Chapter 1

Introduction

1.1 Particles and Fields

Classically, there are two kinds of dynamical systems that we encounter. First, there is the motion of a particle or a rigid body (with a finite number of degrees of freedom) which can be described by a finite number of coordinates. And then, there are physical systems where the number of degrees of freedom is nondenumerably (non-countably) infinite. Such systems are described by fields. Familiar examples of classical fields are the electromagnetic fields described by $\vec{E}(\vec{x}, t)$ and $\vec{B}(\vec{x}, t)$ or equivalently by the potentials $(\phi(\vec{x}, t), \vec{A}(\vec{x}, t))$. Similarly, the motion of a one-dimensional string is also described by a field $\phi(\vec{x}, t)$, namely, the displacement field. Thus, while the coordinates of a particle depend only on time, fields depend continuously on some space variables as well. Therefore, a theory described by fields is usually known as a D+1 dimensional field theory where D represents the number of spatial dimensions on which the field variables depend. For example, a theory describing the displacements of the one-dimensional string would constitute a 1+1 dimensional field theory whereas the more familiar Maxwell's equations (in four dimensions) can be regarded as a 3+1 dimensional field theory. In this language, then, it is clear that a theory describing the motion of a particle can be regarded as a special case, namely, we can think of such a theory as a 0+1 dimensional field theory.

1.2 Metric and Other Notations

In these lectures, we will discuss both nonrelativistic as well as relativistic theories. For the relativistic case, we will use the Bjorken-Drell convention. Namely, the contravariant coordinates are assumed to be

$$x^{\mu} = (t, \vec{x}) \qquad \mu = 0, 1, 2, 3 \tag{1.1}$$

while the covariant coordinates have the form

$$x_{\mu} = \eta_{\mu\nu} x^{\nu} = (t, -\vec{x}) \tag{1.2}$$

Here we have assumed the speed of light to be unity (c=1). The covariant metric can, therefore, be obtained to be diagonal with the signatures

$$\eta_{\mu\nu} = (+,-,-,-) \tag{1.3}$$

The inverse or the contravariant metric clearly also has the same form, namely,

$$\eta^{\mu\nu} = (+,-,-,-) \tag{1.4}$$

The invariant length is given by

$$x^2 = x^\mu x_\mu = \eta^{\mu\nu} x_\mu x_\nu = \eta_{\mu\nu} x^\mu x^\nu = t^2 - \vec{x}^2 \tag{1.5}$$

The gradients are similarly obtained from Eqs. (1.1) and (1.2) to be

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = (\frac{\partial}{\partial t}, \vec{\nabla}) \tag{1.6}$$

$$\partial^\mu = \frac{\partial}{\partial x_\mu} = (\frac{\partial}{\partial t}, -\vec{\nabla}) \tag{1.7}$$

so that the D'Alembertian takes the form

$$\Box = \partial^\mu \partial_\mu = \eta^{\mu\nu} \partial_\mu \partial_\nu = \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 \tag{1.8}$$

1.3 Functionals

In any case, it is evident that in dealing with dynamical systems, we are dealing with functions of continuous variables. In fact, most of the times, we are really dealing with functions of functions which are otherwise known as functionals. If we are considering the motion of a particle in one dimension in a potential, then the Lagrangian is given by

$$L(x,\dot{x}) = \frac{1}{2}m\dot{x}^2 - V(x) \tag{1.9}$$

where $x(t)$ and $\dot{x}(t)$ denote the coordinate and the velocity of the particle and the simplest functional we can think of is the action functional defined as

$$S[x] = \int_{t_i}^{t_f} dt\, L(x,\dot{x}) \tag{1.10}$$

Note that unlike a function whose value depends on a particular point in the coordinate space, the value of the action depends on the entire trajectory along which the integration is carried out.

Thus, a functional has the generic form

$$F[f] = \int dx\, F(f(x)) \tag{1.11}$$

where, for example, we may have

$$F(f(x)) = (f(x))^n \tag{1.12}$$

Sometimes, one loosely also says that $F(f(x))$ is a functional. The notion of a derivative can be extended to the case of functionals in a natural way through the notion of generalized functions. Thus, one defines the functional derivative or the Gateaux derivative from the linear functional

$$F'[v] = \frac{d}{d\epsilon} F[f + \epsilon v]\bigg|_{\epsilon=0} = \int dx\, \frac{\delta F[f]}{\delta f(x)}\, v(x) \tag{1.13}$$

Equivalently, from the working point of view, this simply corresponds to defining

$$\frac{\delta F(f(x))}{\delta f(y)} = \lim_{\epsilon \to 0} \frac{F(f(x) + \epsilon\delta(x-y)) - F(f(x))}{\epsilon} \tag{1.14}$$

It now follows from Eq. (1.14) that

$$\frac{\delta f(x)}{\delta f(y)} = \delta(x-y) \tag{1.15}$$

The functional derivative satisfies all the properties of a derivative, namely, it is linear and associative,

$$\begin{aligned} \frac{\delta}{\delta f(x)}(F_1[f] + F_2[f]) &= \frac{\delta F_1[f]}{\delta f(x)} + \frac{\delta F_2[f]}{\delta f(x)} \\ \frac{\delta}{\delta f(x)}(F_1[f] F_2[f]) &= \frac{\delta F_1[f]}{\delta f(x)}\, F_2[f] + F_1[f]\, \frac{\delta F_2[f]}{\delta f(x)} \end{aligned} \tag{1.16}$$

It also satisfies the chain rule of differentiation. Furthermore, we now see that given a functional $F[f]$, we can Taylor expand it in the form

$$\begin{aligned} F[f] &= \int dx\, P_0(x) + \int dx_1 dx_2\, P_1(x_1, x_2)\, f(x_2) \\ &\quad + \int dx_1 dx_2 dx_3\, P_2(x_1, x_2, x_3)\, f(x_2) f(x_3) + \cdots \end{aligned} \tag{1.17}$$

where

$$
\begin{aligned}
P_0(x) &= F(f(x))\big|_{f(x)=0} \\
P_1(x_1, x_2) &= \left.\frac{\delta F(f(x_1))}{\delta f(x_2)}\right|_{f(x)=0} \\
P_2(x_1, x_2, x_3) &= \left.\frac{1}{2!}\frac{\delta^2 F(f(x_1))}{\delta f(x_2)\delta f(x_3)}\right|_{f(x)=0}
\end{aligned}
\tag{1.18}
$$

and so on.

As simple examples, let us calculate a few particular functional derivatives.

i) Let

$$F[f] = \int dy\ F(f(y)) = \int dy\ (f(y))^n \tag{1.19}$$

where n denotes a positive integer. Then,

$$
\begin{aligned}
\frac{\delta F(f(y))}{\delta f(x)} &= \lim_{\epsilon \to 0} \frac{F(f(y) + \epsilon\delta(y - x)) - F(f(y))}{\epsilon} \\
&= \lim_{\epsilon \to 0} \frac{(f(y) + \epsilon\delta(y - x))^n - (f(y))^n}{\epsilon} \\
&= \lim_{\epsilon \to 0} \frac{(f(y))^n + n\epsilon(f(y))^{n-1}\delta(y - x) + O(\epsilon^2) - (f(y))^n}{\epsilon} \\
&= n(f(y))^{n-1}\delta(y - x)
\end{aligned}
\tag{1.20}
$$

Therefore, we obtain

$$\begin{aligned}\frac{\delta F[f]}{\delta f(x)} &= \int dy\, \frac{\delta F(f(y))}{\delta f(x)} \\ &= \int dy\, n(f(y))^{n-1}\delta(y-x) \\ &= n(f(x))^{n-1} \end{aligned} \tag{1.21}$$

ii) Let us next consider the one-dimensional action in Eq. (1.10)

$$S[x] = \int_{t_i}^{t_f} dt'\, L(x(t'), \dot{x}(t')) \tag{1.22}$$

with

$$\begin{aligned} L(x(t), \dot{x}(t)) &= \frac{1}{2}m(\dot{x}(t))^2 - V(x(t)) \\ &= T(\dot{x}(t)) - V(x(t)) \end{aligned} \tag{1.23}$$

In a straightforward manner, we obtain

$$\begin{aligned}\frac{\delta V(x(t'))}{\delta x(t)} &= \lim_{\epsilon \to 0} \frac{V(x(t') + \epsilon\delta(t'-t)) - V(x(t'))}{\epsilon} \\ &= V'(x(t'))\delta(t'-t) \end{aligned} \tag{1.24}$$

where we have defined

$$V'(x(t')) = \frac{\partial V(x(t'))}{\partial x(t')}$$

and

$$\begin{aligned}\frac{\delta T(\dot{x}(t'))}{\delta x(t)} &= \lim_{\epsilon\to 0}\frac{T(\dot{x}(t')+\epsilon\frac{d}{dt'}\delta(t'-t))-T(\dot{x}(t'))}{\epsilon}\\ &= m\dot{x}(t')\frac{d}{dt'}\delta(t'-t)\end{aligned} \tag{1.25}$$

It is clear now that

$$\begin{aligned}\frac{\delta L(x(t'),\dot{x}(t'))}{\delta x(t)} &= \frac{\delta(T(\dot{x}(t'))-V(x(t')))}{\delta x(t)}\\ &= m\dot{x}(t')\frac{d}{dt'}\delta(t'-t)-V'(x(t'))\delta(t'-t)\end{aligned} \tag{1.26}$$

Consequently, in this case, we obtain for $t_i \le t \le t_f$

$$\begin{aligned}\frac{\delta S[x]}{\delta x(t)} &= \int_{t_i}^{t_f} dt'\ \frac{\delta L(x(t'),\dot{x}(t'))}{\delta x(t)}\\ &= \int_{t_i}^{t_f} dt'\ (m\dot{x}(t')\frac{d}{dt'}\delta(t'-t)-V'(x(t'))\delta(t'-t))\\ &= -m\ddot{x}(t)-V'(x(t))\\ &= -\frac{d}{dt}\frac{\partial L(x(t),\dot{x}(t))}{\partial\dot{x}(t)}+\frac{\partial L(x(t),\dot{x}(t))}{\partial x(t)}\end{aligned} \tag{1.27}$$

The right hand side is, of course, reminiscent of the Euler-Lagrange equation. In fact, we note that

$$\frac{\delta S[x]}{\delta x(t)} = -\frac{d}{dt}\frac{\partial L}{\partial\dot{x}(t)}+\frac{\partial L}{\partial x(t)}=0 \tag{1.28}$$

gives the Euler-Lagrange equation as a functional extremum of the action. This is nothing other than the principle of least action expressed in a compact notation in the language of functionals.

1.4 Review of Quantum Mechanics

In this section, we will describe very briefly the essential features of quantum mechanics assuming that the readers are familiar with the subject. The conventional approach to quantum mechanics starts with the Hamiltonian formulation of classical mechanics and promotes observables to noncommuting operators. The dynamics, in this case, is given by the time-dependent Schrödinger equation

$$i\hbar\frac{\partial|\psi(t)\rangle}{\partial t} = H|\psi(t)\rangle \tag{1.29}$$

where H denotes the Hamiltonian operator of the system. Equivalently, in the one dimensional case, the wave function of a particle satisfies

$$\begin{aligned} i\hbar\frac{\partial\psi(x,t)}{\partial t} &= H(x)\psi(x,t) \\ &= (-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x))\psi(x,t) \end{aligned} \tag{1.30}$$

where we have identified

$$\psi(x,t) = \langle x|\psi(t)\rangle \tag{1.31}$$

with $|x\rangle$ denoting the coordinate basis states. This, then, defines the time evolution of the system.

The main purpose behind solving the Schrödinger equation lies in determining the time evolution operator which generates the time translation of the system. Namely, the time evolution operator transforms the quantum mechanical state at an earlier time, t_2, to a future time, t_1, as

$$|\psi(t_1)\rangle = U(t_1,t_2)|\psi(t_2)\rangle \tag{1.32}$$

Clearly, for a time independent Hamiltonian, we see from Eq. (1.29) (the Schrödinger equation) that for $t_1 > t_2$,

$$U(t_1,t_2) = e^{-\frac{i}{\hbar}(t_1-t_2)H} \tag{1.33}$$

More explicitly, we can write

$$U(t_1,t_2) = \theta(t_1-t_2)e^{-\frac{i}{\hbar}(t_1-t_2)H} \tag{1.34}$$

It is obvious that the time evolution operator is nothing other than the Greens function for the time dependent Schrödinger equation and satisfies

$$(i\hbar\frac{\partial}{\partial t_1} - H)U(t_1,t_2) = i\hbar\delta(t_1-t_2) \tag{1.35}$$

Determining this operator is equivalent to finding its matrix elements in a given basis. Thus, for example, in the coordinate basis defined by

$$X|x\rangle = x|x\rangle \tag{1.36}$$

we can write

$$\langle x_1|U(t_1,t_2)|x_2\rangle = U(t_1,x_1;t_2,x_2) \tag{1.37}$$

If we know the function $U(t_1,x_1;t_2,x_2)$ completely, then the time evolution of the wave function can be written as

$$\psi(x_1,t_1) = \int dx_2\; U(t_1,x_1;t_2,x_2)\psi(x_2,t_2) \tag{1.38}$$

It is interesting to note that the dependence on the intermediate times drops out in the above equation as can be easily checked.

Our discussion has been within the framework of the Schrödinger picture so far where the quantum states $|\psi(t)\rangle$ carry time dependence while the operators are time independent. On the other hand, in the Heisenberg picture, where the quantum states are time independent, we can identify using Eq. (1.32)

$$\begin{aligned} |\psi\rangle_H &= |\psi(t=0)\rangle_S = |\psi(t=0)\rangle \\ &= e^{\frac{i}{\hbar}tH}|\psi(t)\rangle = e^{\frac{i}{\hbar}tH}|\psi(t)\rangle_S \end{aligned} \tag{1.39}$$

In this picture, the operators carry all the time dependence. For example, the coordinate operator in the Heisenberg picture is related to the coordinate operator in the Schrödinger picture through the relation

$$X_H(t) = e^{\frac{i}{\hbar}tH} X e^{-\frac{i}{\hbar}tH} \tag{1.40}$$

The eigenstates of this operator satisfying

$$X_H(t)|x,t\rangle_H = x|x,t\rangle_H \tag{1.41}$$

are then easily seen to be related to the coordinate basis in the Schrödinger picture through

$$|x,t\rangle_H = e^{\frac{i}{\hbar}tH}|x\rangle \tag{1.42}$$

It is clear now that for $t_1 > t_2$

$$\begin{aligned}
{}_H\langle x_1,t_1|x_2,t_2\rangle_H &= \langle x_1|e^{-\frac{i}{\hbar}t_1H}e^{\frac{i}{\hbar}t_2H}|x_2\rangle \\
&= \langle x_1|e^{-\frac{i}{\hbar}(t_1-t_2)H}|x_2\rangle \\
&= \langle x_1|U(t_1,t_2)|x_2\rangle \\
&= U(t_1,x_1;t_2,x_2)
\end{aligned} \tag{1.43}$$

Thus, we see that the matrix elements of the time evolution operator are nothing other than the time ordered transition amplitudes between the coordinate basis states in the Heisenberg picture.

Finally, there is the interaction picture where both the quantum states as well as the operators carry partial time dependence. Without going into any technical detail, let us simply note here that the interaction picture is quite useful in the study of non-trivially interacting theories. In any case, the goal of the study of quantum mechanics in any of these pictures is to construct the matrix elements of the time evolution operator which as we have seen can be identified with transition amplitudes between the coordinate basis states in the Heisenberg picture.

1.5 References

Dirac, P.A.M., "Principles of Quantum Mechanics", Oxford Univ. Press.

Schiff, L.I., "Quantum Mechanics", McGraw-Hill Publishing.