psi\middle|\frac{d\mathcal{O}}{dt}\middle|\psi\right\rangle + \left\langle \psi\middle|\mathcal{O}\middle|\frac{d\psi}{dt}\right\rangle$$

$$= \left\langle \frac{\partial\psi}{\partial t}\middle|\mathcal{O}\middle|\psi\right\rangle + \left\langle \psi\middle|\frac{d\mathcal{O}}{dt}\middle|\psi\right\rangle + \left\langle \psi\middle|\mathcal{O}\middle|\frac{\partial\psi}{\partial t}\right\rangle . \tag{2-61}$$

But as long as our operators are functions of $x^i$ and $t$ or their derivatives (where $x \neq x(t)$), using similar logic to that of Box 2.1, we can take the total time derivative of the operator in the bottom row of (2-61) as a partial time derivative. This is what we do in (2-26).

### 2.10.2 For Relations Like (2-37)

We can generalize. Consider an entity, call it $\tilde{\mathcal{O}}$, that is a function of fields which are in turn functions of $x^i$ and $t$. We will temporarily assume $\tilde{\mathcal{O}}$ is a classical entity, and later extrapolate to quantum operators and quantum fields.

$$\tilde{\mathcal{O}} = f(x,t)\,g(x,t)\,h(x,t) \qquad x \neq x(t) \tag{2-62}$$

$\tilde{\mathcal{O}}$ in (2-62) is analogous to $\mathcal{O}^H = U^\dagger \mathcal{O}^S U$ in (2-37). $f$, $g$, and $h$ are analogous to $U^\dagger$, $\mathcal{O}^S$ and $U$.

$$\frac{d\tilde{\mathcal{O}}}{dt} = \frac{d\tilde{\mathcal{O}}}{df}\frac{df}{dt} + \frac{d\tilde{\mathcal{O}}}{dg}\frac{dg}{dt} + \frac{d\tilde{\mathcal{O}}}{dh}\frac{dh}{dt} = gh\frac{df}{dt} + fh\frac{dg}{dt} + fg\frac{dh}{dt}$$

$$= gh\left(\frac{\partial f}{\partial t}\frac{dt}{dt} + \frac{\partial f}{\partial x}\frac{dx}{dt}\right) + fh\left(\frac{\partial g}{\partial t}\frac{dt}{dt} + \frac{\partial g}{\partial x}\frac{dx}{dt}\right) + fg\left(\frac{\partial h}{\partial t}\frac{dt}{dt} + \frac{\partial h}{\partial x}\frac{dx}{dt}\right) = gh\frac{\partial f}{\partial t} + fh\frac{\partial g}{\partial t} + fg\frac{\partial h}{\partial t} . \tag{2-63}$$

Note the equivalence of the ends of the first and second rows in (2-63). So, as long as $\tilde{\mathcal{O}}$ is a function of fields (as it typically is in QFT), the partial time and total time derivatives on the RHS can be interchanged in these sorts of expressions. This holds true as long as $x$ is not a function of $t$.

If $f$, $g$, and $h$ are operators (as in QFT), we have to more careful about the order above (always keeping the factor with $f$ in it to the left of the factor with $g$, and $g$ to the left of $h$.)

## *2.11 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 $\alpha$ 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 $\phi$?

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 $X_i$ . The operation $\partial_\mu\partial^\mu = \partial^\mu\partial_\mu$ is called the d'Alembertian operator, and is the 4D Minkowski coordinates analogue of the 3D Laplacian operator $\partial_i\partial_i = \partial^i\partial^i$ of Cartesian coordinates.

    c) Then find $\partial^\mu\partial_\mu (x^\alpha x_\alpha)$, where physical length of the interval of $x^\alpha$ is $\sqrt{x^\alpha x_\alpha}$ , i) by expressing all terms in $t$ and $X_i$, and ii) solely using 4D component notation. (For the last part, note, from a), that $\partial x^\alpha / \partial x_\beta = g^{\alpha\beta}$ and from (2-5), $\partial x_\alpha / \partial x^\beta = g_{\alpha\beta}$ .)

5.  Obtain your answer to the following question by inspection of the final equation in Box 2-2, and then ask yourself whether or not your conclusion feels right intuitively.

    If $\phi^r$ were a sinusoid, how would the physical momentum density of a short wavelength wave compare to that of a longer one?

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 $\phi^r(x^i,t)$. $r = 1,2,3$, here, as there is a component of displacement, measured in length units, in each of the three spatial directions. $\phi^r$ and $x^i$ are both measures of length, but the $x^i$ are fixed locations in space, whereas the $\phi^r$ 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 $\phi^r$ (actually, it is in terms of the time derivatives of $\phi^r$ and $\phi_r$)? What is the Lagrangian density for the field? Use (2-13) to find the differential equation of motion for the displacement $\phi^r$. 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, where $\mu, \nu$ take on only values 1,2,3. In that case, it is better expressed as $g_{ij} = \delta_{ij}$ 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 $p_\mu$. 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 $\tau$ and fill in the boxes using 4D momentum, etc.) Keep it simple by treating only a free particle (no potential involved.)

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 $U = e^{-iH(t-t_0)}$ and $|\psi_E\rangle = |Ae^{-i(Et_0-\mathbf{p}\cdot\mathbf{x})}\rangle$, an energy eigenstate at time $t_0$. What is $U|\psi_E\rangle$? Does $U$ here act as a translator of the state in time? That is, does it have the effect of moving the state that was fixed in time forward in time, and turning it into a dynamic entity rather than a static one? If we operate on this new dynamic state with $U^\dagger$, would we turn it back into a static state? Is that not what we do when we operate on a Schrödinger picture state to turn it into a (static) Heisenberg picture state? (Earlier in the chapter we took $t_0 = 0$ to make things simpler.)

13.  (Problem added in revision of 2$^{nd}$ edition). Express the components $p^\mu$ of 4-momentum for a photon. Assume it is traveling in the $x^1$ direction. Use natural units where speed of light $c = 1$. (Hint: Use energy expressed in terms of frequency $f$ and 3-momentum in terms of wave length $\lambda$.) Then show that even though $p^\mu \neq 0$, $(p)^2 = p^\mu p_\mu = 0$. (Hint: Use speed of light expressed in terms of frequency and wave length.) Does this make sense in light of (2-55), given what we know about the photon mass? Then express $p^\mu$ in terms of $\omega = 2\pi f$ and wave number $k = 2\pi/\lambda$ where $\hbar = h/2\pi$. (= 1 in natural units).