MECHANICS

James T. Wheeler

August 20, 2002
Modified Fall, 2003

Abstract

Original notes based on the text by L. D. Landau and E. M. Lifshitz
(3rd edition)
1 The equations of motion

1.1 Generalized coordinates

A particle is “a body whose dimensions may be neglected in describing its motion.” This means that we do not try to characterize its shape or what it is made of, but only those properties that depend only on location and change of location: position, time, momentum, velocity and so on.

The position, velocity and acceleration are given by a vectors:

\[ \mathbf{r} = (x, y, z) \text{ (in cartesian coordinates)} \]  
\[ \dot{\mathbf{r}} = \frac{d\mathbf{r}}{dt} = \mathbf{v} \]  
\[ \ddot{\mathbf{r}} = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2} = \mathbf{a} \]  

Multiple particles require a vector for the position of each particle, \( \mathbf{r}_i \).

The number of quantities required to fully specify a system is the number of degrees of freedom. For one particle in space there are three degrees of freedom; for \( N \) particles there are \( 3N \).

It is always possible to describe a system of particles in several ways. In general, to describe a system with \( 3N \) degrees of freedom, we need to specify \( 3N \) different quantities, but this may be done in several ways. For example, we might specify the position of a particle using polar coordinates instead of cartesian coordinates:

\[ \mathbf{r} = (\rho, \varphi, z) \]  

Any set of quantities which completely specify a system are called generalized coordinates for the system. There will always be as many generalized coordinates as there are degrees of freedom.

Knowing where a particle is does not tell us everything we can know about its mechanical properties. Because particles change position with time, we must also specify velocity. Most motion is adequately described by knowing the initial position and velocity; this corresponds to satisfying a second order differential equation. (This is not necessary, but it usually works. For example, by hooking up an accelerometer to a car, we could rig up the car so that the accelerator was pressed down proportionally to the rate of change of acceleration. Then the motion would satisfy

\[ \mathbf{a} = k \frac{d\mathbf{a}}{dt} \]  

2
Since this is a third order equation, a complete specification of the motion will require the initial position, velocity and acceleration. Such situations are unusual.)

### 1.2 The principle of least action

The principle of least action gives the most general formulation of classical motion. Given a function \( L(q_i, \dot{q}_i, t) \) called the Lagrangian, we write the action:

\[
S = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t) dt
\]

The principle of least action says that a classical system will follow a path that makes \( S \) extremal (either maximum or minimum). For most classical systems \( S \) is actually a minimum, but relativistic mechanics follows a maximum. For the problem of finding the path of the motion, it doesn’t matter which of these is the case.

Consider the case when we have one degree of freedom. We can find a differential equation that satisfies the principle of least action as follows. Suppose \( q(t) \) is the actual path of motion that makes \( S \) minimum. Denote the minimum value by \( S[q(t)] \). Then for any function \( \delta q(t) \), the path \( q(t) + \delta q(t) \) must produce a larger value of \( S \):

\[
S[q(t) + \delta q(t)] > S[q(t)]
\]

Just as a function is at a minimum if its derivative vanishes, we find the path, \( q(t) \), that makes \( S \) a minimum by demanding that a small change \( \delta q(t) \) in \( q(t) \) doesn’t change \( S \). That is, \( q(t) \) makes \( S \) minimum if

\[
S[q(t) + \delta q(t)] - S[q(t)] = 0
\]

for any small change \( \delta q(t) \). Notice that every path we consider should have the same endpoints:

\[
q(t_1) + \delta q(t_1) = q(t_1) \quad (9)
\]
\[
q(t_2) + \delta q(t_2) = q(t_2) \quad (10)
\]

or simply,

\[
\delta q(t_1) = \delta q(t_2) = 0 \quad (11)
\]
Now examine what this means in detail. We can expand:

\[
0 = S[q(t) + \delta q(t)] - S[q(t)] = \int_{t_1}^{t_2} L(q(t) + \delta q(t), \dot{q}(t) + \delta \dot{q}(t), t)dt - \int_{t_1}^{t_2} L(q(t), \dot{q}(t), t)dt \tag{12}
\]

Now, remembering that \( \delta q(t) \) is small, we can expand \( L \) about \( q(t) \) in a Taylor series. Remember that \( t \) is just a parameter here. Picking any particular value of \( t \), \( L(q, \dot{q}) \) is just an ordinary function of two variables:

\[
L(q(t) + \delta q(t), \dot{q}(t) + \delta \dot{q}(t), t) = L(q(t), \dot{q}(t), t) + \frac{\partial L}{\partial q(t)} \delta q(t) + \cdots \tag{14}
\]

\[
+ \frac{\partial L}{\partial \dot{q}(t)} \delta \dot{q}(t) + \cdots \tag{15}
\]

Since we can take \( \delta q \) to be as small as we like, we can neglect terms higher than first order. Then

\[
0 = \int_{t_1}^{t_2} L(q(t) + \delta q(t), \dot{q}(t) + \delta \dot{q}(t), t)dt - \int_{t_1}^{t_2} L(q(t), \dot{q}(t), t)dt \tag{16}
\]

\[
= \int_{t_1}^{t_2} \left( L(q(t), \dot{q}(t), t) + \frac{\partial L}{\partial q(t)} \delta q(t) + \frac{\partial L}{\partial \dot{q}(t)} \delta \dot{q}(t) \right) dt \tag{17}
\]

\[
- \int_{t_1}^{t_2} L(q(t), \dot{q}(t), t)dt \tag{18}
\]

\[
= \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q(t)} \delta q(t) + \frac{\partial L}{\partial \dot{q}(t)} \delta \dot{q}(t) \right) dt \tag{19}
\]

Next, we integrate the second term by parts. It is important to see here that \( \delta \dot{q}(t) \) is the difference in the velocity along the path \( q(t) + \delta q(t) \) from the velocity along the path \( q(t) \). Therefore,

\[
\delta \dot{q}(t) = \frac{d(q + \delta q)}{dt} - \frac{dq}{dt} \tag{20}
\]

\[
= \frac{dq}{dt} + \frac{d(\delta q)}{dt} - \frac{dq}{dt} = \frac{d(\delta q)}{dt} \tag{21}
\]

\[
= \frac{d(\delta q)}{dt} \tag{22}
\]
Now we can perform the integration by parts:

\[ I = \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial \dot{q}(t)} \right) \dot{q}(t) \, dt \]  

(23)

\[ = \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial \dot{q}(t)} \frac{d(\delta q)}{dt} \right) \, dt \]  

(24)

\[ = \int_{t_1}^{t_2} \left( \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}(t)} \dot{q}(t) \right) - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}(t)} \right) \delta q \right) \, dt \]  

(25)

\[ = \left( \left. \frac{\partial L}{\partial \dot{q}(t)} \dot{q}(t_2) \right|_{t_2} - \left. \frac{\partial L}{\partial \dot{q}(t)} \dot{q}(t_1) \right|_{t_1} \right) \]  

(26)

\[ - \int_{t_1}^{t_2} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}(t)} \right) \delta q \, dt \]  

(27)

\[ = - \int_{t_1}^{t_2} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}(t)} \right) \delta q \]  

(28)

where the integrated term vanishes because the variation vanishes at the endpoints. Returning to the full equation, we have

\[ 0 = S[q(t) + \delta q(t)] - S[q(t)] \]

\[ = \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial \dot{q}(t)} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}(t)} \right) \right) \delta q(t) \, dt \]  

(29)

We claim that the expression in brackets must vanish. The proof is as follows. Suppose there exists a time \( t_a \) for which

\[ \left( \frac{\partial L}{\partial \dot{q}(t)} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}(t)} \right) \right) \bigg|_{t_a} \neq 0 \]

Suppose, for concreteness that \( \left( \frac{\partial L}{\partial \dot{q}(t)} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}(t)} \right) \right) |_{t_a} > 0 \). Then if the Lagrangian and its derivatives are continuous, then we can always choose \( \varepsilon \) small enough that \( \frac{\partial L}{\partial \dot{q}(t)} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}(t)} \right) \) remains positive on the interval \( t \in [t_a - \varepsilon, t_a + \varepsilon] \). Now choose \( \delta q(t) \) to satisfy (for example)

\[ \delta q = \begin{cases} 1 + \sin \left( \frac{\pi}{\varepsilon} (t - t_a) + \frac{\pi}{2} \right) & t \in [t_a - \varepsilon, t_a + \varepsilon] \\ 0 & t \notin [t_a - \varepsilon, t_a + \varepsilon] \end{cases} \]

Then we must have

\[ 0 = \int_{t_{a-\varepsilon}}^{t_{a+\varepsilon}} \left( \frac{\partial L}{\partial \dot{q}(t)} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}(t)} \right) \right) \left( 1 + \sin \left( \frac{\pi}{\varepsilon} (t - t_a) + \frac{\pi}{2} \right) \right) \, dt \]  

(30)
But this is impossible because the integrand is positive definite. Therefore, our assumption was wrong: there is not time \( t_\alpha \) such that \( \frac{\partial L}{\partial q(t)} - \frac{d}{dt} \left( \frac{\partial L}{\partial q(t)} \right) \) is nonzero, i.e.,

\[
\frac{\partial L}{\partial q(t)} - \frac{d}{dt} \left( \frac{\partial L}{\partial q(t)} \right) = 0
\]  

(31)

at every point in the interval. This equation is called the Lagrange equation or the Euler-Lagrange equation.

In general, the Lagrange equation is a second order differential equation. If we have more than one degree of freedom, we can apply the same argument to each degree of freedom independently, so we just get one copy of the equation for each degree of freedom:

\[
\frac{\partial L}{\partial q_i(t)} - \frac{d}{dt} \left( \frac{\partial L}{\partial q_i(t)} \right) = 0
\]  

(32)

The Lagrangian for two non-interacting systems is simply given by adding the Lagrangians of the two systems.

Because the Euler-Lagrange equation is linear in the Lagrangian, multiplying the action by a constant leaves the Euler-Lagrange equations unchanged.

Adding a total time derivative does not change the equations of motion. The simplest way to see this is to integrate the added term. Let

\[
L' = L + \frac{df}{dt}
\]  

(33)

for any function \( f \). Then

\[
S' = \int_{t_1}^{t_2} L' dt
\]  

(34)

\[
= \int_{t_1}^{t_2} \left( L + \frac{df}{dt} \right) dt
\]  

(35)

\[
= f(t_2) - f(t_1) + S
\]  

(36)

Since our variation of the action does not change the endpoints,

\[
\delta (f(t_2) - f(t_1)) = 0
\]

and the equation of motion is unchanged.
The same result follows from the variation as well. Consider the variation

\[ 0 = \delta S' = \delta \int_{t_1}^{t_2} L' \, dt \]

\[ = \delta S + \int_{t_1}^{t_2} \delta \frac{df}{dt} \, dt \]

Substituting

\[ \frac{df}{dt} = \frac{\partial f}{\partial \dot{q}} \ddot{q} + \frac{\partial f}{\partial \ddot{q}} \dot{\ddot{q}} + \frac{\partial f}{\partial t} \]

into the extra term we find

\[ \int_{t_1}^{t_2} \delta \frac{df}{dt} \, dt = \int_{t_1}^{t_2} \delta \left( \frac{\partial f}{\partial \dot{q}} \ddot{q} + \frac{\partial f}{\partial \ddot{q}} \dot{\ddot{q}} + \frac{\partial f}{\partial t} \right) \, dt \]

\[ = \int_{t_1}^{t_2} dt \left( \frac{\partial^2 f}{\partial \dot{q}^2} \delta q + \frac{\partial^2 f}{\partial \ddot{q} \dot{q}} \delta \dot{q} \right) \ddot{q} + \frac{\partial f}{\partial \dot{q}} \delta q \]

\[ + \int_{t_1}^{t_2} dt \left( \frac{\partial^2 f}{\partial \ddot{q} \dot{q}} \delta \dot{q} + \frac{\partial^2 f}{\partial \ddot{q} \ddot{q}} \delta \ddot{q} \right) \dot{\ddot{q}} \]

\[ + \int_{t_1}^{t_2} dt \left( \frac{\partial^2 f}{\partial \ddot{q} \ddot{q}} \delta \ddot{q} + \frac{\partial^2 f}{\partial \ddot{q} \ddot{q}} \delta \ddot{q} \right) \]

\[ = \int_{t_1}^{t_2} dt \delta q \left( \frac{\partial^2 f}{\partial \dot{q}^2} \ddot{q} - \frac{d}{dt} \left( \frac{\partial^2 f}{\partial \ddot{q} \dot{q}} \right) - \frac{d}{dt} \frac{\partial f}{\partial \dot{q}} \right) \]

\[ + \int_{t_1}^{t_2} dt \delta q \left( \frac{\partial^2 f}{\partial \ddot{q} \dot{q}} \ddot{q} - \frac{d}{dt} \left( \frac{\partial^2 f}{\partial \ddot{q} \ddot{q}} \right) + \frac{d^2}{dt^2} \frac{\partial f}{\partial \ddot{q}} \right) \]

\[ + \int_{t_1}^{t_2} dt \delta q \left( \frac{\partial^2 f}{\partial \ddot{q} \ddot{q}} \ddot{q} - \frac{d}{dt} \frac{\partial^2 f}{\partial \ddot{q} \ddot{q}} \right) \]

The integrand of the extra term is therefore proportional to

\[ I = \frac{\partial^2 f}{\partial \dot{q}^2} \ddot{q} + \frac{\partial^2 f}{\partial \ddot{q} \dot{q}} \dot{\ddot{q}} + \frac{\partial^2 f}{\partial \ddot{q} \ddot{q}} \ddot{q} \]

\[ - \frac{d}{dt} \left( \frac{\partial f}{\partial \dot{q}} + \dot{q} \frac{\partial^2 f}{\partial \ddot{q} \dot{q}} + \ddot{q} \frac{\partial^2 f}{\partial \ddot{q} \ddot{q}} + \frac{\partial^2 f}{\partial \ddot{q} \ddot{q}} \right) \]

\[ + \frac{d^2}{dt^2} \frac{\partial f}{\partial \ddot{q}} \]
\[
= \frac{d}{dt} \frac{\partial f}{\partial q} - \frac{d}{dt} \left( \frac{\partial f}{\partial q} + \frac{d}{dt} \left( \frac{\partial f}{\partial \dot{q}} \right) \right)
\]
(49)
\[
+ \frac{d^2}{dt^2} \frac{\partial f}{\partial q} = 0
\]
(50)

so there is no contribution.

### 1.3 Functional differentiation

What distinguishes a functional such as the action \( S[x(t)] \) from a function \( f(x(t)) \), is that \( f(x(t)) \) is a number for each value of \( t \), whereas the value of \( S[x(t)] \) cannot be computed without knowing the entire function \( x(t) \). If we think of functions and functionals as maps, a compound function, \( f(x(t)) \), is the composition of two maps from the reals to the reals,

\[
f : R \to R
\]
(52)
\[
x : R \to R
\]
(53)
giving a third map

\[
f \circ x : R \to R
\]
(54)

A functional, by contrast, maps an entire function space into \( R \),

\[
S : F \to R
\]
(55)
\[
F = \{ f | f : R \to R \}
\]
(56)

In this section we develop the **functional derivative**, that is, the generalization of differentiation to functionals.

We would like the functional derivative to formalize finding the extremum of an action integral, so it makes sense to review the variation of an action. The usual argument is that we replace \( x(t) \) by \( x(t) + h(t) \) in the functional \( S[x(t)] \), then demand that to first order in \( h(t) \),

\[
\delta S \equiv S[x + h] - S[x] = 0
\]
(57)

We want to replace this statement by the demand that at the extremum, the first functional derivative of \( S[x] \) vanishes,

\[
\frac{\delta S[x(t)]}{\delta x(t)} = 0
\]
(58)
This gives us the answer, namely

\[
\frac{\delta S[x(t)]}{\delta x(t)} = \frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt}\frac{\partial L(x, \dot{x})}{\partial \dot{x}} \quad (59)
\]

All we have to do is figure out how to define the functional derivative so that we get this answer.

Suppose \( S \) is given by

\[ S[x(t)] = \int L(x(t), \dot{x}(t))\,dt \quad (60) \]

Then, recalling the steps of the previous section, we replace \( x \) by \( x + h \) and subtract \( S \). As before, after integration by parts, this gives

\[
\delta S \equiv \int L(x + h, \dot{x} + \dot{h})\,dt - \int L(x, \dot{x}(t))\,dt \quad (61)
\]

We are tempted to define a functional derivative by analogy with the derivative function, replacing

\[
\frac{df}{dx} = \lim_{\varepsilon \to 0} \frac{f(x + \varepsilon) - f(x)}{\varepsilon}
\]

with

\[
\frac{\delta S[x(t)]}{\delta x(t)} = \lim_{h(t) \to 0} \frac{S[x(t) + h(t)] - S[x(t)]}{h(t)}
\]

but this is not well-defined, since there may be times, \( t \), when \( h(t) = 0 \).

To overcome this problem, we consider a one parameter set of functions, \( x(t, \alpha) \), and perform an ordinary differentiation. Let

\[ x(t, \alpha) = x(t) + \alpha h(t) \]

where \( \alpha \) is independent of \( t \). Then define:

\[
\frac{dS[x(t, \alpha)]}{d\alpha}\bigg|_{\alpha=0} = \lim_{\beta \to 0} \frac{S[x(t) + (\alpha + \beta) h(t)] - S[x(t) + \alpha h(t)]}{\beta}\bigg|_{\alpha=0}
\]

\[
= \lim_{\beta \to 0} \frac{S[x(t) + \beta h(t)] - S[x(t)]}{\beta}
\]

\[
= \lim_{\beta \to 0} \frac{1}{\beta} \left( \int_{t_1}^{t_2} L(x(t) + \beta h(t), \dot{x}(t) + \beta \dot{h}(t), t)\,dt - \int_{t_1}^{t_2} L(x(t), \dot{x}(t), t)\,dt \right)
\]

\[
= \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}(t)} \right) h(t)\,dt
\]

9
where the last steps parallel the steps leading up to eq.(29).

Next, we need a systematic way to free the expression in parenthesis from the integral. Notice that we would have the right answer if we simply replaced $h(t)$ by a Dirac delta function, $\delta(t-t')$, defining

$$x(t, t', \alpha) = x(t) + \alpha \delta(t-t')$$

Then we could define

$$\frac{\delta S[x(t)]}{\delta x(t')} = \frac{dS[x(t, t', \alpha)]}{d\alpha} \bigg|_{\alpha=0}$$

$$= \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}(t)} \right) \delta(t-t') dt$$

$$= \frac{\partial L}{\partial q(t')} - \frac{d}{dt'} \frac{\partial L}{\partial \dot{q}(t')}$$

This approach suffers a fatal flaw: we are not allowed to evaluate functionals on distributions. For example, for any bounded function $f(t)$,

$$S[f(t)] = \int_{t_1}^{t_2} f^2(t) dt$$

is a well-defined functional, but if we try to replace $f(t)$ by the distribution $\delta(t)$, we find

$$S[\delta(t)] = \int_{t_1}^{t_2} \delta^2(t) dt = \delta(0) = \infty$$

Fortunately, there is a simple remedy.

Let $h_n(t, t')$ be any collection of functions such that

$$\lim_{n \to \infty} h_n(t, t') = \delta(t-t')$$

Then we define

$$x(t, t', \alpha, n) = x(t) + \alpha h_n(t, t')$$

and

$$\frac{\delta S}{\delta x}(t') = \lim_{n \to \infty} \frac{dS[x(t, t', \alpha, n)]}{d\alpha} \bigg|_{\alpha=0}$$

$$= \lim_{n \to \infty} \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial x(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}(t)} \right) h_n(t, t') dt$$

$$= \frac{\partial L}{\partial x(t')} - \frac{d}{dt'} \frac{\partial L}{\partial \dot{x}(t')}$$

The integrals are now well-defined.
1.3.1 Higher order derivatives

To define higher order derivatives, we can simply carry the expansion of $L$ to higher order.

\[
S = \int L(x, \dot{x}) \, dt
\]

\[
\delta S = \int L(x + \delta x, \dot{x} + \delta \dot{x}) \, dt - \int L(x, \dot{x}) \, dt
\]

\[
= \int L(x, \dot{x}) + \frac{\partial L}{\partial x} \delta x + \frac{\partial L}{\partial \dot{x}} \delta \dot{x} + \frac{1}{2!} \frac{\partial^2 L}{\partial x \partial \dot{x}} (\delta x)^2
\]

\[
+ \frac{\partial^2 L}{\partial x \partial \dot{x}} x \delta x \delta \dot{x} + \frac{1}{2!} \frac{\partial^2 L}{\partial \dot{x} \partial x} (\delta \dot{x})^2 - L(x, \dot{x}) \, dt
\]

\[
= \int \frac{\partial L}{\partial x} \delta x + \frac{\partial L}{\partial \dot{x}} \delta \dot{x} + \frac{1}{2!} \frac{\partial^2 L}{\partial x(t) \partial x(t)} (\delta x)^2
\]

\[
+ \frac{\partial^2 L}{\partial x(t) \partial \dot{x}(t)} (\delta x \delta \dot{x} + \frac{1}{2!} \frac{\partial^2 L}{\partial \dot{x}(t) \partial x(t)} (\delta \dot{x})^2
\]

Now, to integrate each of the various terms by parts, we need to insert some delta functions. Look at one term at a time:

\[
\int \frac{\partial L}{\partial \dot{x}} \delta \dot{x} = - \int \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \delta x
\]

\[
I_2 = \frac{1}{2!} \int dt \frac{\partial^2 L}{\partial x \partial \dot{x}} (\delta x)^2
\]

\[
= \frac{1}{2} \int dt \int dt' \delta(t - t') \frac{\partial^2 L}{\partial x(t) \partial x(t')} \delta x(t) \delta x(t')
\]

\[
I_3 = \int dt \frac{\partial^2 L}{\partial x \partial \dot{x}} \delta x \delta \dot{x}
\]

\[
= \int dt \int dt' \delta(t - t') \frac{\partial^2 L}{\partial x \partial \dot{x}} \left( \frac{1}{2} \delta x(t) \delta \dot{x}(t') + \delta x(t') \delta \dot{x}(t) \right)
\]

\[
= - \frac{1}{2} \int dt \int dt' \left( \frac{d}{dt'} \left( \delta(t - t') \frac{\partial^2 L}{\partial x \partial \dot{x}} \right) \right)
\]

\[
+ \frac{d}{dt} \left( \delta(t - t') \frac{\partial^2 L}{\partial x \partial \dot{x}} \right) \delta x(t) \delta x(t')
\]

\[
= - \frac{1}{2} \int dt \int dt' \left( \frac{\partial}{\partial t'} \delta(t - t') \frac{\partial^2 L}{\partial x \partial \dot{x}} \right)
\]
\[
+ \frac{\partial}{\partial t} \delta(t - t') \frac{\partial^2 L}{\partial x \partial \dot{x}} + \delta(t - t') \frac{d}{dt} \left( \frac{\partial^2 L}{\partial x \partial \dot{x}} \right) \delta x(t) \delta x(t')
\]
\[
= -\frac{1}{2} \int \! dt \int \! dt' \left( \delta(t - t') \frac{d}{dt} \frac{\partial^2 L}{\partial x \partial \dot{x}} \right) \delta x(t) \delta x(t')
\]

and finally,

\[
I_4 = \frac{1}{2!} \int \! dt \! \int \frac{\partial^2 L}{\partial \dot{x}(t) \partial \dot{x}(t)} \left( \delta \dot{x}(t) \right)^2
\]
\[
= \frac{1}{2} \int \! dt \! \int \! dt' \delta(t - t') \frac{\partial^2 L}{\partial \dot{x}(t) \partial \dot{x}(t)} \delta \dot{x}(t) \delta \dot{x}(t')
\]
\[
= \frac{1}{2} \int \! dt \! \int \! dt' \frac{\partial}{\partial t} \frac{\partial}{\partial t'} \left( \delta(t - t') \frac{\partial^2 L}{\partial \dot{x}(t) \partial \dot{x}(t)} \right) \delta x(t) \delta x(t')
\]
\[
= \frac{1}{2} \int \! dt \! \int \! dt' \frac{\partial}{\partial t} \frac{\partial}{\partial t'} \left( \delta(t - t') \frac{\partial^2 L}{\partial \dot{x} \partial x} \right) \delta x(t) \delta x(t')
\]
\[
+ \frac{1}{2} \int \! dt \! \int \! dt' \left( \frac{\partial}{\partial t} \delta(t - t') \frac{d}{dt} \frac{\partial^2 L}{\partial \dot{x} \partial x} \right) \delta x(t) \delta x(t')
\]
\[
= \frac{1}{2} \int \! dt \! \int \! dt' \left( \frac{\partial^2}{\partial t^2} \delta(t - t') \frac{\partial^2 L}{\partial \dot{x} \partial x} \right) \delta x(t) \delta x(t')
\]
\[
+ \frac{1}{2} \int \! dt \! \int \! dt' \left( \frac{\partial}{\partial t} \delta(t - t') \frac{d}{dt} \frac{\partial^2 L}{\partial \dot{x} \partial x} \right) \delta x(t) \delta x(t')
\]

Combining,

\[
\delta S = \int \! dt \left( \frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right) \delta x
\]
\[
+ \frac{1}{2} \int \! dt \! \int \! dt' \delta(t - t') \frac{\partial^2 L}{\partial x \partial \dot{x}} \delta x(t) \delta x(t')
\]
\[
- \frac{1}{2} \int \! dt \! \int \! dt' \left( \delta(t - t') \frac{d}{dt} \frac{\partial^2 L}{\partial \dot{x} \partial x} \right) \delta x(t) \delta x(t')
\]
\[
+ \frac{1}{2} \int \! dt \! \int \! dt' \left( \frac{\partial^2}{\partial t^2} \delta(t - t') \frac{\partial^2 L}{\partial \dot{x} \partial x} \right) \delta x(t) \delta x(t')
\]
\[
+ \frac{1}{2} \int \! dt \! \int \! dt' \left( \frac{\partial}{\partial t} \delta(t - t') \frac{d}{dt} \frac{\partial^2 L}{\partial \dot{x} \partial x} \right) \delta x(t) \delta x(t')
\]
\[
\frac{\delta S}{\delta x} = \frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \\
\frac{\delta^2 S}{\delta x(t) \delta x(t')} = \left( \frac{\partial^2 L}{\partial x \partial x} - \frac{d}{dt} \frac{\partial^2 L}{\partial x \partial \dot{x}} \right) \delta (t - t') \\
- \frac{d}{dt} \left( \frac{\partial^2 L}{\partial \dot{x} \partial \dot{x}} \right) \frac{\partial \delta (t - t')}{\partial t} - \frac{\partial^2 L}{\partial \dot{x} \partial \dot{x}} \frac{\partial^2 \delta (t - t')}{\partial t^2}
\]

This agrees with DeWitt when there is only one function \( x \). The result may be used to define a second functional derivative. Set

\[
\delta S = \int dt \frac{\delta S}{\delta x} \delta x + \frac{1}{2} \int dt dt' \frac{\delta^2 S}{\delta x(t) \delta x(t')} \delta x(t) \delta x(t') + \cdots 
\]

Then we identify:

\[
\frac{\delta S}{\delta x} = \frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \\
\frac{\delta^2 S}{\delta x(t) \delta x(t')} = \left( \frac{\partial^2 L}{\partial x \partial x} - \frac{d}{dt} \frac{\partial^2 L}{\partial x \partial \dot{x}} \right) \delta (t - t') \\
- \frac{d}{dt} \left( \frac{\partial^2 L}{\partial \dot{x} \partial \dot{x}} \right) \frac{\partial \delta (t - t')}{\partial t} - \frac{\partial^2 L}{\partial \dot{x} \partial \dot{x}} \frac{\partial^2 \delta (t - t')}{\partial t^2}
\]

Third and higher order derivatives may be defined by extending this procedure.

The result may also be found by taking two independent variations from the start.

### 1.4 Galileo’s relativity principle

Landau and Lifshitz make the following argument. It is a reasonable assumption (but an assumption nonetheless) that space and time are homogeneous and space is isotropic, in the absence of sources of force (or at a great distance from sources - this is what we mean by “an isolated system”). As a consequence, we should be able to find a Lagrangian with the same properties (though not all Lagrangians that describe motion in a homogeneous, isotropic space have the property, since we can always add the time derivative of an arbitrary function). To be homogeneous in space and time, \( L \) must be independent of position and time, so \( L(\mathbf{r}, \mathbf{v}, t) \) reduces to \( L(\mathbf{v}) \). Isotropy means that \( L \) cannot depend on the direction of \( \mathbf{v} \), so
we must have a function of the magnitude, \( L(v^2) \). The Lagrange equations then reduce to
\[
\frac{d}{dt} \left( \frac{\partial L}{\partial v} \right) = 0
\]  
so that integrating, we have a constant vector, \( \mathbf{k} \), such that
\[
\mathbf{k} = \frac{\partial L}{\partial v} = 2 \frac{dL}{d(v^2)} v
\]
As long as the derivative is nonzero, this means that the velocity, \( \mathbf{v} = \mathbf{v}_0 \), is constant in both direction and magnitude. The motion is therefore given by the straight line,
\[
\mathbf{x} = \mathbf{x}_0 + \mathbf{v}_0 t
\]
where \( \mathbf{x}_0 \) and \( \mathbf{v}_0 \) provide the expected number of possible initial conditions.

We can prove a converse to this result. For simplicity, let’s work with one degree of freedom.

We start by assuming that there exists a class of special frames of reference such that the motion of a free particle - a particle not influenced by any forces - is a straight line, and that the velocity remains constant.

In such a special frame of reference, the motion is given by \( q(t) = q_0 + v_0 t \) (so that \( v(t) = v_0 \)). We require this to be a solution to the equation of motion:
\[
0 = \frac{\partial L}{\partial q(t)} - \frac{d}{dt} \left( \frac{\partial L}{\partial v(t)} \right)
\]
Expand \( L(q, v) \) in a double Taylor series:
\[
L = \sum_{m,n=0}^{\infty} \frac{1}{m!n!} \left. \frac{\partial^2 L}{\partial q^m \partial v^n} \right|_{q=v=0} q^m v^n
\]
\[
= \sum_{m,n=0}^{\infty} \frac{1}{m!n!} a_{mn} q^m v^n
\]
where we have defined constants \( a_{mn} = \left. \frac{\partial^2 L}{\partial q^m \partial v^n} \right|_{q=v=0} \). Now compute:
\[
0 = \frac{\partial L}{\partial q(t)} - \frac{d}{dt} \left( \frac{\partial L}{\partial v(t)} \right)
\]
\[
\begin{align*}
\sum_{m,n=0}^{\infty} \frac{1}{m!n!} m a_{mn} q^{m-1} v^n - \frac{d}{dt} \sum_{m,n=0}^{\infty} \frac{1}{m!n!} n a_{mn} q^m v^{n-1} & = 0 \quad (76) \\
\sum_{m,n=0}^{\infty} \frac{1}{m!n!} m a_{mn} q^{m-1} v^n - \sum_{m,n=0}^{\infty} \frac{1}{m!n!} n m a_{mn} q^m v^{n-1} & = 0 \quad (77) \\
- \sum_{m,n=0}^{\infty} \frac{1}{m!n!} n (n-1) a_{mn} q^n v^{n-2} \frac{dv}{dt} & = 0 \quad (78) \\
\sum_{m,n=0}^{\infty} \frac{1}{m!n!} m (1-n) a_{mn} q^{m-1} v^n & = 0 \quad (79) \\
- \sum_{m,n=0}^{\infty} \frac{1}{m!n!} n (n-1) a_{mn} q^m v^{n-2} \frac{dv}{dt} & = 0 \quad (80)
\end{align*}
\]

By assumption, we require the solution of this equation to be a straight line,

\[q = q_0 + v_0 t\]

for any initial conditions. Therefore we must have \(\frac{dv}{dt} = 0\). Substituting into the equation of motion, we therefore demand

\[
0 = \sum_{m,n=0}^{\infty} \frac{1}{m!n!} m (1-n) a_{mn} q^{m-1} v^n
\]

\[
= \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \frac{1}{m!n!} m (1-n) a_{mn} q^{m-1} v^n
\]

\[
= \sum_{n=0}^{\infty} \sum_{s=0}^{\infty} \frac{1}{s!n!} (1-n) a_{s+1,n} q^n v^n
\]

Substituting the arbitrary initial conditions \(q_0, v_0\) for \(q\) and \(v\) at time \(t = 0\), we have

\[
0 = \sum_{n=0}^{\infty} \sum_{s=0}^{\infty} (s+1) (1-n) a_{s+1,n} q_0^n v_0^n
\]

Since \(q_0\) and \(v_0\) are arbitrary and independent, each term must vanish separately.

We therefore have, for all \(n, s = 0, 1, \ldots, \infty\):

\[0 = (s+1) (1-n) a_{s+1,n}\]
or since \( s + 1 \) is nonzero,

\[
0 = (1 - n) a_{s+1,n}
\]

and all coefficients \( a_{mn} \) in the Taylor series for \( L \) vanish except \( a_{0n} \) for all \( n \), and \( a_{s1} \) for all \( s \). The Lagrangian must therefore be of the form

\[
L(q, v) = \sum_{n=0}^{\infty} a_{0n} v^n + \left( \sum_{s=0}^{\infty} a_{s1} q^s \right) v
\]

Since \( a_{0n} \) and \( a_{s1} \) are arbitrary, the remaining sums are arbitrary analytic functions \( f(v) \) and \( g(q) \), respectively, so

\[
L(q, v) = f(v) + g(q)v
\]

Notice that the second term may be written as

\[
g(q)v = \frac{dG(q)}{dt}
\]

where

\[
G(q) = \int_q^q g(q)dq
\]

and is therefore a total time derivative. Since a total derivative does not contribute to the equations of motion, we may equivalently write simply

\[
L' = L(v)
\]

The only further constraint on \( L(v) \) is that it must be at least quadratic in \( v \) – otherwise the equation of motion vanishes identically. The simplest Lagrangian that results in straight-line motion is therefore, \( L' = \alpha v^2 \) for any constant \( \alpha \).

### 1.4.1 Galilean transformations

Now suppose we have an inertial frame of reference, so that a free particle moves with constant velocity:

\[
x = x_0 + v_0 t
\]
Are there other inertial frames of reference? In any other inertial frame, we must have
\[ \mathbf{X} = \mathbf{X}_0 + \mathbf{V}_0 t \] (85)
Comparing these, we have
\[ \mathbf{X} - \mathbf{x} = \mathbf{X}_0 - \mathbf{x}_0 + (\mathbf{V}_0 - \mathbf{v}_0) t \] (86)
If we define
\[ \mathbf{u} = \mathbf{V}_0 - \mathbf{v}_0 \] (87)
\[ \mathbf{r}_0 = \mathbf{X}_0 - \mathbf{x}_0 \] (88)
then the coordinates in the two inertial frames are related by
\[ \mathbf{X} = \mathbf{x} + \mathbf{r}_0 + \mathbf{u} t \] (89)
Therefore, there is a 6-parameter set of inertial frames of reference, which differ from one another by \( \mathbf{r}_0 \) in their choice of origin, and move with respect to one another with a constant velocity \( \mathbf{u} \).

1.5 The Lagrangian for a free particle
We have established that the action is extremal for straight line motion if and only if the Lagrangian is a function of velocity only, up to a total time derivative. This establishes a set of inertial observers within the context of variational mechanics. Now we narrow the range of free particle Lagrangians further by demanding Galilean invariance of the action. This will insure that the Lagrangian equations of motion are independent of our choice of inertial frame of reference.

Suppose we are in the \( \mathbf{x} \) frame, and the Lagrangian is given by \( L(v) \), where \( v = |\dot{\mathbf{x}}| \). Then, since the Lagrangian in the \( \mathbf{X} \) frame must be the same up to a total time derivative, we require
\[ L(V^2) - L(v^2) = \frac{df}{dt} \] (90)
where the two velocities are related by \( \mathbf{v} = \mathbf{V} - \mathbf{u} \). It is sufficient to choose \( u \) infinitesimal and expand \( L \) to first order in a Taylor series about \( v \):
\[ L(V^2) = L(v^2 + 2\mathbf{v} \cdot \mathbf{u} + u^2) = L(v^2) + \frac{dL}{d(v^2)} 2\mathbf{v} \cdot \mathbf{u} + O(u^2) \] (91)
Then the extra term has to be a total time derivative:

\[
\frac{df}{dt} = \frac{dL}{d(v^2)} 2\mathbf{v} \cdot \mathbf{u} = \frac{dL}{d(v^2)} \frac{d\mathbf{x}}{dt} \cdot \mathbf{u} = \frac{d}{dt} \left(\frac{dL}{d(v^2)} 2\mathbf{x} \cdot \mathbf{u}\right) - \frac{d}{dt} \left(\frac{dL}{dv^2}\right) 2\mathbf{x} \cdot \mathbf{u}
\]

(93)

(94)

(95)

The first term on the right is a total time derivative, but the second has got to go:

\[-\frac{d}{dt} \left(\frac{dL}{dv^2}\right) 2\mathbf{x} \cdot \mathbf{u} = 0
\]

(96)

Since the direction of \(\mathbf{u}\) is arbitrary,

\[\frac{d}{dt} \left(\frac{dL}{dv^2}\right) = 0
\]

(97)

so

\[\frac{dL}{dv^2} = \frac{m}{2} = \text{const.}
\]

(98)

where we have chosen the constant with a bit of forethought. Integrating with respect to \(v^2\) we get

\[L = \frac{1}{2}mv^2 + \text{const.}
\]

(99)

as the necessary form for the Lagrangian of a single free particle to be invariant under infinitesimal Galilean transformations. This form is also sufficient for full Galilean invariance, because we can now consider an arbitrary value of \(\mathbf{u}\) in \(L(V^2)\):

\[L(V^2) = L(v^2 + 2\mathbf{v} \cdot \mathbf{u} + u^2) = \frac{1}{2}m(v^2 + 2\mathbf{v} \cdot \mathbf{u} + u^2) = \frac{1}{2}mv^2 + \frac{d}{dt}(m\mathbf{r} \cdot \mathbf{u} + \frac{1}{2}mu^2t)
\]

(100)

(101)

(102)

where we use the constancy of \(\mathbf{u}\) in the last step.
For a system of $N$ free particles, we just add the free Lagrangians for each particle:

$$L = \sum_{a=1}^{N} \frac{1}{2} m_a v_a^2$$  

(103)

Notice that since the Lagrangian is arbitrary up to an overall multiple, we are free to choose the unit of mass. Thus, $L$ only depends on the ratios of the masses. For example, if we choose $m_1$ as our mass standard, we can write

$$L' = \frac{1}{m_1} L = \sum_{a=1}^{N} \frac{1}{2} \left( \frac{m_a}{m_1} \right) v_a^2$$  

(104)

It is easy to see that $L'$ gives the same equations of motion as $L$.

Here’s a cool trick. Notice that the velocity is

$$v = \frac{d\mathbf{r}}{dt}$$  

(105)

so that

$$v^2 = \mathbf{v} \cdot \mathbf{v} = \frac{d\mathbf{r}}{dt} \cdot \frac{d\mathbf{r}}{dt}$$  

(106)

$$= \frac{d\mathbf{r} \cdot d\mathbf{r}}{dt^2}$$  

(107)

Now, $d\mathbf{r} \cdot d\mathbf{r}$ is just the infinitesimal line element:

$$ds^2 = d\mathbf{r} \cdot d\mathbf{r}$$  

(108)

If you’ve ever seen any general relativity, you know that this is just one way to specify the metric $ds^2 = \sum g_{ij} dx^i dx^j$. In any case, we know what this is for several special cases:

$$ds^2 = dx^2 + dy^2 + dz^2$$  

(109)

$$= d\rho^2 + \rho^2 d\varphi^2 + dz^2$$  

(110)

$$= dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\varphi^2$$  

(111)

and it’s not too hard to figure out what it is in any coordinate system if someone hands us the coordinate transformations.

19
This means that the free particle Lagrangian is easy to write in polar or spherical coordinates:

$$L_{\text{polar}} = \frac{1}{2} m \left( \left( \frac{dp}{dt} \right)^2 + \rho^2 \left( \frac{d\varphi}{dt} \right)^2 + \left( \frac{dz}{dt} \right)^2 \right)$$

$$L_{\text{spherical}} = \frac{1}{2} m \left( \left( \frac{d\rho}{dt} \right)^2 + r^2 \left( \frac{d\theta}{dt} \right)^2 + r^2 \sin^2 \theta \left( \frac{d\varphi}{dt} \right)^2 \right)$$

These forms will prove useful later.

### 1.6 The Lagrangian for a system of particles

Next we need to allow particles to interact with one another or respond to an applied potential. We start with the case of a closed system. Many interactions (but not all) may be described by simply adding a position-dependent function to the Lagrangian. Denote the function by $-U(r_1, r_2, \ldots)$ where $r_a$ is the position of the $a$th particle. Then

$$L = \sum_{a=1}^{N} \frac{1}{2} m_a v_a^2 - U(r_1, r_2, \ldots, r_N)$$

We call $T = \sum_{a=1}^{N} \frac{1}{2} m_a v_a^2$ the kinetic energy of the system and $U(r_1, r_2, \ldots)$ the potential energy. The Lagrangian is then the difference between $T$ and $U$.

We work out the equations of motion:

$$0 = \frac{\partial L}{\partial r_a(t)} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{r}_a(t)} \right)$$

$$= -\frac{\partial U}{\partial r_a(t)} - m \frac{d\dot{r}_a}{dt}$$

or

$$m\ddot{r}_a = -\frac{\partial U}{\partial r_a}$$

which we recognize immediately as Newton’s second law. Landau and Lifshitz point out the causality violation inherent in these equations. Because $U$ depends only on the positions of the particles, if one particle is displaced
the force on all other particles changes *immediately*. This is not in agreement with our experience - it takes time for the effect to travel from one particle to another. The situation is corrected in field theory. Maxwell’s formulation of electromagnetism (and also the Dirac/Feynman formulation) predict that electromagnetic interactions travel at the speed of light. Similarly, general relativity and the modern field theory formulations of the weak and strong interactions correctly predict a delay in these interactions. However, for our purposes, the speed of interaction is so fast that we may neglect any time delay, which for macroscopic, laboratory-scale distances is measured in billionths of a second.

Notice that replacing $t$ by $-t$ leaves Newton’s law unchanged. This is called *time reversal invariance*. It is also violated in Nature, but only by the weak interaction.

We can find the new form of $L$ in an arbitrary set of generalized coordinates. If the positions are given by

$$x_a = f_a(q_1, q_2, \ldots, q_N)$$

then the velocities are

$$\dot{x}_a = \sum_{b=1}^{3N} \frac{\partial f_a}{\partial q_b} \dot{q}_b$$

Here the index $a$ ranges over the total number of degrees of freedom of the system, including both the number of particles, $N$, and the three independent coordinates, 1, 2, 3. We now substitute into $L$:

$$L = \sum_{a=1}^{3N} \frac{1}{2} m_a v_a^2 - U(r_1, r_2, \ldots, r_N)$$

$$= \sum_{a=1}^{3N} \frac{1}{2} m_a \left( \sum_{b=1}^N \frac{\partial f_a}{\partial q_b} \right)^2 - U(q_1, q_2, \ldots, q_N)$$

$$= \sum_{a=1}^{3N} \frac{1}{2} m_a \left( \sum_{b=1}^N \frac{\partial f_a}{\partial q_b} \dot{q}_b \right) \left( \sum_{c=1}^N \frac{\partial f_a}{\partial q_c} \dot{q}_c \right) - U(q)$$

$$= \sum_{b=1}^{3N} \sum_{c=1}^{3N} \left( \sum_{a=1}^{3N} \frac{1}{2} m_a \frac{\partial f_a}{\partial q_b} \left( \frac{\partial f_a}{\partial q_c} \right) \right) \dot{q}_b \dot{q}_c - U(q)$$

$$= \frac{1}{2} \sum_{b,c=1}^{3N} a_{bc} \dot{q}_b \dot{q}_c - U(q)$$
where we define

\[ a_{bc}(q_1, q_2, \ldots, q_N) = \sum_{a=1}^{3N} m_a \left( \frac{\partial f_a}{\partial q_b} \right) \left( \frac{\partial f_a}{\partial q_c} \right) \]  

(125)

Notice that \( a_{bc} \) depends on the coordinates but not the velocities, so the new Lagrangian is \textit{always a quadratic function of the velocities},

\[ L = \frac{1}{2} \sum_{b,c=1}^{3N} a_{bc} \dot{q}_b \dot{q}_c - U(q) \]  

(126)

\textit{as long as the potential is independent of velocity.}

For an open system, we can imagine adding the Lagrangian of the external sources to make a closed system again. Then, if we let \( L_A \) be the open system Lagrangian without the external forces, and let \( L_B \) be the Lagrangian of the external sources, the total \( L_{total} = L_A + L_B \) describes a closed system. Now, imagine that we vary \( L_{total} \) with respect to the \textit{external} particles and solve for their motion,

\[ q_B = q_B(t) \]  

(127)

\[ \dot{q}_B = \frac{dq_B(t)}{dt} \]  

(128)

Then, at least in principle, we can substitute these explicit functions of time back into \( L_{total} \) to give

\[ L_{total} = T_A(q_A, \dot{q}_A) + T_B(q_B(t), \dot{q}_B(t)) - U(q_A, q_B(t)) \]  

(129)

But \( T_B(q_B(t), \dot{q}_B(t)) \) is just a function of time, so we can write it as a time derivative:

\[ T_B(q_B(t), \dot{q}_B(t)) = \frac{d}{dt} \left( \int_{t_0}^{t} T_B(q_B(t'), \dot{q}_B(t')) dt' \right) \]  

(130)

and therefore it doesn’t contribute to the equations of motion for the remaining particles. We then have the effective Lagrangian

\[ L'_{total} = T_A(q_A, \dot{q}_A) - U(q_A, q_B(t)) \]  

(131)

where the only effect of the external forces is in the form of the potential, which may depend on time through the external positions. Therefore, in principle, we can always describe an open system by allowing the potential to depend on time in some appropriate way.
2 Conservation Laws

A system with $2s$ degrees of freedom may be solved for the $s$ positions and $s$ velocities as functions of time. However, it is often more useful to find quantities which do not change in time, that is, conserved quantities. In general there are $2s$ quantities which are independent of time, and therefore depend only on the initial conditions - indeed, we could take the initial conditions themselves, though certain other quantities are more useful.

Generally, any quantities built from the positions and velocities which stay constant in time are called integrals of the motion. There will be $2s - 1$ of these, which together with the initial time, make up the full $2s$ degrees of freedom of the system. Certain of these are more interesting than others. The ones that are interesting are (usually) those that have one or both of two important properties: (1) they come from some symmetry of the problem, and/or (2) they are additive. Such quantities are said to be conserved.

2.1 Energy

The first conservation law is related to time translation invariance, or as Landau and Lifshitz put it, the homogeneity of time. This holds whenever the Lagrangian does not depend explicitly on time. This, in turn, is true for any closed system and for open systems where the external forces are constant. In either case we have $\frac{dL}{dt} = 0$, so we can write the total time derivative of $L$ using the chain rule:

$$\frac{dL}{dt} = \sum_{i=1}^{3N} \frac{\partial L}{\partial q_i} \dot{q}_i + \sum_{i=1}^{3N} \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i$$  \hspace{1cm} (132)

Using the Lagrange equations to replace

$$\frac{\partial L}{\partial q_i} = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i}$$  \hspace{1cm} (133)

in the first term, we get

$$\frac{dL}{dt} = \sum_{i=1}^{3N} \left( \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i + \sum_{i=1}^{3N} \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i$$  \hspace{1cm} (134)
\[= \sum_{i=1}^{3N} \left( \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i + \sum_{i=1}^{3N} \frac{\partial L}{\partial q_i} \dot{q}_i \quad (135)\]

\[= \sum_{i=1}^{3N} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i \quad (136)\]

\[= \frac{d}{dt} \sum_{i=1}^{3N} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i \quad (137)\]

Bringing both terms to the same side, we have

\[\frac{d}{dt} \left( \sum_{i=1}^{3N} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i \right) = L \quad (138)\]

so that

\[E \equiv \sum_{i=1}^{3N} \left( \frac{\partial L}{\partial \dot{q}_i} \right) \dot{q}_i - L = \text{const.} \quad (139)\]

The quantity \(E\) is called the energy, and it is conserved.

Of course we knew that, but we’re used to writing the energy as \(T + U\).

Let’s see if we can recover this form for \(E\). When \(L\) is independent of time, it takes the form (in general coordinates):

\[L = T - U \quad (140)\]

\[= \frac{1}{2} \sum_{b,c=1}^{N} a_{bc} \dot{q}_b \dot{q}_c - U(q) \quad (141)\]

To compute \(E\) we need the derivative

\[\sum_{i=1}^{3N} \left( \frac{\partial L}{\partial \dot{q}_i} \right) = \sum_{i=1}^{3N} \left( \frac{\partial (T(q, \dot{q}) - U(q))}{\partial \dot{q}_i} \right) \dot{q}_i \quad (142)\]

\[= \sum_{i=1}^{3N} \left( \frac{\partial T(q, \dot{q})}{\partial \dot{q}_i} \right) \dot{q}_i \quad (143)\]

\[= \sum_{i=1}^{3N} \left( \frac{1}{2} \sum_{b,c=1}^{N} a_{bc} \dot{q}_b \dot{q}_c \right) \dot{q}_i \quad (144)\]
\[ \delta L = \sum_{a=1}^{N} \frac{\partial L}{\partial x_a} \cdot \delta x_a = \sum_{a=1}^{N} \frac{\partial L}{\partial x_a} \cdot \epsilon \]

\[ = \sum_{a=1}^{N} \left( \epsilon \cdot \frac{\partial}{\partial x_a} \right) L \]

The operator \( \left( \epsilon \cdot \frac{\partial}{\partial x_a} \right) \) is just the directional derivative in the \( \epsilon \) direction.

Now suppose that, for some direction \( \epsilon \), \( \delta L = 0 \). If we take the dot product of \( \epsilon \) with the Lagrange equations, we have

\[ 0 = \epsilon \cdot \left( \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}_a} \right) - \frac{\partial L}{\partial x_a} \right) \]

\[ = \frac{d}{dt} \left( \epsilon \cdot \frac{\partial L}{\partial \dot{x}_a} \right) \]

Summing over all particles

\[ 0 = \sum_a \epsilon \cdot \left( \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}_a} \right) - \frac{\partial L}{\partial x_a} \right) \]
we see the quantity
\[ \epsilon \cdot P = \epsilon \cdot \sum_a \frac{\partial L}{\partial \dot{x}_a} \] (157)
is conserved. This is the component of momentum in the \( \epsilon \) direction. If the Lagrangian is invariant under every constant displacement, then
\[ 0 = \delta L = \sum_{a=1}^{N} \frac{\partial L}{\partial x_a} \cdot \epsilon \] (158)
\[ = \epsilon \cdot \sum_{a=1}^{N} \frac{\partial L}{\partial x_a} \] (159)
for every \( \epsilon \). We conclude that
\[ \sum_{a=1}^{N} \frac{\partial L}{\partial x_a} = 0 \] (160)
and use of the Lagrange equations leads to
\[ P = \sum_a \frac{\partial L}{\partial \dot{x}_a} = \text{const.} \] (161)
When this is true we say that momentum is conserved.
As for the energy, we can recover the usual expression for the momentum by substituting the typical form of \( L \). We find:
\[ P = \sum_a \frac{\partial L}{\partial \dot{x}_a} \] (162)
\[ = \sum_a \frac{\partial}{\partial \dot{x}_a} \left( \frac{1}{2} \sum_{b,c=1}^{N} a_{bc} \dot{x}_b \dot{x}_c - U(q) \right) \] (163)
\[ = \sum_a \left( \frac{1}{2} \sum_{b,c=1}^{N} a_{bc} \delta_{ab} \dot{x}_c + \frac{1}{2} \sum_{b,c=1}^{N} a_{bc} \dot{x}_b \delta_{ac} \right) \] (164)
Now, noticing from the definition of $a_{bc}$ that it is symmetric, $a_{bc} = a_{cb}$, we see that these last two terms are the same. Therefore, the generalized momentum,

$$\mathbf{P} = \sum_a \left( \sum_{b=1}^N a_{ab} \dot{x}_b \right)$$

is conserved. $\mathbf{P}$ takes a more familiar form when we write it in Cartesian coordinates where $a_{ab} = m \delta_{ab}$ and the kinetic energy $T$ is just $\frac{1}{2} \sum_{a=1}^N m_a (\dot{x}_a)^2$. We then have the usual expression,

$$\mathbf{P} = \sum_a m \dot{x}_a$$

Notice that whichever form applies, the momentum is additive: the total momentum is just the sum of the momenta of all of the different particles, as long as we define the generalized momentum of a single particle to be

$$\mathbf{p}_a = \frac{\partial L}{\partial \dot{x}_a}$$

which reduces to the usual expression, $m_a v_a$ in Cartesian coordinates. Also notice that then generalized momentum sometimes depends on position! This may seem surprising, but recall the simple case of a mass swung in a circle on a string. There, the momentum is $mr \dot{\theta}$, which depends on the coordinate $r$ as well as the generalized velocity $\dot{\theta}$. We will see other examples later.

Finally, suppose that the component of momentum in the $\epsilon$ direction is constant. Then using the Euler-Lagrange equation,

$$0 = \frac{d}{dt} (\epsilon \cdot \mathbf{p}_a)$$

$$= \frac{d}{dt} \left( \epsilon \cdot \frac{\partial L}{\partial \dot{x}_a} \right)$$

$$= \epsilon \cdot \frac{\partial L}{\partial x_a}$$

\[27\]
In Cartesian coordinates, this is simply the absence of force in the $\varepsilon$ direction

$$0 = \varepsilon \cdot \frac{\partial L}{\partial \dot{x}_a} = \varepsilon \cdot \left( -\frac{\partial U}{\partial x_a} \right) = \varepsilon \cdot F$$

If this holds for all directions $\varepsilon$, then the total force vanishes,

$$\sum_{a=1}^{N} F_a = -\sum_{a=1}^{N} \frac{\partial U}{\partial x_a} = 0$$

(168)

The result is equivalent to Newton’s third law.

### 2.3 Center of Mass

Notice that the homogeneity of space guarantees that the center of mass of a closed system is conserved. However, since the total momentum depends on velocity, the value of the total momentum will be different in different inertial frames of reference. Consider two intertial frames, $K_1$ and $K_2$, with $K_2$ moving with velocity $u$ relative to $K_1$. Then the velocity $V$ of any particle in $K_2$ is related to that particle’s velocity $v$ in $K_1$ by

$$V = v + u$$

Therefore, since the total momentum in each of two frames $K_1$ and $K_2$ is given by

$$P_1 = \sum_{a=1}^{N} m_a v_a$$

$$P_2 = \sum_{a=1}^{N} m_a v_a$$

we can express $P_2$ in terms of $P_1$ by

$$P_2 = \sum_{a=1}^{N} m_a (v_a + u) = \sum_{a=1}^{N} m_a v_a + \sum_{a=1}^{N} m_a u$$

= $P_1 + \sum_{a=1}^{N} m_a u$
Let’s define the total mass of the system by

\[ M \equiv \sum_{a=1}^{N} m_a \]

This expresses the additivity of mass. Then we can write this relationship as

\[ \mathbf{P}_2 = \mathbf{P}_1 + M \mathbf{u} \]

where \( \mathbf{u} \) is the velocity of \( K_2 \) relative to \( K_1 \).

Since \( \mathbf{u} \) may be any velocity vector at all, we can always find a frame, \( K_{CM} \), in which the total momentum \( \mathbf{P}_{CM} \) is zero. Starting with \( \mathbf{P}_1 \), we transform to that inertial frame \( K_0 \) moving relative to frame \( K_1 \) with velocity

\[ \mathbf{u}_0 \equiv -\frac{\mathbf{P}_1}{M} \]

and find that

\[ \mathbf{P}_{CM} = \mathbf{P}_1 + M \mathbf{u}_0 \]

\[ = \mathbf{P}_1 + M \left( -\frac{\mathbf{P}_1}{M} \right) \]

\[ = 0 \]

The system is said to be at rest with respect to this new frame. The velocity of the system in any other frame \( K \) is now just \( \mathbf{u} \), where \( \mathbf{u} \) is the velocity of the frame \( K \) relative to the frame \( K_{CM} \), while the total momentum in \( \mathbf{P} = M \mathbf{u} \). Continuing in a general frame \( K \), we can write the velocity of the system as

\[ \mathbf{u} = \frac{\mathbf{P}}{M} \]

\[ = \frac{1}{M} \sum_{a=1}^{N} m_a \mathbf{v}_a \]

\[ = \frac{1}{M} \sum_{a=1}^{N} m_a \frac{d\mathbf{x}_a}{dt} \]

\[ = \frac{d}{dt} \left( \frac{1}{M} \sum_{a=1}^{N} m_a \mathbf{x}_a \right) \]
The term in parentheses is a *mass-weighted average position vector* for the system. It is called the *center of mass*:

\[
R = \left( \frac{1}{M} \sum_{a=1}^{N} m_a x_a \right)
\]

The velocity of the system in \( K \) is the time rate of change of the center of mass:

\[
u = \frac{dR}{dt}
\]

Since \( u \) is constant, we can integrate this equation to find the motion of the center of mass:

\[R = R_0 + ut\]

This says that the center of mass moves in a straight line at constant speed. This statement is equivalent to the conservation of momentum.

In the center of mass frame \( K_{CM} \), we have \( u = 0 \), and the center of mass is at a fixed position. We can then perform a translation of the origin in \( K_{CM} \) by \(-R_0\) so that the center of mass position is fixed at \( \mathbf{0} \).

The energy of a mechanical system in its rest frame, \( K_{CM} \), is called its *internal energy*. Let the internal energy in the rest frame be

\[E_i = \sum_a \frac{1}{2} m_a v_a^2 + U\]

Then, in any other frame, the velocity \( v_a \) of each particle as

\[V_a = v_a + u\]

where \( V_a \) is the velocity of the particle in \( K \) and \( u \) is the velocity of \( K \) relative to \( K_{CM} \). We can therefore write the total energy as

\[E = \sum_a \frac{1}{2} m_a (v_a + u)^2 + U\]

\[= \sum_a \frac{1}{2} m_a (v_a^2 + 2u \cdot v_a + u_0^2) + U\]

\[= \frac{1}{2} \sum_a m_a v_a^2 + u \cdot \sum_a m_a v_a + \frac{1}{2} \sum_a m_a u^2 + U\]
Therefore, in a general frame, the energy is the sum of four terms. The first and last sum to give the internal energy. The second term is the relative velocity of \( K \) and \( K_{CM} \), dotted into the average momentum in the center of mass frame,

\[
\sum_a m_a \mathbf{v}_a = \mathbf{P}_{CM} = 0
\]

(this condition is how we define the center of mass frame!). The third term is just the contribution to the energy in the general frame due to the center of mass motion.

Therefore, the energy in the center of mass frame, and hence the internal energy \( E_i \) is:

\[
E_i = \frac{1}{2} \sum a m_a \mathbf{v}_a^2 + U
\]

and we can write the total energy as

\[
E = E_i + \frac{1}{2} M \mathbf{u}^2
\]

### 2.4 Angular momentum

Conservation of angular momentum follows from the isotropy of space. This means that there is no preferred direction in space. As a consequence, an experiment will yield the same results no matter how we rotate it. This, in turn, means that under an infinitesimal rotation, \( \delta \varphi \) there should be no change in the Lagrangian, \( \delta L \).

Let’s implement this last conclusion. Let \( \mathbf{r}_a \) be the location of a particle relative to the center of rotation and let \( \delta \varphi \) be a vector directed along the axis of the rotation with magnitude equal to the infinitesimal angle of rotation. The position vector \( \mathbf{r}_a \) will change by (see figure 5 in L&L)

\[
\delta \mathbf{r}_a = \delta \varphi \times \mathbf{r}_a
\]

Since the angle \( \delta \varphi \) is constant, the change in velocity is just

\[
\delta \mathbf{v}_a = \frac{d}{dt} (\delta \mathbf{r}_a)
= \delta \varphi \times \dot{\mathbf{r}}_a
= \delta \varphi \times \mathbf{v}_a
\]
With these we can compute the change in $L$

\[
\delta L = \sum_{a=1}^{N} \frac{\partial L}{\partial \dot{r}_a} \delta \dot{r}_a + \sum_{a=1}^{N} \frac{\partial L}{\partial \dot{r}_a} \delta \dot{r}_a
\]

\[
= \sum_{a=1}^{N} \frac{\partial L}{\partial \dot{r}_a} \cdot (\delta \varphi \times \dot{r}_a) + \sum_{a=1}^{N} \frac{\partial L}{\partial \dot{r}_a} \cdot (\delta \varphi \times v_a)
\]

Using the definition of the generalized momentum,

\[
p_a = \frac{\partial L}{\partial \dot{r}_a}
\]

and the Lagrange equation,

\[
\frac{\partial L}{\partial \dot{r}_a} = \frac{d}{dt} \frac{\partial L}{\partial \dot{r}_a} = \frac{dp_a}{dt}
\]

the change in $L$ becomes

\[
\delta L = \sum_{a=1}^{N} \frac{dp_a}{dt} \cdot (\delta \varphi \times \dot{r}_a) + \sum_{a=1}^{N} p_a \cdot (\delta \varphi \times v_a)
\]

Using the cyclic property of the triple product,

\[
a \cdot (b \times c) = c \cdot (a \times b) = b \cdot (c \times a)
\]

we rewrite $\delta L$:

\[
\delta L = \sum_{a=1}^{N} \delta \varphi \cdot \left( \dot{r}_a \times \frac{dp_a}{dt} \right) + \sum_{a=1}^{N} \delta \varphi \cdot (v_a \times p_a)
\]

\[
= \delta \varphi \cdot \sum_{a=1}^{N} \left( \dot{r}_a \times \frac{dp_a}{dt} + v_a \times p_a \right)
\]

\[
= \delta \varphi \cdot \frac{d}{dt} \left( \sum_{a=1}^{N} (r_a \times p_a) \right)
\]

Now, suppose that the Lagrangian is unchanged by the rotation. Then we can set $\delta L = 0$. If this is true for arbitrary values of $\delta \varphi$ then we must have

\[
\frac{d}{dt} \left( \sum_{a=1}^{N} r_a \times p_a \right) = 0
\]
The term in parentheses is therefore a constant of the motion. We define the total angular momentum:

\[ \mathbf{M} = \sum_{a=1}^{N} (\mathbf{r}_a \times \mathbf{p}_a) = \text{const.} \]

The total angular momentum is conserved for any closed system. Notice that \( \mathbf{M} \) is additive: we compute \( \mathbf{r}_a \times \mathbf{p}_a \) for each particle, and add the results.

The (constant) value of the total angular momentum is different in different inertial frames of reference. If we change the origin of the inertial frame, so that each position vector is changed by the same constant vector, \( \mathbf{r}_a' = \mathbf{r}_a + \mathbf{a} \) then the total angular momentum changes to:

\[
\mathbf{M}' = \sum_{a=1}^{N} (\mathbf{r}_a' \times \mathbf{p}_a) \\
= \sum_{a=1}^{N} ((\mathbf{r}_a + \mathbf{a}) \times \mathbf{p}_a) \\
= \mathbf{M} + \sum_{a=1}^{N} (\mathbf{a} \times \mathbf{p}_a) \\
= \mathbf{M} + \mathbf{a} \times \sum_{a=1}^{N} \mathbf{p}_a \\
= \mathbf{M} + \mathbf{a} \times \mathbf{P}
\]

where \( \mathbf{P} \) is the total linear momentum. Therefore the only frames in which the total angular momentum doesn’t depend on the origin of the coordinate system are those with \( \mathbf{P} = 0 \), that is, one of the rest frames of the system.

If we change the motion of the frame of reference instead of the origin, then we replace \( \mathbf{v}_a' = \mathbf{v}_a + \mathbf{u} \) to find

\[
\mathbf{M}' = \sum_{a=1}^{N} (\mathbf{r}_a \times \mathbf{p}_a')
\]

33
\[
\sum_{a=1}^{N} m_a \mathbf{r}_a \times \mathbf{v}'_a = \sum_{a=1}^{N} m_a \mathbf{r}_a \times (\mathbf{v}_a + \mathbf{u}) = \mathbf{M} + \sum_{a=1}^{N} m_a \mathbf{r}_a \times \mathbf{u} = \mathbf{M} + \mathbf{M} \mathbf{R} \times \mathbf{u}
\]

Just like the energy, we can talk about the \textit{intrinsic angular momentum} of the system as the value of \( \mathbf{M} \) in the center of mass system. If we regard \( \mathbf{M} = \mathbf{M}_{\text{int}} \) as this intrinsic angular momentum then \( \mathbf{u} \) is the average velocity of the system in the primed frame and \( \mathbf{P} = \mu \mathbf{u} \) is the total linear momentum of the system in the primed system. So we can write

\[ \mathbf{M}' = \mathbf{M}_{\text{int}} + \mathbf{R} \times \mathbf{P} \]

Like the total linear momentum, it is possible for only one or more components of total angular momentum to be conserved. Thus, for cylindrical symmetry, we have conserved angular momentum about the symmetry axis, while for spherical symmetry, the angular momentum about the center of the symmetry is conserved.

### 2.5 Scale Invariance (“Mechanical similarity”)

Frequently, a mechanical problem will allow some variables to be scaled by a constant factor. For example, consider a simple harmonic oscillator. Since the potential energy is \( \frac{1}{2}kx^2 \), the Lagrangian is

\[ L = \frac{1}{2}mv^2 - \frac{1}{2}kx^2 \quad (169) \]

Suppose we rescale distances, \( x \), by a factor \( \alpha \), so that we replace

\[
\begin{align*}
    x & \to \alpha x \\
    v & = \frac{dx}{dt} \to \alpha \frac{dx}{dt} = \alpha v
\end{align*}
\]

Then the Lagrangian changes by an \textit{overall} factor of \( \alpha^2 \) :

\[ L \to \frac{1}{2}m (\alpha v)^2 - \frac{1}{2}k (\alpha x)^2 = \alpha^2 L \quad (172) \]
But we know that an overall constant doesn’t change the equation of motion. Therefore, the solution of the rescaled problem can be found from the solution to the original problem. Specifically, the solution to the original problem is

$$x = x_0 \sin \sqrt{\frac{k}{m}} t$$

(173)

so if we set $y = \alpha x$ the solution of the analogous problem for $y$ is

$$y = \alpha x = \alpha x_0 \sin \sqrt{\frac{k}{m}} t$$

(174)

$$= y_0 \sin \sqrt{\frac{k}{m}} t$$

(175)

This trivial example simply shows that the motion of a simple oscillator is independent of its initial amplitude. Shortly we will show some more interesting examples. First, let’s work out what happens in general.

Suppose the potential energy for a system of particles satisfies the scaling condition

$$U(\alpha x_1, \alpha x_2, \ldots, \alpha x_N) = \alpha^k U(x_1, x_2, \ldots, x_N)$$

(176)

for some value of $k$. Any function with this property is called homogeneous of degree $k$. This is true for any power law, $U(x) = Ax^k$, but works for many other functions as well, for example,

$$U(x_1, x_2, x_3) = \frac{(x_3 \cdot x_3)^s}{\sqrt{(x_1 + x_2) \cdot (x_1 + x_2)}}$$

(177)

since

$$U(\alpha x_1, \alpha x_2, \alpha x_3) = \frac{(\alpha x_3 \cdot x_3)^s}{\sqrt{(\alpha x_1 + \alpha x_2) \cdot (\alpha x_1 + \alpha x_2)}} = \alpha^{2s} \frac{(x_3 \cdot x_3)^s}{\alpha \sqrt{(x_1 + x_2) \cdot (x_1 + x_2)}}$$

(178)

$$= \alpha^{2s-1} U(x_1, x_2, x_3)$$

(179)

In this example, $U$ is homogeneous of degree $2s - 1$.

Now, suppose we have a potential which is a homogeneous function of order $k$. We can rescale the distance and time by arbitrary factors $\alpha$ and
\[ \beta : \]
\[ \begin{align*}
    x_a & \rightarrow \alpha x_a \\
    t & \rightarrow \beta t
\end{align*} \]  
(181)  
(182)

Then velocities will scale to
\[ v_a = \frac{dx_a}{dt} \rightarrow \frac{d(\alpha x_a)}{d(\beta t)} = \frac{\alpha}{\beta} v_a \]  
(183)

and the Lagrangian
\[ L = \sum_a \frac{1}{2} m_a v_a^2 - U(x_1, x_2, \ldots, x_N) \]  
(184)

changes to:
\[ L' = \sum_a \frac{1}{2} m_a \left( \frac{\alpha}{\beta} v_a \right)^2 - U(\alpha x_1, \alpha x_2, \ldots, \alpha x_N) \]  
(185)
\[ = \left( \frac{\alpha}{\beta} \right)^2 \sum_a \frac{1}{2} m_a v_a^2 - \alpha^k U(x_1, x_2, \ldots, x_N) \]  
(186)

We demand that the new Lagrangian be a multiple of the old one. This will be true if and only if
\[ \left( \frac{\alpha}{\beta} \right)^2 = \alpha^k \]  
(187)
\[ \beta^2 = \alpha^{2-k} \]  
(188)
\[ \beta = \alpha^{1 - \frac{k}{2}} \]  
(189)

and then we have
\[ L' = \alpha^k L \]  
(190)

We can relate all of the primed and unprimed quantities. For any lengths (which transform like \( x_a \)) we have
\[ l' = \alpha l \]  
(191)

or
\[ \frac{l'}{l} = \alpha \]  
(192)
We write the other quantities in terms of this one. Thus:

\[
\frac{l'}{l} = \beta = \alpha^{1 - \frac{k}{2}} = \left(\frac{l'}{l}\right)^{1 - \frac{k}{2}}
\]

(193)

\[
\frac{E'}{E} = \frac{T' + U'}{T + U} = \alpha^k = \left(\frac{l'}{l}\right)^k
\]

(194)

\[
\frac{M'}{M} = \frac{(\sum r \times p)'}{(\sum r \times p)} = \frac{\alpha^2}{\beta} = \left(\frac{l'}{l}\right)^{2 - 1 + \frac{k}{2}} = \left(\frac{l'}{l}\right)^{1 + \frac{k}{2}}
\]

(195)

and so on.

Invariance under rescaling of the variables may be done with the parameters of the Lagrangian as well. Thus, we could rescale the mass by a factor, or any parameters in the potential. The only requirement is that the net effect of all rescalings must leave the Lagrangian changed by no more than an overall constant.

Let’s look at some examples.

For gravity near Earth’s surface, the force is a constant \( mg \), and the potential is therefore linear,

\[
U(z) = mgz
\]

(196)

This means that \( k = 1 \) for scalings of the coordinate \( z \). For a pendulum moving in this potential, we can immediately find the dependence of the period on the length. When the length changes from \( l \) to \( l' \), corresponding times will change according to

\[
\frac{l'}{l} = \left(\frac{l'}{l}\right)^{1 - \frac{k}{2}} = \left(\frac{l'}{l}\right)^{1 - \frac{1}{2}} = \left(\frac{l'}{l}\right)^{\frac{1}{2}}
\]

(197)

(198)

Suppose we know the period \( T_0(= t) \) when the length is \( L_0(= l) \). Then the period \( T(= t') \) at any length \( L(= l') \) is

\[
T = T_0 \sqrt{\frac{L}{L_0}} = \sqrt{\frac{L}{\lambda}}
\]

(199)

where the constant \( \lambda = \frac{\sqrt{L}}{T_0} \) is just the gravitational acceleration, \( g \).
We can also use scaling to find the dependence on the local acceleration of gravity. Suppose we rescale $g$ by a factor $\lambda$. Then the potential changes by the same factor,

\[
\frac{U'}{U} = \lambda
\]

In order for the Lagrangian to remain invariant, we can rescale the time so that the kinetic energy also scales by $\lambda$. This means that we set

\[
\frac{t'}{t} = \frac{1}{\sqrt{\lambda}}
\]

because then

\[
KE' = \frac{1}{2}m\left(\frac{dx}{dt'}\right)^2 = \lambda \times KE
\]

If the period of a pendulum is $T$ in the unprimed variables, it is therefore $\frac{1}{\sqrt{\lambda}}$ in the primed variables. This means that we can write

\[
\frac{T'}{T} = \left(\frac{g}{g'}\right)^2
\]

Combining this with the dependence on length found above,

\[
\frac{T'}{T} = \left(\frac{L'}{L}\right)^{\frac{1}{2}}
\]

we deduce

\[
T \propto \sqrt{\frac{L}{g}}
\]

All this without solving an equation!

Now consider an inverse square law, either for the Coulomb force or Newton’s gravitational attraction, so that the potential is given by

\[
U(x_1, x_2) = -\frac{K}{\sqrt{(x_1 - x_2)^2}}
\]  

with $K$ an appropriate constant ($G_{\text{Newton}} M m$ or $k_{\text{Coulomb}} q q'$ for example). Then, noting that $k = -1$, we can easily find the relationship between the orbital period and the size of an orbit. If we increase the size of the orbit
(as measured, say, by the greatest separation of two points on the orbit) from \( L \) to \( L_0 \), then the periods will scale by

\[
\frac{T}{T_0} = \left( \frac{L}{L_0} \right)^{1-\frac{3}{2}} = \left( \frac{L}{L_0} \right)^{1-\frac{3}{2}} \tag{202}
\]

\[
= \left( \frac{L}{L_0} \right)^{\frac{3}{2}} \tag{203}
\]

which we can write as

\[
T^2 = \lambda L^3 \tag{204}
\]

We recognize this immediately as Kepler’s law.

This technique is used to good advantage in quantum field theory, where it is related to renormalization. One derives the appropriate form of the renormalization group equation to specify how various properties of the system change under certain rescalings of the parameters.

### 2.6 Virial theorem

Finally, let’s look at the *virial theorem*. The kinetic energy satisfies

\[
\sum v_a \cdot \frac{\partial T}{\partial v_a} = \sum_a v_a \cdot \frac{\partial}{\partial v_a} \left( \frac{1}{2} \sum_b m_b v_b^2 \right) \tag{205}
\]

\[
= \sum_a v_a \cdot \left( \frac{1}{2} \sum_b 2m_b v_b \delta_{ab} \right) \tag{206}
\]

\[
= \sum_a v_a \cdot (m_a v_a) \tag{207}
\]

\[
= 2T \tag{208}
\]

Therefore, using the definition of the momentum, then using the product rule,

\[
2T = \sum v_a \cdot \frac{\partial T}{\partial v_a} \tag{209}
\]

\[
= \sum v_a \cdot \frac{\partial L}{\partial v_a} \tag{210}
\]

\[
= \sum v_a \cdot p_a \tag{211}
\]
\[
\sum \frac{dr_a}{dt} \cdot p_a = \sum \frac{dp_a}{dt}
\]

(212)

\[
\sum dr_a \cdot p_a - \sum r_a \cdot \frac{dp_a}{dt}
\]

(213)

Now, let’s average this over time, for a long time:

\[
2\bar{T} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau 2T
\]

(214)

\[
= \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \left( \frac{d}{dt} \sum r_a \cdot p_a - \sum r_a \cdot \frac{dp_a}{dt} \right)
\]

(215)

\[
= \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \frac{d}{dt} \sum r_a \cdot p_a - \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \sum r_a \cdot \frac{dp_a}{dt}
\]

(216)

\[
= \lim_{\tau \to \infty} \frac{1}{\tau} \left( \sum r_a(\tau) \cdot p_a(\tau) - \sum r_a(0) \cdot p_a(0) \right)
\]

(217)

\[
- \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \sum r_a \cdot \frac{dp_a}{dt}
\]

(218)

Suppose the motion is bounded. Then neither the position nor the momentum ever exceed some finite values and because the denominator of the first term \((\tau)\) becomes arbitrarily large, the first term vanishes. We then have

\[
2\bar{T} = - \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \sum r_a \cdot \frac{dp_a}{dt}
\]

(219)

Next, using Newton’s second law,

\[
\frac{dp_a}{dt} = - \frac{\partial U}{\partial r_a}
\]

(220)

we write the integral as

\[
2\bar{T} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \sum r_a \cdot \frac{\partial U}{\partial r_a}
\]

(221)

Now we need Euler’s theorem for homogeneous functions. A homogeneous function which satisfies \(U(\alpha r_a) = \alpha^k U(r_a)\) is said to be homogeneous of degree \(k\). Euler’s theorem then says that

\[
\sum r_a \cdot \frac{\partial U}{\partial r_a} = \alpha U(r_a)
\]

(222)
If you want to just accept this, you can skip the next paragraph, but I’m going to go ahead here and try to prove Euler’s theorem for a function of one variable. If you want an extra problem, try to generalize my proof to more than one variable!

So let’s try to prove Euler’s theorem for a homogeneous function of degree \( k \). Let \( f(\alpha x) = \alpha^k f(x) \) and compute

\[
x \frac{df}{dx} = x \lim_{\varepsilon \to 0} \frac{f(x + \varepsilon) - f(x)}{\varepsilon}
\] (223)

Now at any fixed \( x \) we can let \( \alpha x = x + \varepsilon \), so that \( \alpha = 1 + \frac{\varepsilon}{x} \). Then

\[
x \frac{df}{dx} = x \lim_{\varepsilon \to 0} \frac{f(x + \varepsilon) - f(x)}{\varepsilon}
= x \lim_{\varepsilon \to 0} \frac{f(\alpha x) - f(x)}{\varepsilon}
= x \lim_{\varepsilon \to 0} \frac{\alpha^k f(x) - f(x)}{\varepsilon}
= x \lim_{\varepsilon \to 0} \frac{(1 + \frac{\varepsilon}{x})^k f(x) - f(x)}{\varepsilon}
= x \lim_{\varepsilon \to 0} \frac{\frac{k\varepsilon}{x} f(x)}{\varepsilon}
= k f(x)
\] (224)

(225)

(226)

(227)

(228)

So we can see what’s happening. We just use the factor \( \alpha \) to produce the extra \( \varepsilon \), then we can pull it out of the argument. Then, putting everything back in terms of \( \varepsilon \) we get the result.

Now let’s get back to the Virial Theorem. We showed that

\[
2\bar{T} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \sum \mathbf{r}_a \cdot \frac{\partial U}{\partial \mathbf{r}_a}
\] (231)

and we now assume that \( U \) is homogeneous of degree \( k \). This means that the sum can be replaced by \( kU \):

\[
2\bar{T} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau kU(r_a)
= k \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau U(r_a)
= k\bar{U}
\] (232)
and we have the result that twice the average kinetic energy equals $k$ times
the average potential energy:

$$2\bar{T} = k\bar{U} \quad (235)$$

For the harmonic oscillator this means that the potential and kinetic energy
are equal, while for an inverse square law the average kinetic energy is minus
one-half the average potential, $\bar{T} = -\frac{1}{2}\bar{U}$. 
3 Integration of the equations of motion

Now we’ll solve some problems.

3.1 Motion in one dimension

This whole class of problems is soluble. In Cartesian coordinates we have

\[ L = \frac{1}{2} m \dot{x}^2 - U(x) \]  

(236)

The easiest way to solve this is to use the conservation of energy (there is no angular momentum in one dimension, and since the general potential \( U(x) \) is external, momentum is not conserved). Energy conservation gives:

\[ E = \frac{1}{2} m \dot{x}^2 + U(x) \]  

(237)

where \( E \) is constant. Solving for the velocity, we get

\[ \dot{x} = \frac{dx}{dt} = \sqrt{\frac{2}{m} (E - U(x))} \]  

(238)

so we can solve:

\[ t = \int_{t_0}^{x(t)} \frac{dx}{\sqrt{\frac{2}{m} (E - U(x))}} \]  

(239)

There are essentially two types of motion possible, depending on what happens to the factor \( E - U(x) \) in the denominator. From energy conservation and the fact that the kinetic energy term is necessarily non-negative, we have

\[ E = \frac{1}{2} m \dot{x}^2 + U(x) \geq U(x) \]  

(240)

This puts limits on what values of \( x \) are possible. Any value of \( x \) for which \( U(x) \) is larger than the total energy is forbidden. Upon reaching such a value, the system must turn around and go back the way it came. Thus, any values of \( x \) such that

\[ E = U(x) \]  

(241)

are called turning points. At any turning point, the velocity must be zero. If the motion occurs in a region bounded by two turning points, then to
motion says in that finite region of space and is called finite. If there are
no turning points, or only one, the motion can continue indefinitely and is
called infinite.

For finite motion, the particle will generally move back and forth over
the region. We can ask for the period of this motion. If the two turning
points are given by \( x_1 \) and \( x_2 \), the period is given by twice the time it takes
to go from one to the other:

\[
T = \sqrt{2m} \int_{x_1}^{x_2} \frac{dx}{\sqrt{E - U(x)}}
\]

(242)

It is easy to tell from a graph of \( U(x) \) what the possible motions are.

Notice that we could also have gotten the result by directly integrating
the equation of motion. Starting with

\[
\frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} = 0
\]

leads to

\[
m \frac{d\dot{x}}{dt} = -\frac{dU}{dx}
\]

This can be integrated for any potential \( U(x) \) by multiplying both sides by
the velocity:

\[
m \dot{x} \frac{d\dot{x}}{dt} = -\frac{dU}{dx} \frac{dx}{dt}
\]

\[
m \int \dot{x} d\dot{x} = -\int \frac{dU}{dx} dx
\]

The integrations immediately give conservation of energy.

### 3.2 Determination of the potential from the period of oscillation

Suppose we know the period of a 1-dimensional system for every value of
the energy, so we have the function \( T(E) \). There is a clever way to ex-
tract information about the potential. A similar method is sometimes used
for scattering problems to find the scattering potential from the angular
distribution of scattering. We’ll follow the text closely here.
For simplicity, suppose $U(x)$ has only one minimum. Choose the coordinates so that the minimum of the potential is at the origin, and adjust the energy scale so that the minimum value of $U(x)$ is zero. We want to change variables in the equation for the period and regard $x$ as a function of $U$. Then $x(U)$ is double valued between the turning points $x_1$ and $x_2$, since the motion runs over the same values of $U$ when going from $x_1$ to 0 as it does when going from 0 to $x_2$. Each value of $U$ gives a value of $x$ in each of these halves of the motion. Therefore, to get a unique integral, we divide the integral into two parts:

$$T = \sqrt{2m} \int_{x_1}^{x_2} \frac{dx}{\sqrt{E - U(x)}}$$

$$= \sqrt{2m} \int_{x_1}^{0} \frac{dx}{\sqrt{E - U(x)}} + \sqrt{2m} \int_{0}^{x_2} \frac{dx}{\sqrt{E - U(x)}}$$

$$= \sqrt{2m} \int_{x_1}^{0} \frac{dx(U)}{\sqrt{E - U}} + \sqrt{2m} \int_{0}^{x_2} \frac{dx}{\sqrt{E - U}} \quad (243)$$

On each of these intervals, $U(x)$ is single-valued, so we can define the inverse function, $x(U)$. Call the two functions we get this way $x_1(U)$ and $x_2(U)$. The

$$T = \sqrt{2m} \int_{x_1}^{0} \frac{dx_1(U)}{\sqrt{E - U}} + \sqrt{2m} \int_{0}^{x_2} \frac{dx_2(U)}{\sqrt{E - U}}$$

$$= \sqrt{2m} \int_{E=U(x_1)}^{0} \frac{dx_1(U)}{dU} \frac{dU}{\sqrt{E - U}}$$

$$+ \sqrt{2m} \int_{E=U(x_2)}^{0} \frac{dx_2(U)}{dU} \frac{dU}{\sqrt{E - U}} \quad (244)$$

Now multiply by $\frac{1}{\sqrt{\alpha - E}}$ and integrate over energy from 0 to $\alpha$

$$I = \int_{0}^{\alpha} dE \frac{1}{\sqrt{\alpha - E}} T(E)$$

$$= \sqrt{2m} \int_{0}^{\alpha} dE \int_{E}^{0} \frac{1}{\sqrt{\alpha - E}} \frac{dx_1(U)}{dU} \frac{dU}{\sqrt{E - U}}$$

$$+ \sqrt{2m} \int_{0}^{\alpha} dE \int_{0}^{E} \frac{1}{\sqrt{\alpha - E}} \frac{dx_2(U)}{dU} \frac{dU}{\sqrt{E - U}} \quad (246)$$

$$+ \sqrt{2m} \int_{0}^{\alpha} dE \int_{0}^{E} \frac{1}{\sqrt{\alpha - E}} \frac{dx_2(U)}{dU} \frac{dU}{\sqrt{E - U}} \quad (247)$$
Let’s look at the two (equivalent) integrals. When we interchange the order of integration, we have to adjust the lower limit on the energy integral.

\[\int_0^\alpha dE \int_0^E \frac{dx_2}{\sqrt{\alpha - E} \sqrt{E - U}} = \int_0^\alpha \frac{dx_1}{dU} dU \int_U^\alpha \frac{dE}{\sqrt{\alpha - E} \sqrt{E - U}}\]  \hspace{1cm} (248)

Now we can rewrite:

\[\int_U^\alpha \frac{dE}{\sqrt{\alpha - E} \sqrt{E - U}} = \int_0^{\alpha - U} \frac{dE'}{\sqrt{\alpha - U - E' \sqrt{E'}}}\]

\[= \int_0^\beta \frac{dE'}{\sqrt{\beta - E' \sqrt{E'}}}\]

\[= \int_0^\beta \frac{dE'}{\sqrt{\beta \sqrt{1 - E'/\beta}} \sqrt{E'/\beta}}\]

\[= \int_0^1 \frac{dz}{\sqrt{1 - z \sqrt{z}}}\]  \hspace{1cm} (249)

where \(z = \frac{E'}{E}\). Writing the integral in this form, with a dimensionless variable over a constant range, we see that its value is just some constant. We can easily find the constant. Let \(s = \sqrt{z}\). Then

\[ds = \frac{dz}{2\sqrt{z}}\]

\[\int_0^1 \frac{dz}{\sqrt{1 - z \sqrt{z}}} = \int_0^1 \frac{2ds}{\sqrt{1 - s^2}}\]  \hspace{1cm} (250)

Now substitute \(s = \sin \alpha\) and \(ds = \cos \alpha \, d\alpha\) to find

\[\int_0^1 \frac{dz}{\sqrt{1 - z \sqrt{z}}} = \int_0^1 \frac{2ds}{\sqrt{1 - s^2}}\]

\[= \int_0^{\frac{\pi}{2}} 2d\alpha\]

\[= \pi\]  \hspace{1cm} (251)

Returning to our original problem, we have

\[\int_0^\alpha \frac{T(E)dE}{\sqrt{\alpha - E}} = -\sqrt{2m\pi} \int_0^\alpha dx_1(U) + \sqrt{2m\pi} \int_0^\alpha dx_2(U)\]
\[ \sqrt{2m\pi} (x_2(\alpha) - x_1(\alpha)) = \int_0^\alpha \frac{T(E)dE}{\sqrt{\alpha - E}} \]

\[ (x_2(\alpha) - x_1(\alpha)) = \frac{1}{\pi\sqrt{2m}} \int_0^\alpha \frac{T(E)dE}{\sqrt{\alpha - E}} \]

This last equation holds for any \( \alpha \); in particular we may call the variable \( U \) instead, so

\[ x_2(U) - x_1(U) = \frac{1}{\pi\sqrt{2m}} \int_0^U \frac{T(E)dE}{\sqrt{\alpha - E}} \] (252)

The solution is not unique unless we know that \( U \) is symmetric, in which case we know that \( x_2(U) = -x_1(U) \). Defining \( x(U) = x_2(U) \) we then have

\[ x(U) = \frac{1}{2\pi\sqrt{2m}} \int_0^U \frac{T(E)dE}{\sqrt{U - E}} \] (253)

### 3.3 The reduced mass

Landau and Lifshitz omit an argument here; we can solve the problem of a central force between two particles completely, but they do not address more general forces. Let’s see if we can fill the gap. Suppose we have a closed system with two particles. For such a closed system,

\[ U = U(x_1, x_2) \] (254)

and the Lagrangian is

\[ L = \frac{1}{2} m_1 \dot{x}_1^2 + \frac{1}{2} m_2 \dot{x}_2^2 - U(x_1, x_2) \] (255)

However, since the system is isolated, total linear and angular momentum must be conserved. Therefore, in the center of mass system we may write the constants of motion:

\[ P = m_1 v_1 + m_2 v_2 \] (256)
\[ M = m_1 x_1 \times v_1 + m_2 x_2 \times v_2 \] (257)

We rewrite everything in terms of two new variables:

\[ r = x_1 - x_2 \] (258)
\[ s = x_1 + x_2 \] (259)

47
Setting the origin at the center of mass, we have
\[ m_1 x_1 + m_2 x_2 = 0 \]  \hspace{1cm} (260)
so that
\[ x_1 = -\frac{m_2}{m_1} x_2 \]
and therefore
\[ r = x_1 - x_2 = -\frac{m_2}{m_1} x_2 - x_2 = -\frac{m_2 + m_1}{m_1} x_2 \]  \hspace{1cm} (261)
or
\[ x_2 = -\frac{m_1 r}{m_1 + m_2} \]  \hspace{1cm} (262)
\[ x_1 = -\frac{m_2}{m_1} x_2 = \frac{m_2 r}{m_1 + m_2} \]  \hspace{1cm} (263)
We also find
\[ s = x_1 + x_2 = \frac{m_2 - m_1}{m_1 + m_2} r \]  \hspace{1cm} (264)
so that \( s \) in this frame of reference is not independent of \( r \). Thus, our freedom to choose the origin of our coordinate system totally eliminates \( s \) from the problem. In particular, we can write the potential as
\[ U(x_1, x_2) = U\left(\frac{m_2 r}{m_1 + m_2}, -\frac{m_1 r}{m_1 + m_2}\right) = U(r) \]  \hspace{1cm} (265)
This means that the potential depends only on the vector from one particle to the other. We can argue from isotropy that it can’t depend on the direction, but since isotropy leads us to conservation of momentum, we should be able to make the argument using the conservation law. We have
\[ 0 = \frac{d}{dt} M = \frac{d}{dt} \left( m_1 x_1 \times v_1 + m_2 x_2 \times v_2 \right) \]
\[ = m_1 \frac{dx_1}{dt} \times v_1 + m_2 \frac{dx_2}{dt} \times v_2 + m_1 x_1 \times \frac{dv_1}{dt} + m_2 x_2 \times \frac{dv_2}{dt} \]
\[ = m_1 x_1 \times \frac{dv_1}{dt} + m_2 x_2 \times \frac{dv_2}{dt} \]  \hspace{1cm} (266)
Using Newton’s second law to replace the accelerations:

\[
0 = -m_1 \dot{x}_1 \times \frac{dU}{dx_1} - m_2 \dot{x}_2 \times \frac{dU}{dx_2}
\]

\[
m_1 \dot{x}_1 \times \frac{dU}{dx_1} = -m_2 \dot{x}_2 \times \frac{dU}{dx_2}
\]

\[
m_1 m_2 \dot{r} \times \frac{dU}{dx_1} = \frac{m_1 m_2}{m_1 + m_2} \dot{r} \times \frac{dU}{dx_2}
\]

\[
0 = \dot{r} \times \left( \frac{dU}{dx_1} - \frac{dU}{dx_2} \right)
\]

\[
= \dot{r} \times \left( \frac{m_1 + m_2}{m_2} \frac{dU}{dr} + \frac{m_1 + m_2}{m_1} \frac{dU}{dr} \right)
\]

\[
= \frac{(m_1 + m_2)^2}{m_1 m_2} \dot{r} \times \frac{dU}{dr}
\]

(267)

The vanishing of the last expression means that the gradient of \( U \) must be in the direction of the vector \( \dot{r} \). If we let \( r \) be a coordinate measuring distances in the \( \dot{r} \) direction, then the gradient of \( U \) will depend only on that coordinate. Therefore, we have

\[
U = U(r)
\]

(268)

This justifies the implicit claim of Landau and Lifshitz that every two particle problem reduces to a central force problem. The Lagrangian now simplifies to

\[
L = \frac{1}{2} m_1 \dot{x}_1^2 + \frac{1}{2} m_2 \dot{x}_2^2 - U(x_1, x_2)
\]

\[
= \frac{1}{2} m_1 \left( \frac{m_2 \dot{r}}{m_1 + m_2} \right)^2 + \frac{1}{2} m_2 \left( \frac{m_1 \dot{r}}{m_1 + m_2} \right)^2 - U(r)
\]

\[
= \left( \frac{1}{2} \frac{m_1 m_2^2}{(m_1 + m_2)^2} + \frac{1}{2} \frac{m_2 m_1^2}{(m_1 + m_2)^2} \right) \dot{r}^2 - U(r)
\]

\[
= \frac{1}{2} \left( \frac{m_1 m_2 (m_1 + m_2)}{(m_1 + m_2)^2} \right) \dot{r}^2 - U(r)
\]

\[
= \frac{1}{2} m \dot{r}^2 - U(r)
\]

(269)

where we define the reduced mass,

\[
m = \frac{m_1 m_2}{m_1 + m_2}
\]

(270)
We have now reduced the two-particle problem to a one particle problem with a potential $U(r)$.

### 3.4 Motion in a central field

The problem of single particle motion with a potential dependant only on $r$ is called the central field problem. We can solve it completely.

Start with the angular momentum, $\mathbf{M} = \mathbf{r} \times \mathbf{p}$. Because $\mathbf{M}$ is conserved, it cannot change direction. If we choose coordinates so that $\mathbf{M}$ lies along the $z$ axis, then it always lies along the $z$ axis. But the motion of the particle is specified by the vector $\mathbf{r}$, and $\mathbf{r} \cdot \mathbf{M} = 0$. Therefore, $\mathbf{r}$ must lie in the $xy$ plane for the entire motion. Even if we don’t make these particular coordinate choices, it should be clear that the motion must lie in a plane orthogonal to $\mathbf{M}$.

Now, specify the position $\mathbf{r}$ in the plane of motion using polar coordinates:

$$\mathbf{r} = r(r, \varphi)$$

The velocity then has components

$$\mathbf{v} = (\dot{r}, r \dot{\varphi})$$

and the Lagrangian is

$$L = \frac{1}{2} m (\dot{r}^2 + r^2 \dot{\varphi}^2) - U(r)$$

Now we’ve got the problem down to only two variables, and we can still derive information from the magnitude of the total angular momentum. It is simplest to recognize that it is the momentum conjugate to the angular coordinate $\varphi$. Thus, noting that

$$\frac{\partial L}{\partial \varphi} = 0$$

we immediately have

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{\varphi}} = 0$$

from the Lagrange equation. This just says that the momentum conjugate to $\varphi$ is conserved because $\varphi$ is cyclic. Therefore,

$$J = |\mathbf{M}| = \frac{\partial L}{\partial \dot{\varphi}} = mr^2 \dot{\varphi} = \text{const.}$$
It is easy to see that this is the angular momentum. The relation allows us to solve for \( \dot{\varphi} \) in terms of \( r \):

\[
\dot{\varphi} = \frac{J}{mr^2}
\]

Substituting into \( L \) we have

\[
L = \frac{1}{2} m(r^2 + \frac{J^2}{m^2 r^2}) - U(r)
\]

reducing the problem to one degree of freedom – a problem which we have already solved.

Writing the energy associated with \( L \)

\[
E = T + U
\]

\[
= \frac{1}{2} m r^2 + \frac{J^2}{2mr^2} + U(r)
\]

\[
= \frac{1}{2} (mr^2 + U'(r))
\]

\[
U' = U(r) + \frac{J^2}{2mr^2}
\]

We can use this to write an integral solution for \( r(t) \) as before, just as if it were a typical one dimensional problem with potential \( U' \). Landau and Lifshitz call the contribution the “centrifugal energy”. This is appropriate, because the term is actually part of the kinetic energy of the system. American textbooks usually talk about the resulting “centrifugal force”,

\[
F_{\text{centrifugal}} = -\frac{d}{dr} \left( \frac{J^2}{2mr^2} \right) \hat{r}
\]

\[
= \frac{J^2}{2mr^3} \hat{r}
\]

The actual force provided by the true potential \( U \) in order to balance this term is the negative of this,

\[
F = -\frac{J^2}{2mr^3} \hat{r}
\]

The consequent acceleration is

\[
a = \frac{F}{m} = -\frac{J^2}{2m^2 r^3} \hat{r}
\]

51
This is called the centripetal acceleration.

Treating this as a case of one dimensional motion, we could immediately write an integral solution by using the conservation of energy,

$$E = \frac{1}{2} mr^2 + U'$$

(271)

then solving for $dt$ and integrating:

$$t = \int dt = \int_{r_0}^{r(t)} \frac{dx}{\sqrt{\frac{2}{m} (E - U - \frac{J^2}{2mr^2})}}$$

(272)

This gives the period and turning points of the motion. Notice that bounded motion oscillates in the radial coordinate $r$ between values $r_{\text{min}}$ and $r_{\text{max}}$ given by the roots of the denominator of the integrand.

However, this is not the easiest way to solve the central force problem. It turns out to be easier to understand the motion if we find the radius of the orbit as a function of the angle rather than as a function of the time. This gives us a geometric picture of the orbit. To accomplish the change of variable, we use the angular momentum again,

$$\frac{d\varphi}{dt} = \frac{J}{mr^2}$$

Now rewrite the conservation of energy in terms of $r$ and $\varphi$,

$$E = \frac{1}{2} m \left( \frac{dr}{dt} \right)^2 + U(r) + \frac{J^2}{2mr^2}$$

$$= \frac{1}{2} m \left( \frac{dr}{d\varphi} \frac{d\varphi}{dt} \right)^2 + U(r) + \frac{J^2}{2mr^2}$$

$$= \frac{1}{2} \frac{J^2}{mr^4} \left( \frac{dr}{d\varphi} \right)^2 + U(r) + \frac{J^2}{2mr^2}$$

(273)

Solving for $d\varphi$ and integrating,

$$\left( \frac{dr}{d\varphi} \right)^2 = \frac{2m^2r^4}{J^2} \left( E - U - \frac{J^2}{2mr^2} \right)$$

$$\int d\varphi = \int_{r_0}^{r(\varphi)} \frac{dr}{\sqrt{\frac{2m^2r^6}{J^2} \left( E - U - \frac{J^2}{2mr^2} \right)}}$$

$$= \int_{r_0}^{r(\varphi)} \frac{Jdr}{r^2 \sqrt{2m (E - U) - \frac{J^2}{r^2}}}$$

(274)
From this we can again immediately find the turning points of the motion, for any central potential \( U(r) \), this time as a function of angle. If the potential increases monotonically, and if it does not decrease near \( r = 0 \) faster than \( \frac{J^2}{r^2} \) increases, the effective potential \( U' \) will be a single well shape and the motion will oscillate back and forth between some minimum and some maximum values for \( r \). We can ask whether this motion closes on itself (perhaps after several complete orbits). The motion closes if and only if

\[
r(\varphi + 2\pi n) = r(\varphi)
\]

for some value of \( n \). We can show that the motion closes if \( U \) varies as \( 1/r \) or as \( r^2 \).

In particular, the motion does not close is the potential has \( 1/r^2 \) dependence. Such a contribution to the potential effectively adds a constant to the \( J^2/r^2 \) term in the integral and messes up the exact balance of coefficients required for closed orbits. Such a term arises from the corrections that general relativity gives to the gravitation law, and it is this effect that leads the the famous advance in Mercury’s perihelion.

### 3.5 Kepler’s problem

We are particularly interested in the case of an inverse square law force, given by a \( 1/r \) potential:

\[
U = -\frac{\alpha}{r}
\]

The effective potential \( U' \) is then

\[
U' = -\frac{\alpha}{r} + \frac{J^2}{2mr^2}
\]

It is easy to see that \( U' \) tends toward \( +\infty \) as \( r \) goes to zero, and that \( U' \) goes to zero at infinity. In between, there is a single minimum (as long as \( \alpha \) is positive), given by setting

\[
\frac{dU'}{dr} = \frac{\alpha}{r^2} - \frac{J^2}{mr^3} = 0
\]

Solving,

\[
r_0 = \frac{J^2}{\alpha m}
\]
The motion will oscillate asymmetrically about this value.

Now let’s perform the “elementary integration” of the equation of the orbit. We have

\[
\varphi = \int_{r_0}^{r(\varphi)} \frac{J dr}{r^2 \sqrt{2mE + \frac{2\alpha}{r} - \frac{J^2}{r^2}}} \quad (275)
\]

and it makes sense to define a new variable,

\[
u = \frac{J}{r},
\]

\[du = -\frac{J}{r^2} \, dr\]

so we have

\[
\varphi = -\int_{J/r_0}^{u=J/r(\varphi)} \frac{du}{\sqrt{2mE + \frac{2\alpha}{J}u - u^2}} \quad (276)
\]

Now complete the square in the denominator

\[
2mE + \frac{2\alpha}{J}u - u^2 = -\left(u - \frac{\alpha}{J}\right)^2 + \left(\frac{\alpha}{J}\right)^2 + 2mE
\]

\[
= -\left(u - \frac{\alpha}{J}\right)^2 + \left(\frac{\alpha}{J}\right)^2 \left(1 + \frac{2EJ^2}{\alpha^2}\right)
\]

and define a new variable \(y\) and the eccentricity, \(e\):

\[
y = u - \frac{\alpha}{J}
\]

\[e = \sqrt{1 + \frac{2EJ^2}{\alpha^2}}
\]

In terms of these the integral becomes

\[
\varphi = \int_{y_0}^{y(\varphi)} \frac{dy}{\sqrt{\left(\frac{\alpha e}{J}\right)^2 - y^2}} \quad (277)
\]

\[
= \int_{y_0}^{y(\varphi)} \frac{\left(\frac{J}{\alpha e}\right) dy}{\sqrt{1 - \left(\frac{Jy}{\alpha e}\right)^2}} \quad (278)
\]

and the obvious substitution is

\[
\frac{Jy}{\alpha e} = \sin \theta
\]
This gives

\[
\varphi = \int_{\theta_0}^{\theta(\varphi)} \frac{\cos \theta \, d\theta}{\sqrt{1 - \sin^2 \theta}} = \theta(\varphi) - \theta_0 \tag{279}
\]

\[
= \sin^{-1} \left( \frac{J_y}{mae} \right) - \theta_0 \tag{280}
\]

We can take the constant to be zero, since it just gives an initial value to the angle \( \varphi \). Instead, let’s take it to be \( \frac{\pi}{2} \) to make our answer agree with the form given in the text. This just turns the \( \sin \varphi \) into \( \cos \varphi \), that is,

\[
\sin (\varphi + \theta_0) = \sin (\varphi + \theta_0) = \sin \left( \varphi + \frac{\pi}{2} \right) = \cos \varphi
\]

Then retracing our substitutions we get

\[
\cos \varphi = \frac{J_y}{mae} \tag{281}
\]

\[
= \frac{J}{mae} \left( u - \frac{ma}{J} \right) \tag{282}
\]

\[
= \frac{J}{mae} \left( \frac{J}{r} - \frac{ma}{J} \right) \tag{283}
\]

and solving for \( r \),

\[
\frac{J}{r} = \frac{ma}{J} + \frac{mae}{J} \cos \varphi \tag{284}
\]

\[
r = \frac{ma}{J} \left( 1 + e \cos \varphi \right) \tag{285}
\]

Now let’s figure out what the shape of the orbit is. First, we can write \( r \) in terms of the constant

\[
p = \frac{J^2}{ma} > 0 \tag{286}
\]

so we have simply

\[
r = \frac{p}{1 + e \cos \varphi} \tag{287}
\]
or

\[ p = r (1 + e \cos \varphi) \quad (288) \]
\[ = \sqrt{x^2 + y^2 + ex} \quad (289) \]

since \( x = r \cos \varphi \). If we solve for the square root, then square, we have:

\[ \sqrt{x^2 + y^2} = p - ex \quad (290) \]
\[ x^2 + y^2 = p^2 - 2pex + e^2x^2 \quad (291) \]
\[ (1 - e^2) x^2 + 2pex + y^2 = p^2 \quad (292) \]

This relationship holds regardless of the magnitude of the constants.

There are now three distinct cases, depending on whether the energy is negative, zero, or positive. Since \( e = \sqrt{1 + \frac{2E_{\text{mag}}}{m}} \), the change in sign makes a crucial difference in the magnitude of the eccentricity:

\[
\begin{align*}
E < 0 & \quad e < 1 \\
E = 0 & \quad e = 1 \\
E > 0 & \quad e > 1 
\end{align*}
\quad (293)
\]

We handle each case in turn.

3.5.1 Case I: \( E < 0, \ e < 1 \)

Starting with

\[ (1 - e^2) x^2 + 2pex + y^2 = p^2 \quad (294) \]

we note that \( 1 - e^2 > 0 \) and we complete the square:

\[ (1 - e^2) x^2 + 2pex + \frac{p^2e^2}{1 - e^2} + y^2 = p^2 + \frac{p^2e^2}{1 - e^2} \quad (295) \]
\[ \left( \sqrt{1 - e^2} x + \frac{pe}{\sqrt{1 - e^2}} \right)^2 + y^2 = \frac{p^2}{1 - e^2} \quad (296) \]
\[ (1 - e^2) \left( x + \frac{pe}{1 - e^2} \right)^2 + y^2 = \frac{p^2}{1 - e^2} \quad (297) \]

If we shift the \( x \) axis by a constant

\[ x' = x + \frac{pe}{1 - e^2} \quad (298) \]
we see that this is just the equation of an ellipse

\[
(1 - e^2) (x')^2 + y^2 = \frac{p^2}{1 - e^2}
\]  

(299)

(Notice that \(x^2 + y^2 = \text{const.}\) would give a circle, and the factor \((1 - e^2)\) just contracts the \(x\) component of the motion.).

### 3.5.2 Case 2: \(E = 0\) and \(e = 1\)

This is a special borderline case. The general equation reduces:

\[
p^2 = (1 - e^2) x^2 + 2pex + y^2 = 2px + y^2
\]

(300)

(301)

so that

\[
x = -\frac{1}{2p} y^2 + \frac{p}{2}
\]

(302)

and the motion is a parabola that goes off along the negative \(x\)-axis, with the bottom curve of the parabola at \(x = p/2\). Thus, the motion is unbounded.

### 3.5.3 Case 3: \(E > 0\) and \(e > 1\)

In this case we also expect unbounded motion. Starting again with

\[
p^2 = -(e^2 - 1) x^2 + 2pex + y^2
\]

(303)

written to display the positive quantity \(e^2 - 1 > 0\), we again complete the square:

\[
p^2 = -(e^2 - 1) x^2 + 2pex + y^2
\]

(304)

\[
= -(e^2 - 1) x^2 + 2pex + \left(\frac{e^2 p^2}{e^2 - 1}\right) + \left(\frac{e^2 p^2}{e^2 - 1}\right) + y^2
\]

(305)

\[
= -\left(\sqrt{e^2 - 1}x + \frac{ep}{\sqrt{e^2 - 1}}\right)^2 + \frac{e^2 p^2}{e^2 - 1} + y^2
\]

(306)

\[
= -(e^2 - 1) \left(x + \frac{ep}{e^2 - 1}\right)^2 + \frac{e^2 p^2}{e^2 - 1} + y^2
\]

(307)

Rearranging, we have

\[
(e^2 - 1) \left(x + \frac{ep}{e^2 - 1}\right)^2 - y^2 = \frac{e^2 p^2}{e^2 - 1} - p^2 = \frac{p^2}{e^2 - 1}
\]

(308)
Once again we shift the $x$ axis,

$$x' = x + \frac{ep}{e^2 - 1} \quad (309)$$

to get the equation of an hyperbola:

$$(e^2 - 1) (x')^2 - y^2 = \frac{p^2}{e^2 - 1} \quad (310)$$

Recall that for the hyperbola lying between 45-degree diagonal lines (which therefore make an angle with one another of 90 degrees), the equation is

$$x^2 - y^2 = \text{const.} \quad (311)$$

so we have the same figure with the $x$ axis scaled. This makes the angle between the two asymptotes different from 90, but otherwise we still have the hyperbola.

### 3.5.4 Conic Sections

It is often stated in Classical Mechanics texts that the inverse square law orbits are conic sections. By *conic section* we mean any of the curves that result when a plane is intersected with a cone. These curves are easy to find. First, we write the equation of a plane. A plane through the origin may be described as the set of all vectors orthogonal to a given vector, $\mathbf{n}$, the normal to the plane:

$$\mathbf{n} \cdot \mathbf{x} = 0 \quad (312)$$

We can displace the plane by shifting the coordinates by a constant vector $\mathbf{a}$,

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{a}) = 0 \quad (313)$$

or if we define $d = -\mathbf{n} \cdot \mathbf{a}$, simply the general linear equation,

$$\mathbf{n} \cdot \mathbf{x} + d = 0 \quad (314)$$

or in terms of components, $\mathbf{n} = (a, b, c)$

$$ax + by + cz + d = 0 \quad (315)$$
Now consider a cone lying along the $z$ axis. In cylindrical coordinates it is the set of points satisfying $z = kr$ for some constant $k$. We can write this as

$$z^2 = k x^2 + k y^2$$ \hspace{1cm} (316)

Combining the two equations,

$$k x^2 + k y^2 = \left( -\frac{1}{c} (ax + by + d) \right)^2$$ \hspace{1cm} (317)

$$kc^2 x^2 + kc^2 y^2 = a^2 x^2 + 2abxy + b^2 y^2 + 2adx + 2bdy + d^2$$ \hspace{1cm} (318)

$$0 = (a^2 - kc^2) x^2 + (b^2 - kc^2) y^2 + (2ab) xy + (2ad) x + (2bd) y + d^2$$ \hspace{1cm} (319)

Renaming the constants (and noticing that we can give any term an arbitrary sign) this is just the general quadratic polynomial equation in two variables:

$$c_1 x^2 + c_2 y^2 + c_3 xy + c_4 x + c_5 y + c_6 = 0$$ \hspace{1cm} (320)

It is easy to visualize the possible curves that this describes: a line, a circle, an ellipse, a parabola, a hyperbola.

### 3.5.5 Runge-Lenz vector

The Runge-Lenz vector for the $U = -\frac{\alpha}{r}$ problem (note sign difference of $U$ from LL) is the vector

$$A = v \times M - \alpha \frac{r}{r}$$

It is easy to demonstrate that $A$ is conserved:

$$\frac{dA}{dt} = \frac{d}{dt} \left( v \times M - \alpha \frac{r}{r} \right)$$

$$= \dot{v} \times M - \alpha \frac{\dot{v}}{r} + \alpha \frac{r \cdot \dot{r}}{r}$$

We can replace $M = m r \times v$ and $\dot{r} = \frac{r}{r} \cdot v$ so we have

$$\frac{dA}{dt} = \dot{v} \times (m r \times v) - \alpha \frac{\dot{v}}{r} + \alpha \frac{r \cdot \dot{r}}{r}$$

$$= m r (\dot{v} \cdot v) - m v (\dot{v} \cdot r) - \alpha \frac{\dot{v}}{r} + \alpha \frac{r \cdot \dot{r}}{r}$$

$$= m r (\dot{v} \cdot v) - m v (\dot{v} \cdot r) - \alpha \frac{\dot{v}}{r} + \alpha \frac{r \cdot \dot{r}}{r}$$
Finally, since the equation of motion requires $\dot{v} = -\frac{\alpha r}{mr^3}$ the derivative reduces to zero:

$$\frac{dA}{dt} = mr \left( -\frac{\alpha r}{mr^3} \cdot v \right) + m v \left( \frac{\alpha r}{mr^3} \cdot r \right) - \frac{\alpha}{r} v + \frac{\alpha}{r^2} \left( \frac{r}{r} \cdot v \right)$$

$$= -r \left( \frac{\alpha}{r^3} r \cdot v \right) + v \frac{\alpha}{r} - \frac{\alpha}{r} + \frac{\alpha}{r^2} \left( \frac{r}{r} \cdot v \right)$$

$$= 0$$

The Runge-Lenz vector is therefore conserved.
4 Collisions between particles

4.1 Disintegration of particles

To begin, we look at the decay of a particle at rest into two particles. It is easiest to start in the rest frame of the initial particle and later transform to a general frame. In the center of mass frame, $K_0$, the total linear momentum of the particle vanishes and the energy is just the internal energy, $E_i$ (recall that this is the definition of internal energy):

$$E_0 = E_i$$
$$P_0 = 0$$

For the two resulting particles after the decay we have

$$E_1 = E_{1i} + \frac{(p_1)^2}{2m_1}$$
$$p_1 = m_1v_1$$
$$E_2 = E_{2i} + \frac{(p_2)^2}{2m_2}$$
$$p_2 = m_2v_2$$

Since no external forces act during the decay, these quantities are conserved and we have:

$$E_0 = E_1 + E_2$$
$$P_0 = p_1 + p_2$$

Substituting the initial value $P_0 = 0$ into the momentum equation gives $p_1 = -p_2$. Let the magnitude of $p_1$ and $p_2$ be $p_0$. Then the energy equation gives

$$E_i = E_{1i} + \frac{(p_1)^2}{2m_1} + E_{2i} + \frac{(p_2)^2}{2m_2}$$

$$= E_{1i} + \frac{p_0^2}{2m_1} + E_{2i} + \frac{p_0^2}{2m_2}$$

Rearranging,

$$\varepsilon \equiv E_i - E_{1i} - E_{2i}$$

61
= \frac{p_0^2}{2m_1} + \frac{p_0^2}{2m_2}
= \frac{1}{2} \left( \frac{1}{m_1} + \frac{1}{m_2} \right) p_0^2
= \frac{1}{2} \left( \frac{m_1 + m_2}{m_1 m_2} \right) p_0^2
= \frac{p_0^2}{2m}

where \( m \) is the reduced mass.

This gives the magnitude of the resulting momenta. We see from \( p_1 = -p_2 \) that their directions are opposite in this frame of reference. We may assume that there is no preferred direction (isotropy of space!) so the particular direction of the decay is equally likely to be any direction.

Now let’s find the angle of the decay in a different frame of reference. We’ll call it the laboratory frame because often we are interested in scattering problems in which one particle is accelerated to high velocity then made to collide with a stationary target. (However, the largest colliders accelerate both particles, so the lab frame and the center of mass frame are the same.) Suppose that in the laboratory frame the center of mass moves with velocity \( \mathbf{V} \). Let the velocity of particle 1 be \( \mathbf{v}_0 = \frac{p_1}{m_1} \) in the center of mass frame and \( \mathbf{v} \) in the laboratory frame. Then

\[ \mathbf{v} = \mathbf{v}_0 + \mathbf{V} \]

Rearranging and squaring,

\[ \mathbf{v} - \mathbf{V} = \mathbf{v}_0 \]
\[ v^2 - 2vV \cos \theta + V^2 = v_0^2 \]

where \( \theta \) is the angle between the relative motion of the two frames, \( \mathbf{V} \), and the velocity \( \mathbf{v} \) of the particle in the lab frame. Let the angle between \( \mathbf{V} \) and \( \mathbf{v}_0 \) be \( \theta_0 \). If we draw the three vectors (see Fig. 14 in LL), we immediately have the relations

\[ v \sin \theta = v_0 \sin \theta_0 \]
\[ v \cos \theta = V + v_0 \cos \theta_0 \]
These give the relationship between $\theta$ and $\theta_0$. Bringing $V$ to the left in the second equation and taking the ratio,

$$\tan \theta_0 = \frac{\sin \theta_0}{\cos \theta_0} = \frac{v \sin \theta}{v \cos \theta - V} = \frac{\sin \theta}{\cos \theta - \alpha}$$

where we have set $\alpha = \frac{V}{v_0}$. Finally,

$$\theta_0 = \tan^{-1} \left( \frac{\sin \theta}{\cos \theta - \alpha} \right)$$

In a scattering experiment we are interested in the distribution of quantities we measure. For example, suppose we want to find the distribution of particle energies in the lab frame for the decay above. We assume that in the center of mass frame, the distribution is uniform – every time we measure particle 1 it will have the same energy, $E_1 = E_{1i} + \frac{p_0^2}{2m_1}$, and particle 1 will come from the decay isotropically. The fraction of particles entering a given solid angle,

$$d^2\Omega = \sin \theta_0 \, d\theta_0 \, d\varphi$$

is just the fraction of the total solid angle that $d^2\Omega$ comprises. The total solid angle of a sphere is $4\pi$, so the fraction is

$$\frac{d^2\Omega}{4\pi}$$

Since there is no dependence of the distribution on $\varphi$, but only on $\theta_0$, we can integrate over $d\varphi$ to get $2\pi$. Then the probability that a scatter will enter a band between $\theta_0$ and $\theta_0 + d\theta_0$ in the center of mass system is

$$f = \frac{d^2\Omega}{4\pi} = \frac{2\pi \sin \theta_0 \, d\theta_0}{4\pi} = \frac{1}{2} \sin \theta_0 \, d\theta_0$$

Now let’s transform to the laboratory frame and find the distribution of kinetic energies. We have $T_1 = \frac{1}{2}m_1v^2$, and the velocity $v$ is related to the center of mass frame by

$$v = v_0 + V$$

63
Squaring,
\[ v^2 = v_0^2 + 2v_0V \cos \theta_0 + V^2 \]  
so the kinetic energy is
\[ T_1 = \frac{1}{2} m_1 v^2 \]
\[ = \frac{1}{2} m_1 \left( v_0^2 + 2v_0V \cos \theta_0 + V^2 \right) \]

Now differentiate to find the dependence on \( d\theta_0 \) :
\[ dT = \frac{1}{2} m_1 d \left( v_0^2 + 2v_0V \cos \theta_0 + V^2 \right) \]
\[ = \frac{1}{2} m_1 (-2v_0V \sin \theta_0 d\theta_0) \]
\[ = -m_1 v_0V \sin \theta_0 d\theta_0 \]

Notice that the minus sign shows that energy decreases with increasing angle. The fraction of particles of type 1 with energy between \( T \) and \( T + dT \) is therefore
\[ f = \frac{dT}{2m_1 v_0V} \]

Since the coefficient of \( dT \) is independent of \( T \), the distribution is constant over the full range of energies. Not all energies are allowed, however. Since \( v_0 \) and \( V \) are both fixed, \( T_1 \) has a range:
\[ T_1 = \frac{1}{2} m_1 \left( v_0^2 + 2v_0V \cos \theta_0 + V^2 \right) \]
\[ T_{1\text{min}} = \frac{1}{2} m_1 \left( v_0^2 - 2v_0V + V^2 \right) = \frac{1}{2} m_1 (v_0 - V)^2 \]
\[ T_{1\text{max}} = \frac{1}{2} m_1 \left( v_0^2 + 2v_0V + V^2 \right) = \frac{1}{2} m_1 (v_0 + V)^2 \]

Any energy in this range is equally likely. Notice that this works out right. The size of the energy range is
\[ T_{1\text{max}} - T_{1\text{min}} = \frac{1}{2} m_1 (v_0 + V)^2 - \frac{1}{2} m_1 (v_0 - V)^2 \]
\[ = 2m_1 v_0V \]
4.2 Elastic collisions

4.2.1 Center of mass frame

Next we consider the collision of two particles which fly off without any exchange of energy – an elastic collision. Start again in the center of mass system. Let the initial energy and momentum of the two particles in the center or mass frame be

\[
\begin{align*}
E_1 &= E_{1i} + \frac{(p_1)^2}{2m_1} \\
p_1 &= m_1v_1 \\
E_2 &= E_{2i} + \frac{(p_2)^2}{2m_2} \\
p_2 &= m_2v_2
\end{align*}
\]

and require the total momentum to vanish:

\[
0 = P = p_1 + p_2
\]

The same must be true after the collision. If the final state quantities are indicated by a prime (e.g., \(v'\)), then

\[
p_2' = -p_1'
\]

The fact that no energy is exchanged in the collision means that there is no change in internal energy or total energy, so the kinetic energy must stay the same too. Thus,

\[
E_{1i} + \frac{(p_1)^2}{2m_1} = E_{1i} + \frac{(p_1')^2}{2m_1}
\]

and

\[
p_1 = p_1'
\]
and similarly \( p_2 = p'_2 \); the magnitudes of the momenta are unchanged.

We can also express these relationships in terms of the velocities. The direction of the initial velocities in the center of mass system is the same as the direction of the relative velocity vector,

\[
v = v_1 - v_2 = v\hat{n}
\]

and we can express both initial particle velocities in terms of it. From our first analysis of center of mass we have

\[
v_1 = \frac{m_2}{m_1 + m_2} v\hat{n}
\]

\[
v_2 = -\frac{m_1}{m_1 + m_2} v\hat{n}
\]

Now let’s write expressions for the final velocities. Notice first that since the magnitudes of the momenta are unchanged (as follows from the elasticity), the magnitudes of the velocities are unchanged, \( v_1 = v'_1 \) and \( v_2 = v'_2 \). The only change is from the initial direction, \( \hat{n} \), to the final direction, \( \hat{n}' \). We also have

\[
v = v_1 \hat{n} - v_2 \hat{n} = v\hat{n}
\]

\[
v' = v'_1 \hat{n}' - v'_2 \hat{n}'
\]

\[
= v_1 \hat{n}' - v_2 \hat{n}' = v\hat{n}'
\]

that is, \( v' = v \). Now we can write the final velocities as:

\[
v'_1 = \frac{m_2}{m_1 + m_2} v\hat{n}'
\]

\[
v'_2 = -\frac{m_1}{m_1 + m_2} v\hat{n}'
\]

### 4.2.2 Laboratory frame

Now, if we want these formulas in a different frame of reference, the laboratory frame, then we just need to add the velocity \( \mathbf{V} \) of the center of mass as measured in the lab frame. Let \( \mathbf{u}_1 \) and \( \mathbf{u}_2 \) be the velocities in the lab frame. Then since \( \mathbf{u}_1 = v_1 + \mathbf{V} \) we have

\[
\mathbf{u}_1 = \frac{m_2 v}{m_1 + m_2} + \mathbf{V}
\]
\[ u_2 = - \frac{m_1 v}{m_1 + m_2} + \mathbf{V} \]

\[ u'_1 = \frac{m_2 v'}{m_1 + m_2} + \mathbf{V} \]

\[ u'_2 = - \frac{m_1 v'}{m_1 + m_2} + \mathbf{V} \]

From this, or the definition of the velocity of the center of mass, we have

\[ \mathbf{V} = \frac{m_1 u_1 + m_2 u_2}{m_1 + m_2} \]

### 4.2.3 A series of examples:

The most frequently occurring case is when one of the particles is at rest in the laboratory frame. We therefore consider a collision in which \( m_2 \) is initially at rest. Let \( u_1 \) be the velocity of the other particle. We want to express all other quantities in terms of \( u_1 \) and the initial and final directions. First, choose \( \mathbf{V} \) by demanding that \( m_2 \) be at rest:

\[ 0 = u_2 = - \frac{m_1}{m_1 + m_2} v + \mathbf{V} \quad (338) \]

\[ \mathbf{V} = \frac{m_1}{m_1 + m_2} v \quad (339) \]

\[ = \frac{m_1}{m_1 + m_2} (u_1 - u_2) \quad (340) \]

\[ = \frac{m_1}{m_1 + m_2} u_1 \quad (341) \]

Notice that we used \( v = v_1 - v_2 = u_1 - u_2 = u_1 \), expressing the fact that the relative velocity of the two particles is the same in any inertial frame. Therefore, in the center of mass frame before the collision, we immediately have

\[ v_1 = \frac{m_2}{m_1 + m_2} v = \frac{m_2}{m_1 + m_2} u_1 n_0 \]

\[ v_2 = - \frac{m_1}{m_1 + m_2} v = - \frac{m_1}{m_1 + m_2} u_1 n_0 \]

where the direction \( v \) (and therefore of \( u_1 \)) is \( n_0 \) and the direction of \( v' \) is \( n'_0 \). After the collision

\[ v' = v'_1 - v'_2 = u_1 n'_0 \]
\[ \mathbf{v}'_1 = \frac{m_2}{m_1 + m_2} \mathbf{v}' = \frac{m_2}{m_1 + m_2} u_1 n_0' \]
\[ \mathbf{v}'_2 = -\frac{m_1}{m_1 + m_2} u_1 n_0' \]

with the magnitudes \( v_1 = v'_1 \) and \( v_2 = v'_2 \) unchanged.

Now transform the final velocities back to the lab frame. In the lab frame,

\[ u'_1 = \frac{m_2 \mathbf{v}'}{m_1 + m_2} + \mathbf{V} \]
\[ = \frac{m_2}{m_1 + m_2} u_1 n'_0 + \frac{m_1}{m_1 + m_2} u_1 n_0 \]
\[ = \frac{u_1}{m_1 + m_2} (m_2 n'_0 + m_1 n_0) \]

\[ u'_2 = -\frac{m_1 \mathbf{v}'}{m_1 + m_2} + \mathbf{V} \]
\[ = -\frac{m_1}{m_1 + m_2} u_1 n'_0 + \frac{m_1}{m_1 + m_2} u_1 n_0 \]
\[ = \frac{m_1 u_1}{m_1 + m_2} (-n'_0 + n_0) \]

Collecting these results for the lab frame:

\[ \mathbf{u}_1 = u_1 n_0 \]
\[ \mathbf{u}_2 = 0 \]
\[ \mathbf{u}'_1 = \frac{u_1}{m_1 + m_2} (m_2 n'_0 + m_1 n_0) \]
\[ \mathbf{u}'_2 = \frac{m_1 u_1}{m_1 + m_2} (-n'_0 + n_0) \]

where the unit vectors \( n_0 \) and \( n'_0 \) are defined in the center of mass frame.

We can use these results to compute anything we want to know in either frame.

**Example 1:** Suppose we wish to compute the angle between the final directions of the particles in the lab frame. This will just be

\[ \cos \psi = \frac{\mathbf{u}'_1 \cdot \mathbf{u}'_2}{u_1 u_2} \]
\[ = \hat{\mathbf{u}}'_1 \cdot \hat{\mathbf{u}}'_2 \]
\begin{equation}
\begin{aligned}
&= \frac{(m_2 \mathbf{n}_0' + m_1 \mathbf{n}_0) \cdot (-\mathbf{n}_0' + \mathbf{n}_0)}{|m_2 \mathbf{n}_0' + m_1 \mathbf{n}_0| \cdot |\mathbf{n}_0' + \mathbf{n}_0|} \\
&= \frac{(m_2 \mathbf{n}_0' + m_1 \mathbf{n}_0) \cdot (-\mathbf{n}_0' + \mathbf{n}_0)}{\sqrt{m_1^2 + m_2^2 + 2m_1m_2 \cos \theta_0} \sqrt{2 - 2 \cos \theta_0}} \\
&= \frac{(m_1 - m_2) (1 - \cos \theta_0)}{\sqrt{m_1^2 + m_2^2 + 2m_1m_2 \cos \theta_0 \sqrt{2 - 2 \cos \theta_0}}} \\
&= (m_1 - m_2) \left( \frac{1 - \cos \theta_0}{2 (m_1^2 + m_2^2 + 2m_1m_2 \cos \theta_0)} \right)^{1/2}
\end{aligned}
\end{equation}

where we have used \( \mathbf{n}_0 \cdot \mathbf{n}_0' = \cos \theta_0 \). Notice that when \( m_1 = m_2, \cos \psi = 0 \), so the angle between the outgoing particles is always 90°.

**Example 2:** Let’s figure out when \( m_1 \) can scatter backwards. The direction of \( m_1 \) after the collision is
\begin{equation}
u_1' = \frac{u_1}{m_1 + m_2} (m_2 \mathbf{n}_0' + m_1 \mathbf{n}_0)
\end{equation}
and if the scattering is in the same direction as the initial motion, \( \mathbf{n}_0' = -\mathbf{n}_0 \). Then we have
\begin{equation}
u_1' = \frac{u_1}{m_1 + m_2} (m_1 - m_2) \mathbf{n}_0
\end{equation}
In all cases this is along the line of \( \mathbf{n}_0 \), but it is negative only if \( m_2 > m_1 \). Of course, if \( m_1 \) is greater its momentum will overwhelm \( m_2 \) and both particles will continue to move forward.

**Example 3:** How much energy can \( m_2 \) acquire? We have
\begin{equation}
E_2' = \frac{1}{2} m_2 (u_2')^2
\end{equation}
\begin{equation}
= \frac{1}{2} m_2 \left( \frac{m_1 u_1}{m_1 + m_2} \right)^2 (2 - 2 \cos \theta_0)
\end{equation}
\begin{equation}
= m_1 u_1^2 \frac{m_1 m_2}{(m_1 + m_2)^2} (1 - \cos \theta_0)
\end{equation}
\begin{equation}
= \frac{2m_1 m_2}{(m_1 + m_2)^2} E_1 (1 - \cos \theta_0)
\end{equation}
This is maximum when \( \cos \theta_0 = -1 \), giving
\begin{equation}
E_{2\text{max}}' = \frac{4m_1 m_2}{(m_1 + m_2)^2} E_1
\end{equation}

69
This expression, as a function of $m_1$, starts at zero when $m_1$ is zero, rises to a single maximum, then decreases back to zero as $m_1$ tends to infinity. The maximum is given by setting

$$0 = \frac{d}{dm_1} \frac{4m_1m_2}{(m_1 + m_2)^2} = \frac{4m_2}{(m_1 + m_2)^2} - \frac{8m_1m_2}{(m_1 + m_2)^3}$$

or

$$4m_2(m_1 + m_2) - 8m_1m_2 = 0$$
$$m_2 - m_1 = 0$$

so the greatest energy is transferred when the masses are equal. Setting $m_1 = m_2$, we find that all of the energy of $m_1$ can be transferred:

$$E'_{2_{\text{max}}} = E_1$$

When the masses are very different, the transfer is not large. We consider the two extremes. When $m_1 << m_2$ we find

$$E'_{2_{\text{max}}} \approx \frac{4m_1}{m_2}E_1 << E_1$$

while $m_2 << m_1$ leads to

$$E'_{2_{\text{max}}} = \frac{4m_2}{m_1}E_1 << E_1$$

**Example 4:** There are other questions that are not hard to answer. For example, suppose $m_1$ flies off making an angle $\theta_0$ with its initial direction where, as stated above,

$$\cos \theta_0 = n_0 \cdot n'_0$$

What is the angle does $m_1$ make in the lab frame? To answer the question we need only find the angle between the $n_0$ and $u'_0$. Thus,

$$\cos \theta_1 = n_0 \cdot \frac{u'_0}{|u'_0|}$$
Example 5: Finally, there is an easy way to compute the angle, $\theta_2$, between $n_0$ and $u'_2$. First, notice that since

$$u'_2 = \frac{m_1 u_1}{m_1 + m_2} (-n'_0 + n_0) \quad (357)$$

the angle we want is the angle between $n_0 - n'_0$ and $n_0$. Now, the three vectors, $n_0$, $-n'_0$ and $n_0 - n'_0$ form an isosceles triangle because $n_0$ and $-n'_0$ both have length one. Therefore, since the angle between $n_0$ and $n_0 - n'_0$ is $\theta_2$, the angle between $-n'_0$ and $n_0 - n'_0$ is also $\theta_2$. The third angle of the triangle, the angle between $n_0$ and $-n'_0$, is just $\theta_0$. So we immediately have

$$2\theta_2 + \theta_0 = \pi \quad (358)$$

or

$$\theta_2 = \frac{1}{2} (\pi - \theta_0) \quad (359)$$

We may need this result later.

We could also have gotten this result by computing directly. First find a unit vector in the direction of $u'_2$:

$$\hat{m} = \frac{u'_2}{w'_2} = \frac{n_0 - n'_0}{|n_0 - n'_0 + n_0|} \quad (360)$$

$$= \frac{n_0 - n'_0}{\sqrt{2(1 - \cos \theta_0)}} \quad (361)$$

and take the dot product with $n_0$:

$$\cos \theta_2 = n_0 \cdot \hat{m} \quad (362)$$

$$= \frac{n_0 \cdot (n_0 - n'_0)}{\sqrt{2(1 - \cos \theta_0)}} \quad (363)$$

$$= \sqrt{\frac{1 - \cos \theta_0}{2}} \quad (364)$$

$$= \sqrt{\frac{1 + \cos (\pi - \theta_0)}{2}} \quad (365)$$

$$= \cos \frac{1}{2} (\pi - \theta_0) \quad (366)$$

so that once again we find $\theta_2 = \frac{1}{2} (\pi - \theta_0)$.  

71
4.3 Scattering

Our goal in analyzing scattering is to relate our solution of two particle motion in a potential to quantities that are readily measured in the laboratory. Typically, we measure the number of particles scattered through any given angle in the laboratory frame. This number of particles normalized by the incident flux is called the differential cross section.

Let \( dN \), the quantity we measure, be the number of particles per unit time that deflect through an angle between \( \chi \) and \( \chi + d\chi \). Now, imagine a particle beam with a uniform distribution of particles moving with constant velocity in a particular direction. Suppose that \( n \) particles per unit time cross a unit cross-sectional area of this beam (this is just the density of particles in the beam divided by the velocity of the beam, and is called the flux of the beam). The cross section is just the ratio,

\[
d\sigma = \frac{dN}{n}
\]

that is, the number of particles deflected through an angle between \( \chi \) and \( \chi + d\chi \) per unit time, per unit incident flux. Notice that \( \sigma \) has units of area. Typically, values of \( \sigma \) are given in units of barns (as in, “You can’t hit the broad side of a barn.”) with one barn equal to \( 10^{-24} \text{ cm}^2 \). This is about midway between the cross-sectional size of an atom and the cross-sectional size of a nucleus.

Picture a beam of particles hitting a fixed target. We assume that particles in the beam are spread out uniformly, and when the beam is far from the target all of the particles have the same direction and velocity. This means that while a particle in the center of the beam will strike the target particle (and bounce straight back), other particles will miss and fly off at various angles. What makes different particles different is their offset from the center line. This offset, called the impact parameter, determines that particle’s angular momentum. Suppose the particle is very far from the target, moving with velocity \( v_\infty \), and on a line that, if continued, would pass a distance \( \rho \) from the target. Then since

\[
\mathbf{M} = \mathbf{r} \times \mathbf{p}
\]

we have

\[
J = m\rho v_\infty
\]
with the direction of \( \mathbf{M} \) out of the plane of the scattering.

It is now easy to express \( dN \) in terms of our solution to the scattering problem. Let the angle around the line of incident particles be the azimuthal angle \( \phi \) (that is, if we think of \( \chi \) as the angle \( \theta \) of spherical coordinates and the beam line as the \( z \) axis, then \( \phi \) is the usual angle in the \( xy \) plane.) From the potential, can find the dependence of \( \rho \) on \( \chi \) and if we assume that the relationship is one-to-one we can write the number of particles passing between \( \chi \) and \( \chi + d\chi \) in terms of the number passing through the area element between \( \rho \) and \( \rho + d\rho \) and between the annular angle \( \phi \) and \( \phi + d\phi \). We just have

\[
dN = n\rho \, d\phi \, d\rho
\]

(370)

where \( \rho \, d\phi \, d\rho \) is the area of the part of the annulus just described. The cross-section is therefore

\[
d\sigma = \rho \, d\phi \, d\rho
\]

(371)

and using the dependence \( \rho(\chi) \) this becomes

\[
d\sigma = \left| \rho(\chi) \frac{d\rho}{d\chi} \, d\phi \, d\chi \right|
\]

(372)

\[
= \rho(\chi) \left| \frac{d\rho}{d\chi} \right| \, d\phi \, d\chi
\]

(373)

We need the absolute value sign because we want to keep the cross-section positive. However, increasing the impact parameter generally means a the particle will scatter less, so \( \rho \) is a decreasing function of \( \chi \), \( \frac{d\rho}{d\chi} < 0 \).

Finally, we consider the effect of the interaction that causes the scattering of the two particles. This is what determines \( \theta_0 \) in terms of the physical variables. Usually we are interested in scattering from a central potential, \( U(r) \), and this is the case we consider here. We have already solved this problem completely. The motion lies in a plane, and we can express the angle of the particle as a function of the distance between the center of force by the integral

\[
\varphi = \int_{r_0}^{r(\varphi)} \frac{Jdr}{r^2 \sqrt{2m(E-U)-\frac{J^2}{r^2}}}
\]

(374)

where \( \varphi = 0 \) when \( r = r_0 \), and \( J \) is the magnitude of the total angular momentum. Since we can rotate our coordinates any way we like, we can
set the zero of \( \varphi \) wherever we want. Recall that for unbounded motion, there is only one turning point of the motion, \( r_{\text{min}} \). The particle comes in from a great distance, reaches this minimum, then returns outward. It is therefore convenient to take \( \varphi = 0 \) when \( r = r_{\text{min}} \), so that

\[
\varphi = \int_{r_{\text{min}}}^{r_{\text{max}}} \frac{Jdr}{r^2 \sqrt{2m(E - U) - \frac{r^2}{r^2}}} \tag{375}
\]

and the only difference between the approach to \( r_{\text{min}} \) and the retreat back to \( r = \infty \) is the sign of \( \varphi \). The functional dependence \( \varphi(r) \) is the same for both halves of the motion. Notice that once we are given the potential, the full motion depends on only the two constants of the motion, \( J \) and \( E \).

The situation is depicted in Figure 18 on page 48. As the particle approaches (from the right in the diagram), it is deflected away from the original line of motion so it doesn’t ever actually come within a distance \( \rho \) of the target. Instead, the closest approach is \( r_{\text{min}} \). Our equation for \( \varphi \) can now be rewritten in terms of the impact parameter \( \rho \) and the initial velocity \( v_\infty \). With \( J = m\rho v_\infty \) and \( E = \frac{1}{2}mv_\infty^2 \) we have

\[
\varphi(\rho) = \int_{r_{\text{min}}}^{r_{\text{max}}} \frac{m\rho v_\infty dr}{r^2 \sqrt{2m\left(\frac{1}{2}mv_\infty - U\right) - \frac{\rho^2}{r^2}}} \tag{376}
\]

\[
= \int_{r_{\text{min}}}^{r_{\text{max}}} \rho dr \frac{\sqrt{1 - \frac{2U}{mv_\infty^2} - \frac{\rho^2}{r^2}}}{r^2} \tag{377}
\]

The asymptotic value \( \varphi_0 \) of \( \varphi \) is found by setting the upper limit to \( \infty \).

Now look at the figure. The angle \( \varphi \) is the angle between the line of closest approach, \( OA \), and the position vector from \( O \) to the particle. Let the angle \( \theta_0 \) (\( \chi \) in the figure) be the angle by which the particle is deflected in the center of mass frame. This is easily related to \( \varphi \). Note that \( \theta_0 \) is measured from the extension of the particle’s original line of motion through the center of mass, and this line makes an angle \( \varphi_0 \) with \( OA \). Because of the symmetry of the orbit, there is an angle \( \varphi_0 \) on both sides of \( OA \), so we have

\[
\theta_0 + 2\varphi_0 = \pi \tag{378}
\]

and therefore

\[
\theta_0 = \pi - 2\int_{r_{\text{min}}}^{\infty} \frac{\rho dr}{r^2 \sqrt{1 - \frac{2U}{mv_\infty^2} - \frac{\rho^2}{r^2}}} \tag{379}
\]
This gives the center of mass angle of deflection $\theta_0$ as a function of impact parameter $\rho$. The inverse of this function is exactly what we need. Given the potential, we can solve for $\rho(\theta_0)$. After transforming $\theta_0$ to the laboratory frame, we have $\rho(\chi)$ and can compute the cross section.

Since we are dealing with a central potential, there is no dependence of the scattering on $\varphi$. Therefore, $d\sigma$ only depends on $d\chi$ and we can integrate over all $\varphi$ to get

$$d\sigma = 2\pi \rho(\theta_0) \left| \frac{d\rho}{d\theta_0} \right| d\theta_0$$

(380)

Finally, we put the result in terms of the solid angle, $d\Omega = 2\pi \sin \theta_0 d\theta_0$,

$$d\sigma = \frac{\rho(\theta_0)}{\sin \theta_0} \left| \frac{d\rho}{d\theta_0} \right| d\Omega$$

(381)

This gives the differential cross section for a central potential in terms of the solid angle in the center of mass system. To find the cross section in the laboratory frame, we can transform from $\theta_0$ to $\chi$ according to the results of the previous section.

4.4 Rutherford scattering

Perhaps the most important scattering potential is the Coulomb potential, $U = \frac{q_1 q_2}{r}$. In a situation such as Rutherford’s original experiment, the scattered particle is much lighter than the target, and it is a good approximation to assume that the laboratory and center of mass frames are the same. In this case, we can integrate to find $\chi$ directly. We have

$$\theta_0 = \chi = \pi - 2 \int_{r_{\text{min}}}^{\infty} \frac{\rho dr}{\sqrt{1 - \frac{2\alpha}{mv_{\infty}^2 r} - \frac{\rho^2}{r^2}}}$$

(382)

where $r_{\text{min}}$ is the root of $1 - \frac{2\alpha}{mv_{\infty}^2 r} - \frac{\rho^2}{r^2} = 0$. Letting $u = \rho/r$ we find

$$\chi = \pi + 2 \int_{u_{\text{max}}}^{0} \frac{du}{\sqrt{1 - \frac{2\alpha}{mv_{\infty}^2} u - u^2}}$$

(383)

where $u_{\text{max}}$ is the root of $1 - \frac{2\alpha}{mv_{\infty}^2} u - u^2 = 0$. Rewrite

$$1 - \frac{2\alpha}{mv_{\infty}^2} u - u^2 = -\left(u + \frac{\alpha}{mv_{\infty}^2 \rho}\right)^2 + 1 + \frac{\alpha^2}{m^2 v_{\infty}^4 \rho^2}$$

(384)

$$= -y^2 + \beta^2$$

(385)
where we have set $\beta^2 = 1 + \frac{\alpha^2}{m^2 v_\infty^2 \rho^2}$ and $y = u + \frac{\alpha}{m v_\infty \rho}$. Then

$$\chi = \pi + 2 \int_\beta^\infty \frac{dy}{\sqrt{\beta^2 - y^2}}$$

We can integrate if we now set $y = \beta \sin \psi$ :

$$\chi = \pi - 2 \int_{\pi/2}^{\psi_0} \frac{\beta \cos \psi d\psi}{\sqrt{\beta^2 (1 - \sin^2 \psi)}}$$

$$= \pi + 2\psi_0 - \pi$$

$$= 2 \sin^{-1} \left( \frac{\alpha}{\sqrt{1 + \frac{\alpha^2}{m^2 v_\infty^2 \rho^2}}} \right)$$

Simplifying,

$$\sin \frac{\chi}{2} = \frac{\alpha}{\sqrt{m^2 v_\infty^4 \rho^2 + \alpha^2}}$$

$$m^2 v_\infty^4 \rho^2 + \alpha^2 = \frac{\alpha^2}{\sin^2 \frac{\chi}{2}}$$

$$\rho = \frac{\alpha}{m v_\infty^2} \sqrt{\frac{1}{\sin^2 \frac{\chi}{2}} - 1}$$

$$= \frac{\alpha}{m v_\infty^2} \cot \frac{\chi}{2}$$

Now differentiating,

$$\frac{d\rho}{d\chi} = \frac{\alpha}{2 m v_\infty^2 \sin^2 \frac{\chi}{2}}$$

and substituting,

$$d\sigma = \frac{\rho(\chi)}{\sin \chi} \left| \frac{d\rho}{d\chi} \right| d\Omega$$

$$= \frac{1}{\sin \chi} \left( \frac{\alpha}{mv_\infty^2} \right)^2 \frac{\cot \frac{\chi}{2}}{2 \sin^2 \frac{\chi}{2}} d\Omega$$

$$= \left( \frac{1}{mv_\infty^2} \right)^2 \frac{\alpha^2}{4 \sin^4 \frac{\chi}{2}} d\Omega$$

$$= \frac{\alpha^2}{16 E_\infty^2 \sin^4 \frac{\chi}{2}} d\Omega$$
where $E_\infty$ is the initial kinetic energy of the scattered particle. Note especially that the final result, the *Rutherford cross section*,

$$d\sigma_R = \frac{\alpha^2}{16E_\infty^2 \sin^4 \frac{\theta}{2}} d\Omega$$  \hspace{1cm} (398)

is characterized by the *inverse fourth power* of $\sin \frac{\theta}{2}$.

4.5 Small angle scattering

For scattering through small angles, we can make some simplifying approximations in calculating the cross section. We assume that the impact parameter is large so that the particle passes where $U$ is weak. Let the scattering occur in the $xy$ plane, with the incident beam moving in the $x$ direction. Then the scattering angle $\theta_1$ of $m_1$ is small and we can make the approximations

$$\sin \theta_1 \approx \theta_1 \approx \frac{p_{1y}'}{p_1'} \quad (399)$$

$$p_1' \approx m_1 v_\infty \quad (400)$$

Now, using Newton’s 2nd law,

$$p_{1y}' \approx \int_{-\infty}^{\infty} F_y dt \quad (401)$$

$$= -\int_{-\infty}^{\infty} \frac{\partial U}{\partial y} dt \quad (402)$$

$$= -\int_{-\infty}^{\infty} \frac{dU}{dr} \frac{\partial r}{\partial y} dt \quad (403)$$

$$= -\int_{-\infty}^{\infty} \frac{dU}{dr} \frac{y}{r} dt \quad (404)$$

Now (since $U$ is already assumed small), we can write $y = \rho +$ small stuff, and drop the small stuff. Similarly, the velocity is $v = v_\infty +$ small stuff, so we can set $dt \approx dx/v_\infty$ and write

$$p_{1y}' \approx -\frac{\rho}{v_\infty} \int_{-\infty}^{\infty} \frac{dU}{dr} \frac{dx}{r} \quad (405)$$
Finally, change to an integration over $r$:

$$r^2 = x^2 + \rho^2 \quad (406)$$
$$dx = \frac{r}{x} dr \quad (407)$$
$$= \frac{r}{\sqrt{r^2 - \rho^2}} dr \quad (408)$$

so that

$$p_{1y}' \approx -\frac{\rho}{v_\infty} \int_{-\infty}^{\infty} \frac{dU}{dr} \frac{dr}{\sqrt{r^2 - \rho^2}} \quad (409)$$

The scattering angle is therefore given by

$$\theta = \frac{p_{1y}'}{m_1 v_\infty} \quad (410)$$
$$= -\frac{\rho}{m_1 v_\infty^2} \int_{-\infty}^{\infty} \frac{dU}{dr} \frac{dr}{\sqrt{r^2 - \rho^2}} \quad (411)$$
$$= -\frac{2\rho}{m_1 v_\infty^2} \int_{\rho}^{\infty} \frac{dU}{dr} \frac{dr}{\sqrt{r^2 - \rho^2}} \quad (412)$$
5 Small Oscillations

5.1 Small oscillations in one dimension

Almost any system whatsoever has motions comprised of small deviations from equilibrium. These can be treated quite generally. Let’s begin with any system with one degree of freedom, and consider motion near equilibrium. We can expand the potential in a Taylor series about the equilibrium point, $q_0$

$$U(q) = U_0 + \frac{dU}{dq}\bigg|_{q_0} (q - q_0) + \frac{1}{2} \frac{d^2U}{dq^2}\bigg|_{q_0} (q - q_0)^2 + \ldots$$  \hfill (413)

We can simplify this considerably. First, because $q_0$ is a point of equilibrium, it must be an extremum of the potential, so $\frac{dU}{dq}\bigg|_{q_0} = 0$. Furthermore, we may choose a new coordinate, $x = q - q_0$, that vanishes at the equilibrium point. Finally, since we can always change the energy scale by a constant, we can set $U_0 = 0$. This reduces the potential to

$$U(x) = \frac{1}{2} \frac{d^2U}{dx^2}\bigg|_{x_0} x^2 + \ldots$$  \hfill (414)

$$\approx \frac{1}{2} k x^2$$  \hfill (415)

where we have defined $k = \frac{d^2U}{dq^2}\bigg|_{q_0}$. Therefore, near an equilibrium point, any potential has the potential of a simple harmonic oscillator. Notice that $x = 0$ is a minimum of the potential if and only if $k > 0$. If $k < 0$ then $x = 0$ is a maximum. For $U$ maximum, we show below that the equilibrium point is not a stable one.

We can also simplify the kinetic energy. In general, we have

$$T = \frac{1}{2} a(q) \dot{q}^2$$  \hfill (416)

and we can expand $a(q)$ about $q_0$ as well:

$$a(q) = a(q_0) + \frac{da}{dq}\bigg|_{q_0} (q - q_0) + \frac{1}{2} \frac{d^2a}{dq^2}\bigg|_{q_0} (q - q_0)^2 + \ldots$$  \hfill (417)

$$= a(q_0) + \frac{da}{dq}\bigg|_{q_0} x + \frac{1}{2} \frac{d^2a}{dq^2}\bigg|_{q_0} x^2 + \ldots$$  \hfill (418)
However, we also have
\[ \dot{q} = \frac{d}{dt}(x + q_0) = \dot{x} \quad \text{(419)} \]
Since we have kept only terms quadratic in \( x \) in the potential, we must keep only to the same order in the kinetic energy, so we have
\[
T = \frac{1}{2} a(q)\dot{x}^2 \quad \text{(420)}
\]
\[
= \frac{1}{2} \left( a(q_0) + \frac{da}{dq}q_0 + \frac{1}{2} \frac{d^2a}{dq^2}q_0^2 + \ldots \right) \dot{x}^2 \quad \text{(421)}
\]
\[
\approx \frac{1}{2} a(q_0)\dot{x}^2 \quad \text{(422)}
\]
Let’s make this look familiar by defining \( m = a(q_0) \). Although this isn’t necessarily the mass, it acts in much the same way.
Now the Lagrangian is
\[
L = \frac{1}{2} m\dot{x}^2 - \frac{1}{2} kx^2 \quad \text{(423)}
\]
so the Euler-Lagrange equation is
\[
0 = \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} \quad \text{(424)}
\]
\[
= m\ddot{x} - kx \quad \text{(425)}
\]
Rewrite this as
\[ 0 = \ddot{x} + \omega^2 x \quad \text{(426)} \]
where \( \omega = \sqrt{\frac{k}{m}} \). Notice that since the kinetic energy is always positive, \( m \) is positive. Therefore, \( \omega \) is real if \( k \) is positive. For maxima of the potential, \( k \) is negative and \( \omega \) is pure imaginary.
The solution to the equation is
\[
x = c_1 \cos \omega t + c_2 \sin \omega t \quad \text{(427)}
\]
or equivalently,
\[
x = a \cos (\omega t + \alpha) \quad \text{(428)}
\]
In the first case, the initial conditions \( x_0 = x(0) \) and \( v_0 = \dot{x}(0) \) lead to

\[
x_0 = c_1 \cos 0 + c_2 \sin 0
= c_1
\]
\[
v_0 = -\omega c_1 \sin 0 + \omega c_2 \cos 0
= \omega c_2
\]

so that

\[
x = x_0 \cos \omega t + \frac{v_0}{\omega} \sin \omega t
\]  

(433)

For the second form we have

\[
x_0 = a \cos \alpha
v_0 = -\omega a \sin \alpha
\]

(434) (435)

so that

\[
a = \sqrt{x_0^2 + \left(\frac{v_0}{\omega}\right)^2}
\]
\[
\tan \alpha = -\frac{v_0}{\omega x_0}
\]

(436) (437)

The coefficient \( a \) is called the amplitude, and \( \omega \) is called the (angular) frequency. The quantity \((\omega t + \alpha)\) is called the phase, and \( \alpha \) is the initial value of the phase.

A third way to write \( x \) is using a complex exponential. Since \( e^{ix} = \cos x + i \sin x \) we have

\[
x = \text{Re} \left\{ ce^{i(\omega t + \alpha)} \right\}
= \text{Re} \left\{ ce^{i\omega t} e^{i\alpha} \right\}
= \text{Re} \left\{ ce^{i\alpha} e^{i\omega t} \right\}
\]

(438) (439) (440)

The coefficient, \( A = ce^{i\alpha} \), is now an arbitrary complex number, so we can write

\[
x = \text{Re} \left\{ Ae^{i\omega t} \right\}
\]

(441)

Any linear operation on \( x \) commutes with taking the real part, so for many calculations we can work with \( \{ Ae^{i\omega t} \} \) and take the real part afterward.
For example, since differentiation is a linear operation, we have:

\[ v = \frac{dx}{dt} = \frac{d}{dt} \text{Re} \{ A e^{i\omega t} \} = \frac{1}{2} \frac{d}{dt} \{ A e^{i\omega t} + A^* e^{-i\omega t} \} = \frac{1}{2} \left\{ i\omega A e^{i\omega t} - i\omega A^* e^{-i\omega t} \right\} = \frac{1}{2} \left\{ i\omega A e^{i\omega t} \right\} + \frac{1}{2} \left\{ -i\omega A^* e^{-i\omega t} \right\} = \text{Re} \left\{ i\omega A e^{i\omega t} \right\} \]

so we can take the time derivative of \( A e^{i\omega t} \) directly, and take the real part later.

### 5.2 The Forced Simple Harmonic Oscillator

Next, we consider what happens if we have a simple harmonic oscillator subject to a time-dependent force. For this section we will consider a single particle in a pure quadratic potential, rather than a perturbative solution. We take the force to be an arbitrary function of time. Such a force is generally external, but we can include a time-dependent potential for it nonetheless. Let

\[ F(t) = -\frac{\partial}{\partial x} U(x, t) \]

and integrate to find

\[ U(x, t) = -xF(t) \]

The Lagrangian becomes

\[ L = \frac{1}{2} m \dot{x}^2 - \frac{1}{2} k x^2 + xF(t) \]

and the equation of motion is

\[ \ddot{x} + \omega^2 x = \frac{1}{m} F(t) \]
Notice that, because $L$ depends explicitly on $t$, we are no longer guaranteed to have conservation of energy.

To begin with, let’s look at a simple example. Let $F(t)$ be harmonic,

\[
\frac{1}{m} F(t) = f \cos(\gamma t + \beta)
\]

Then we must solve

\[
\ddot{x} + \omega^2 x = f \cos(\gamma t + \beta)
\]

This simple example is more useful than it may appear, because a general force may be expanded in a Fourier series. Since the equation is linear, we can solve for each Fourier mode separately.

Recall that the general solution to an inhomogeneous linear differential equation is the sum of a general solution to the homogeneous equation, $x_h(t)$ and any particular solution, $x_p(t)$, to the full equation. The homogeneous solution gives us the required two initial conditions while the particular solution accommodates the force. We already know the homogeneous solution,

\[
x = a \cos(\omega t + \alpha)
\]

and we seek a particular solution that looks like the driving force. Suppose

\[
x = k \cos(\gamma t + \beta)
\]

for some constant $k$. Substitute into the equation of motion to find:

\[
-\gamma^2 k \cos(\gamma t + \beta) + \omega^2 \omega^2 k \cos(\gamma t + \beta) = f \cos(\gamma t + \beta)
\]

\[
-\gamma^2 k + \omega^2 k = f
\]

\[
k = \frac{f}{\omega^2 - \gamma^2}
\]

The general solution is therefore

\[
x = a \cos(\omega t + \alpha) + \frac{f}{\omega^2 - \gamma^2} \cos(\gamma t + \beta)
\]

Notice that, while the initial conditions are still determined by $a$ and $\alpha$, the particular solution changes their exact form. Indeed, $a$ and $\alpha$ are determined by

\[
x_0 = a \cos \alpha + \frac{f}{\omega^2 - \gamma^2} \cos \beta
\]

\[
v_0 = -\omega a \sin \alpha - \frac{f \gamma}{\omega^2 - \gamma^2} \sin \beta
\]
All other constants in these expressions are known, so we have

\[
a = \left( x_0^2 + \frac{v_0^2}{\omega^2} - 2x_0 \frac{f}{\omega^2 - \gamma^2} \cos \beta \right) + \frac{f^2 \gamma v_0}{\omega^2 (\omega^2 - \gamma^2)} \sin \beta + \frac{f^2}{(\omega^2 - \gamma^2)^2} \right)^{1/2} \tag{454}
\]

\[
\alpha = \tan^{-1} \left( \frac{f\gamma \omega \sin \beta + v_0 (\omega^2 - \gamma^2)}{f \omega \cos \beta - x_0 (\omega^2 - \gamma^2)} \right) \tag{455}
\]

This solution is satisfactory as long as the driving frequency \( \gamma \) is not too close to the natural frequency of the oscillations of the system, \( \omega \). The dramatic increase in the amplitude of oscillations near a natural frequency is called resonance. The result diverges when the driving force acts at the natural frequency of the oscillator. This is unrealistic, since physical oscillators invariably have some dissipation. We will treat friction in a later section. Even in the present case, we can find a solution when \( \gamma \) is extremely close to \( \omega \), which is valid for short times. To do this, let

\[
\gamma = \omega + \varepsilon \tag{456}
\]

where \( \varepsilon \) is small. However, this is not a dimensionless statement. Notice that \( \varepsilon \) multiplied by the time, \( t \), is dimensionless. One option, then, is to consider the motion of the system for times such that \( \varepsilon t << 1 \). Alternatively, since \( \varepsilon/\omega \) is dimensionless, we may demand \( \varepsilon/\omega << 1 \). In fact, we will demand both.

Now let’s rewrite \( x(t) \) by adding and subtracting the same multiple of \( \cos (\omega t + \beta) \):

\[
x = a \cos (\omega t + \alpha) + \frac{f}{m (\omega^2 - \gamma^2)} \cos (\gamma t + \beta) \tag{457}
\]

\[
= a \cos (\omega t + \alpha) + \frac{f}{m (\omega^2 - \gamma^2)} \cos (\omega t + \beta) \tag{458}
\]

\[
+ \frac{f}{m (\omega^2 - \gamma^2)} (\cos (\gamma t + \beta) - \cos (\omega t + \beta)) \tag{459}
\]

Now we define new constants \( a' \) and \( \alpha' \) to absorb the middle term. Then

\[
x = a' \cos (\omega t + \alpha') + \frac{f}{m (\omega^2 - \gamma^2)} (\cos (\gamma t + \beta) - \cos (\omega t + \beta)) \tag{460}
\]
where the arguments of the cosines in the final term differ by order $\varepsilon$.

Now we expand in terms of the small quantity $\varepsilon t$. We will also assume that $\varepsilon << \omega$. We write the denominator of the coefficient as

$$\frac{1}{\omega^2 - \gamma^2} = \frac{1}{\omega^2 - (\omega + \varepsilon)^2} \tag{461}$$

$$= \frac{1}{-2\omega \varepsilon - \varepsilon^2} \tag{462}$$

$$= -\frac{1}{2\omega \varepsilon} \left(1 + \frac{\varepsilon}{2\omega} + \cdots \right) \tag{463}$$

$$= -\frac{1}{2\omega \varepsilon} \left(1 - \frac{\varepsilon}{2\omega} + \frac{\varepsilon^2}{4\omega^2} - \cdots \right) \tag{464}$$

where we have used the Taylor series for $\frac{1}{1+x}$. At the same time, we can expand the cosine term:

$$\cos(\gamma t + \beta) = \cos(\omega t + \beta + \varepsilon t) \tag{465}$$

$$= \cos(\omega t + \beta) \cos(\varepsilon t) - \sin(\omega t + \beta) \sin(\varepsilon t) \tag{466}$$

$$= \left(1 - \frac{\varepsilon^2 t^2}{2} + \cdots \right) \cos(\omega t + \beta)$$

$$- (\varepsilon t + \cdots) \sin(\omega t + \beta) \tag{467}$$

Then

$$\cos(\gamma t + \beta) - \cos(\omega t + \beta) = -\varepsilon t \sin(\omega t + \beta)$$

$$- \frac{\varepsilon^2 t^2}{2} \cos(\omega t + \beta) + \cdots \tag{469}$$

where again we keep terms up to second order in $\varepsilon t$. Combining these results we have

$$x = a' \cos(\omega t + a')$$

$$+ \frac{f}{m (\omega^2 - \gamma^2)} (\cos(\gamma t + \beta) - \cos(\omega t + \beta)) \tag{472}$$

$$= a' \cos(\omega t + a')$$

$$- \frac{f}{2m \omega \varepsilon} \left(1 - \frac{\varepsilon}{2\omega}\right) \left(-\varepsilon t \sin(\omega t + \beta) - \frac{\varepsilon^2 t^2}{2} \cos(\omega t + \beta)\right) \tag{474}$$

$$= a' \cos(\omega t + a')$$

$$- \frac{ft}{2m \omega} \left(-\sin(\omega t + \beta) + \frac{\varepsilon}{2\omega} \sin(\omega t + \beta) - \frac{\varepsilon t}{2} \cos(\omega t + \beta)\right) \tag{476}$$
so the final result is

\[ x = a' \cos(\omega t + \alpha') + \frac{f t}{2m\omega} \sin(\omega t + \beta) \]  
\[ -\frac{f \varepsilon t}{4m\omega^2} \left( \sin(\omega t + \beta) - \omega t \cos(\omega t + \beta) \right) \]  

This is valid as long as both

\[ \varepsilon t << 1 \]  

and \( \frac{ft}{2m\omega} \) remains small. This last quantity is the amplitude of the lowest order correction to the homogeneous solution, and it grows linearly with time. What it means for it to be small depends on what it means for \( x \) to stay small. This, in turn, depends on our original expansion of the potential. If we expanded a general potential as

\[ U(x) = \frac{1}{2} \frac{d^2U}{dx^2} \bigg|_{x=0} x^2 + \frac{1}{3!} \frac{d^3U}{dx^3} \bigg|_{x=0} x^3 + \ldots \]  

then we neglected the cubic term. It is consistent to ignore this third term as long as

\[ \frac{1}{2} \frac{d^2U}{dx^2} \bigg|_{x=0} x^2 >> \frac{1}{3!} \frac{d^3U}{dx^3} \bigg|_{x=0} x^3 \]  

or

\[ x << 3 \left( \frac{dU}{dx^2} \bigg|_{x=0} \right)^{-1} \left( \frac{d^3U}{dx^3} \bigg|_{x=0} \right) \]  

Therefore, the condition on the amplitude near resonance is

\[ \frac{ft}{2m\omega} << 3 \left( \frac{dU}{dx^2} \bigg|_{x=0} \right)^{-1} \left( \frac{d^3U}{dx^3} \bigg|_{x=0} \right) \]  

We need both this and \( \varepsilon t << 1 \) for our approximations to be valid.

Notice that if our system is exactly a harmonic oscillator, so that

\[ x = a \cos(\omega t + \alpha) + \frac{f}{m(\omega^2 - \gamma^2)} \cos(\gamma t + \beta) \]  

is the exact solution to the problem, we can do a little more. Let’s write \( x \) in the complex form

\[ x = Ae^{i\omega t} + Be^{i\gamma t} \]
where the complex constant $B$ contains the resonance. Replacing $\gamma = \omega + \varepsilon$,

$$x = Ae^{i\omega t} \left(1 + \frac{B}{A} e^{i\varepsilon t}\right) \quad (486)$$

Now, the term in parentheses varies slowly because $\varepsilon$ is small. It has squared magnitude

$$\left|1 + \frac{B}{A} e^{i\varepsilon t}\right|^2 = \left(1 + \frac{be^{i\beta}}{a^{e^{i\alpha}}} e^{i\varepsilon t}\right) \left(1 + \frac{be^{-i\beta}}{a^{e^{-i\alpha}}} e^{-i\varepsilon t}\right) \quad (487)$$

$$= \left(1 + \frac{2b}{a} \cos(\varepsilon t + \beta - \alpha) + \frac{b^2}{a^2}\right) \quad (488)$$

We can think of the result,

$$x(t) = \left(a^2 + 2ab \cos(\varepsilon t + \beta - \alpha) + b^2\right)^{1/2} e^{i\omega t} \quad (489)$$

as a simple oscillation, $e^{i\omega t}$ together with a slow modulation of the amplitude between the values $|a - b|$ and $a + b$.

### 5.2.1 General driving force

Now we return to the case of a general driving force, $F(t)$. We assume $F(t) = 0$ for all $t < 0$. We will solve the problem in two ways. First, we will use a Fourier transform, then we will show that the same result follows by a clever trick.

The Fourier transform of a function $f(t)$ is given by

$$g(\gamma) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{i\gamma t} dt \quad (490)$$

It is invertible in the sense that if we know $g(\gamma)$ we can recover $f(t)$ by computing

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(\gamma) e^{-i\gamma t} d\gamma = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t')e^{i\gamma t'} e^{-i\gamma t} dt'd\gamma \quad (491)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t') dt' \int_{-\infty}^{\infty} e^{i\gamma(t'-t)} d\gamma \quad (492)$$

$$= \int_{-\infty}^{\infty} f(t') \delta(t - t') dt' \quad (493)$$

$$= f(t) \quad (494)$$

87
where we have used the harmonic expansion of the Dirac delta function,
\[ \delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\gamma x} \, d\gamma \] (495)

Notice that for the integral of \( g(\gamma) \) we use the opposite sign in the exponent. The expression for \( f \) in terms of \( g \) is called the inverse Fourier transform.

To solve the differential equation
\[ \ddot{x} + \omega^2 x = \frac{1}{m} F(t) \] (496)
we first write the inverse Fourier transforms for both \( x(t) \) and \( F(t) \):
\[ x(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y(\gamma)e^{-i\gamma t} \, d\gamma \] (497)
\[ F(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} G(\gamma)e^{-i\gamma t} \, d\gamma \] (498)

then substitute these into the differential equation. Time derivatives of \( x \) now simply pull down factors of \( i\gamma \), so
\[ \ddot{x} + \omega^2 x = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \gamma^2 y(\omega)e^{-i\gamma t} \, d\gamma + \frac{\omega^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y(\gamma)e^{-i\omega t} \, d\gamma \] (499)
\[ = \frac{1}{m} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} G(\gamma)e^{-i\gamma t} \, d\gamma \] (500)

Collecting all terms together into one integral, we have
\[ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left( -\gamma^2 y(\gamma) + \omega^2 y(\gamma) - \frac{1}{m} G(\gamma) \right) e^{-i\gamma t} \, d\gamma = 0 \] (501)

But this is just the Fourier transform of \( \left( -\gamma^2 y(\gamma) + \omega^2 y(\gamma) - \frac{1}{m} G(\gamma) \right) \), so we can take the inverse Fourier transform of this equation to write
\[ -\gamma^2 y(\gamma) + \omega^2 y(\gamma) - \frac{1}{m} G(\gamma) = 0 \] (502)

The solution for \( y \) is immediate:
\[ y(\gamma) = \frac{G(\gamma)}{m (\omega^2 - \gamma^2)} \] (503)
Finally, we put this expression into the inverse transform to find \( x(t) \), substitute for \( G(\gamma) \), and reverse the order of integration:

\[
x(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{G(\gamma)}{m(\omega^2 - \gamma^2)} \, e^{-i\gamma t} \, d\gamma
\]

\[
= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{m(\omega^2 - \gamma^2)} \left( \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(t')e^{i\gamma t'} \, dt' \right) \, e^{-i\gamma t} \, d\gamma
\]

\[
= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(t')dt' \int_{-\infty}^{\infty} \frac{e^{i\gamma(t'-t)}}{m(\omega^2 - \gamma^2)} \, d\gamma
\]

Now we need to do an integral:

\[
I(t-t') = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-i\gamma(t-t')}}{\omega^2 - \gamma^2} \, d\gamma
\]

Once we have \( I(t-t') \), the solution is given by

\[
x(t) = \int_{-\infty}^{\infty} F(t')I(t-t') \, dt'
\]

We interpret this equation as giving the motion of the system, \( x(t) \), that results from the application of the force \( F(t') \), integrated for all times \( t' \) up to the current moment \( t \). If the expression is to make sense causally, \( I(t-t') \) must vanish if \( t' \) is greater than \( t \). In performing the integral for \( I \), we must make sure this condition holds.

We'll now take a little mathematical aside to evaluate this integral using contour integration.

### 5.2.2 A contour integral

We want to evaluate

\[
I = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-i\gamma\tau}}{\omega^2 - \gamma^2} \, d\gamma
\]

\[
= -\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-i\gamma\tau}}{(\gamma - \omega)(\gamma + \omega)} \, d\gamma
\]

where we have set \( \tau = t - t' \). Since the poles lie on the \( x \)-axis, we need to move them slightly. The choice of which way to shift the poles has
important consequences. Consider what we plan to do. To use contour integration, we need to close the path of integration. The exponential in the numerator allows us to do this. Suppose \( \tau \) is positive. Then the exponential will have a negative real part when \( \gamma \) is in the lower half plane. That is, if \( \gamma = a - ib \) with \( a \) and \( b \) positive real numbers, then

\[
e^{-i\gamma \tau} = e^{-i(a - ib)\tau} = e^{-ia\tau} e^{-b\tau}
\]

so the real part is decaying exponentially.

By contrast, if \( \tau \) is negative, then the exponent has a negative real part when \( \gamma \) is in the upper half plane. But \( \tau \) negative means that \( t - t' < 0 \), and \( \tau \) positive means that \( t - t' > 0 \). Looking at the expression for \( x(t) \),

\[
x(t) = \frac{1}{m} \int_{-\infty}^{\infty} I(t, t') F(t') dt'
\]

we see that if \( I(t, t') \) is nonzero for \( t' > t \), then \( x(t) \) depends on a force that hasn’t acted yet. As noted above, this violates causality! We can insure that this doesn’t happen by pushing the poles in the right way. By writing \( I \) as

\[
I = \frac{1}{2\pi} \lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} \frac{e^{-i\gamma \tau}}{(\gamma - \omega + i\varepsilon)(\gamma + \omega + i\varepsilon)} \, d\gamma
\]

we put both poles in the lower half plane, at \( \gamma = \omega - i\varepsilon \) and at \( \gamma = -\omega - i\varepsilon \). Therefore, only the contour with \( \tau \) positive will give a nonzero integral.

The residue theorem says that the clockwise contour integral around any simple pole gives the \( 2\pi i \) times the residue at the pole:

\[
\oint \frac{f(\gamma)}{\gamma - \gamma_0} \, d\gamma = 2\pi i f(\gamma_0)
\]

We can use the theorem by turning the line integral into a contour integral by adding a semicircle and letting its radius become infinite. When we let the radius of the circle diverge, the factor \( e^{-b\tau} \) goes to zero, so adding the semicircle doesn’t change the value of the integral.

Thus, adding a half circle in the lower half plane to close the contour, we have

\[
I = \frac{1}{2\pi} \lim_{\varepsilon \to 0} \oint \frac{e^{-i\gamma \tau}}{(\gamma - \omega + i\varepsilon)(\gamma + \omega + i\varepsilon)} \, d\gamma
\]

90
\[
\begin{align*}
&= -\frac{2\pi i}{2\pi} \lim_{\varepsilon \to 0} \left( \frac{e^{-i\gamma \tau}}{(\gamma - \omega + i\varepsilon)} \bigg|_{\gamma = -\omega - i\varepsilon} + \frac{e^{-i\gamma \tau}}{(\gamma + \omega + i\varepsilon)} \bigg|_{\gamma = \omega + i\varepsilon} \right) \\
&= -\frac{2\pi i}{2\pi} \left( -\frac{e^{i\omega \tau}}{2\omega} + \frac{e^{-i\omega \tau}}{2\omega} \right) \\
&= -\frac{1}{\omega} \sin \omega \tau
\end{align*}
\] (513)

The minus sign in the second line comes because the contour is clockwise rather than counterclockwise. This result holds for \( \tau > 0 \).

When \( \tau < 0 \), we must close with a semicircle in the upper half plane. Now the contour gives zero because it encloses no poles, so the final result is

\[
I(\tau) = \begin{cases} 
\frac{1}{\omega} \sin \omega \tau & t' < t \\
0 & t' > t
\end{cases}
\] (516)

Now, returning to substitute into the solution for \( x(t) \), we have

\[
x(t) = \frac{1}{m} \int_{-\infty}^{\infty} I(t,t')F(t')dt' \\
= -\frac{1}{m\omega} \int_{-\infty}^{t} F(t')dt' \sin \omega (t-t')
\] (517)

Notice that this is real and causal. We can write the result in another useful form by returning to the exponential notation. Expanding the sine, we have

\[
x(t) = -\frac{1}{2im\omega} \left( \int_{-\infty}^{t} e^{i\omega(t-t')}F(t')dt' - \int_{-\infty}^{t} e^{-i\omega(t-t')}F(t')dt' \right) \\
= -\frac{1}{2im\omega} \left( e^{i\omega t} \int_{-\infty}^{t} e^{-i\omega t'}F(t')dt' - e^{-i\omega t} \int_{-\infty}^{t} e^{i\omega t'}F(t')dt' \right) \\
= -\frac{1}{m\omega} \text{Im} \left( e^{i\omega t} \int_{-\infty}^{t} e^{-i\omega t'}F(t')dt' \right)
\] (519)

\[
= -\frac{1}{2im\omega} \left( \int_{-\infty}^{t} e^{i\omega t} e^{-i\omega t'}F(t')dt' - \int_{-\infty}^{t} e^{-i\omega t} e^{i\omega t'}F(t')dt' \right)
\] (520)

\[
= -\frac{1}{m\omega} \text{Im} \left( e^{i\omega t} \int_{-\infty}^{t} e^{-i\omega t'}F(t')dt' \right)
\] (521)

### 5.2.3 A second approach: a clever trick

There is an easier way to solve this problem, involving a clever trick. Starting with the equation of motion, we add and subtract \( i\omega \ddot{x} \)

\[
\frac{1}{m} F(t) = \ddot{x} + \omega^2 x
\] (522)
\[ \dot{x} + i\omega \dot{x} - i\omega x + \omega^2 x = \frac{d}{dt} (\dot{x} + i\omega x) - i\omega (\dot{x} + i\omega x) \]  \hspace{1cm} (523)

Defining
\[ \xi = \dot{x} + i\omega x \]  \hspace{1cm} (525)

and conversely
\[ x = \frac{1}{\omega} \text{Im} \xi \]  \hspace{1cm} (526)
\[ \dot{x} = \text{Re} \xi \]  \hspace{1cm} (527)

the equation of motion becomes a first order equation for the complex variable \( \xi(t) \):
\[ \frac{d\xi}{dt} - i\omega \xi = \frac{1}{m} F(t) \]

This is straightforward to solve. We can remove the nonderivative term in \( \xi \) by replacing
\[ \xi = A(t) e^{i\omega t} \]  \hspace{1cm} (528)

Then
\[ \dot{\xi} = \dot{A}(t) e^{i\omega t} + i\omega A(t) e^{i\omega t} \]  \hspace{1cm} (529)

Substituting,
\[ \frac{d\xi}{dt} - i\omega \xi = \frac{1}{m} F(t) \]
\[ \dot{A}(t) e^{i\omega t} + i\omega A(t) e^{i\omega t} = \frac{1}{m} F(t) \]
\[ \dot{A}(t) e^{i\omega t} = \frac{1}{m} F(t) \]

Therefore,
\[ \frac{dA}{dt} = e^{-i\omega t} \frac{1}{m} F(t) \]

and we can integrate directly. If the force starts in the far past, we integrate from \( t = -\infty \), so
\[ A(t) = A(-\infty) + \frac{1}{m} \int_{-\infty}^{t} e^{-i\omega t'} F(t') dt' \]  \hspace{1cm} (530)
and the resulting solution for the motion is

$$\xi = A(-\infty)e^{i\omega t} + \frac{1}{m} e^{i\omega t} \int_{-\infty}^{t} e^{-i\omega t'} F(t') dt'$$  \hfill (531)

We could just as easily take the initial time to be $t = 0$ as in Landau and Lifshitz. If we take the initial value of $\xi$ to be zero then the position, $x(t)$ is given by

$$x = \frac{1}{\omega} \text{Im} \xi = \frac{1}{m\omega} \text{Im} \left\{ e^{i\omega t} \int_{-\infty}^{t} e^{-i\omega t'} F(t') dt' \right\}$$  \hfill (532)

in agreement (up to a sign) with our previous result.

Let’s think about the sign ambiguities here. We could just as well have defined

$$\xi = \dot{x} - i\omega x$$  \hfill (533)

but this would give the opposite sign for $x$ in terms of $\xi$, even though the solution would be the same for $\xi$. Similarly, we could have introduced a sign by choosing the opposite sign convention for the Fourier transform,

$$g(\gamma) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\gamma t} dt$$  \hfill (534)

instead of

$$g(\gamma) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{i\gamma t} dt$$  \hfill (535)

and this could change the expression for $x(t)$. In any case, remember that these results are particular solutions, to which we can add an arbitrary homogeneous solution. Whichever sign conventions are chosen, whenever we choose the same initial conditions, the solutions will be equivalent.

Next, we look at the energy of the oscillator. Because there is an external driving force, we no longer have a conservation law, and there is technically no such thing as the energy of the system. However, if at any instant we imagine the force to stop, the oscillator becomes an isolated system again and it will continue with constant energy given by

$$E = \frac{1}{2} m\dot{x}^2 + \frac{1}{2} kx^2$$  \hfill (536)
We therefore can define $E(t)$ to be the energy that the isolated oscillator would have if the force were to stop at time $t$. With this understanding, the energy is given by

$$E(t) = \frac{1}{2} m (\text{Re} \xi)^2 + \frac{1}{2} k \left( \frac{1}{\omega} \text{Im} \xi \right)^2$$  \hspace{1cm} (537)

$$= \frac{1}{2} m \left( (\text{Re} \xi)^2 + (\text{Im} \xi)^2 \right)$$  \hspace{1cm} (538)

$$= \frac{1}{2} m (\xi^* \xi)$$  \hspace{1cm} (539)

Again setting $A(-\infty) = 0$, we have

$$\xi = \frac{1}{m} e^{i\omega t} \int_{-\infty}^{t} e^{-i\omega t'} F(t') dt'$$  \hspace{1cm} (540)

$$\xi^* \xi = \frac{1}{m^2} \left( e^{i\omega t} \int_{-\infty}^{t} e^{-i\omega t'} F(t') dt' \right) \left( e^{-i\omega t} \int_{-\infty}^{t} e^{i\omega t'} F(t') dt' \right)$$  \hspace{1cm} (541)

$$= \frac{1}{m^2} \left| \int_{-\infty}^{t} e^{-i\omega t'} F(t') dt' \right|^2$$  \hspace{1cm} (542)

so the energy is

$$E(t) = \frac{1}{2} m (\xi^* \xi)$$  \hspace{1cm} (543)

$$= \frac{1}{2m} \left| \int_{-\infty}^{t} e^{-i\omega t'} F(t') dt' \right|^2$$  \hspace{1cm} (544)

There is a particularly simple form for the total energy transferred to the oscillator by the force, $E(\infty)$. Since the Fourier transform of the force is just the function

$$G(\gamma) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(t) e^{i\gamma t} \, dt$$  \hspace{1cm} (545)

we can write the total energy transferred as

$$E(\infty) = \frac{1}{2m} \left| \int_{-\infty}^{\infty} e^{-i\omega t'} F(t') dt' \right|^2$$  \hspace{1cm} (546)

$$= \frac{\pi}{m} |G(\omega)|^2$$  \hspace{1cm} (547)
5.3 Oscillations of systems with more than one degree of freedom

Now we consider equilibrium configurations with many degrees of freedom. The way to solve this more general case turns is to transform it into several copies of the single degree of freedom case. To do this we will use some results from matrix algebra.

Consider motion of a system with \(s\) degrees of freedom near a point of stable equilibrium. We again expand the potential, \(U(q) = U(q_1, q_2, \ldots q_s)\), in a Taylor series about the equilibrium point, \(q_0 = (q_{01}, q_{02}, \ldots q_{0s})\):

\[
U(q) = U_0 + \sum_{i=1}^{s} \frac{\partial U}{\partial q_i} \mid_{q_0} (q_i - q_{0i}) + \frac{1}{2} \sum_{i,j=1}^{s} \frac{\partial^2 U}{\partial q_i \partial q_j} \mid_{q_0} (q_i - q_{0i})(q_j - q_{0j}) + \ldots
\]

(548)

As before we can choose new coordinates, \(x_i = q_i - q_{0i}\), and set the energy scale so that \(U_0 = 0\). Because \(q_0\) is a stable point of equilibrium, it must be an extremum of the potential, so \(\frac{dU}{dq} \mid_{q_0} = 0\). Furthermore, stability implies positivity of the matrix of second derivatives, in the sense that the matrix

\[
k_{ij} = \frac{\partial^2 U}{\partial q_i \partial q_j} \mid_{q_0}
\]

(549)

has all of its eigenvalues positive. These definitions reduce the potential to

\[
U(x) = \frac{1}{2} \sum_{i,j=1}^{s} k_{ij} x_i x_j + \ldots
\]

(550)

where the constant matrix \(k_{ij}\) is symmetric and has positive eigenvalues.

For the kinetic energy, we have

\[
T = \frac{1}{2} \sum_{i,j=1}^{s} a_{ij}(q) \dot{q}_i \dot{q}_j
\]

(551)

where, as in the simplest case, the velocities

\[
\dot{q}_i = \frac{d}{dt} (x_i + q_{0i}) = \dot{x}_i
\]

(552)
are considered small so that to second order in small quantities the kinetic energy depends only on the lowest order term in the Taylor series of $a(q)$ about $q_0$:

$$T = \frac{1}{2} \sum_{i,j=1}^{s} a_{ij}(q) \dot{q}_i \dot{q}_j$$  \hspace{1cm} (553)$$

$$= \frac{1}{2} \sum_{i,j=1}^{s} (a_{ij}(q_0) + \ldots) \dot{x}_i \dot{x}_j$$  \hspace{1cm} (554)$$

This time, we define a matrix,

$$m_{ij} = a_{ij}(q_0)$$  \hspace{1cm} (555)$$

which must also have positive eigenvalues because it does when we write it in Cartesian coordinates.

Now to second order the Lagrangian is

$$L = \frac{1}{2} \sum_{i,j=1}^{s} m_{ij} \dot{x}_i \dot{x}_j - \frac{1}{2} \sum_{i,j=1}^{s} k_{ij} x_i x_j$$  \hspace{1cm} (556)$$

If we find the Euler-Lagrange in these coordinates, they are still complicated because the evolution of each coordinate depends on all of the others. We now try to separate the motion into independent parts.

First, replace the coordinates $x_i$ by rotating them. Let $O_{ij}$ be a rotation matrix and define new coordinates $y_i$ by

$$x_i = \sum_{m=1}^{s} O_{im} y_m$$  \hspace{1cm} (557)$$

$$\dot{x}_i = \sum_{m=1}^{s} O_{im} \dot{y}_m$$  \hspace{1cm} (558)$$

Substituting into $L$ we have

$$L = \frac{1}{2} \sum_{i,j=1}^{s} m_{ij} \left( \sum_{m=1}^{s} O_{im} \dot{y}_m \right) \left( \sum_{n=1}^{s} O_{jn} \dot{y}_n \right)$$  \hspace{1cm} (559)$$

$$- \frac{1}{2} \sum_{i,j=1}^{s} k_{ij} \left( \sum_{m=1}^{s} O_{im} y_m \right) \left( \sum_{n=1}^{s} O_{jn} y_n \right)$$  \hspace{1cm} (560)$$
Any real symmetric matrix can be diagonalized by an orthogonal similarity transformation. Choosing to be the particular transformation that diagonalizes $m_{ij}$ puts $m_{ij}$ in the form

$$m_{mn} = \sum_{i,j=1}^{s} O_{mi}^{t} m_{ij} O_{jn}$$

(566)

$$k_{mn} = \sum_{i,j=1}^{s} O_{mi}^{t} k_{ij} O_{jn}$$

(567)

Since $O$ is an orthogonal transformation, $O^{t} = O^{-1}$, so that $m'$ and $k'$ are similarity transformations of $m$ and $k$, respectively. So far, $O$ may be any orthogonal transformation. We now use the following theorem from linear algebra: Any real symmetric matrix can be diagonalized by an orthogonal similarity transformation. Choosing $O$ to be the particular transformation that diagonalizes $m_{ij}$ puts $m_{ij}$ in the form

$$m_{ij} = m_{j} \delta_{ij} = \begin{pmatrix} m_{1} & \cdots & m_{s} \end{pmatrix}$$

(568)

Notice that regardless of the choice for $O$, $k'_{mn}$ is symmetric because $k' = O^{t}kO$ implies

$$(k')^{t} = (O^{t}kO)^{t}$$

(569)

$$= O^{t}k^{t}O$$

(570)

$$= O^{t}kO$$

(571)

$$= k'$$

(572)

97
We can make the kinetic term even simpler by the further transforma-
tion,

\[ y_i = \frac{u_i}{\sqrt{m_i}} \]  \hspace{1cm} (573) \\
\[ \dot{y}_i = \frac{\dot{u}_i}{\sqrt{m_i}} \]  \hspace{1cm} (574)

This turns the “mass matrix”, \( m_{ij} \) into the identity matrix, because the kinetic term is simply

\[ \frac{1}{2} \sum_{m=1}^{s} \sum_{n=1}^{s} m_{mn} \dot{y}_m \dot{y}_n = \frac{1}{2} \sum_{m=1}^{s} \sum_{n=1}^{s} m_{n} \delta_{mn} \dot{y}_m \dot{y}_n \]  \hspace{1cm} (575) \\
\[ = \frac{1}{2} \sum_{m=1}^{s} m_{m} (\dot{y}_m)^2 \]  \hspace{1cm} (576) \\
\[ = \frac{1}{2} \sum_{m=1}^{s} m_{m} \left( \frac{\dot{u}_m}{\sqrt{m_m}} \right)^2 \]  \hspace{1cm} (577) \\
\[ = \frac{1}{2} \sum_{m=1}^{s} \dot{u}_m^2 \]  \hspace{1cm} (578)

Meanwhile, the potential becomes

\[ \frac{1}{2} \sum_{m=1}^{s} \sum_{n=1}^{s} k_{mn} \dot{y}_m \dot{y}_n = \frac{1}{2} \sum_{m=1}^{s} \sum_{n=1}^{s} k_{mn} \frac{\dot{u}_m \dot{u}_n}{m_n} \]  \hspace{1cm} (579) \\
\[ = \frac{1}{2} \sum_{m=1}^{s} \sum_{n=1}^{s} k_{mn} u_m u_n \]  \hspace{1cm} (580)

where \( k_{mn}'' \) remains symmetric.

Reducing the kinetic energy to the simple form \( T = \frac{1}{2} \sum_{m=1}^{s} \dot{u}_m^2 \) is the key step of our derivation, because it means we can do another orthogonal transformation without changing the form of the kinetic energy. For suppose we define

\[ u_i = \sum_{j=1}^{s} O'_{ij} v_j \]  \hspace{1cm} (581)
Then the kinetic energy is

\[
T = \frac{1}{2} \sum_{m=1}^{s} \dot{u}_m^2
\]

(582)

\[
= \frac{1}{2} \sum_{m=1}^{s} \sum_{i,j=1}^{s} O'_{mi}O'_{mj}\dot{v}_i\dot{v}_j
\]

(583)

But

\[
\sum_{m=1}^{s} O'_{mi}O'_{mj} = \sum_{m=1}^{s} O'^{t}_{im}O'_{mj}
\]

(584)

\[
= \sum_{m=1}^{s} O'_{im}^{-1}O'_{mj}
\]

(585)

\[
= \delta_{ij}
\]

(586)

and therefore \( T \) is unchanged in form:

\[
T = \frac{1}{2} \sum_{i,j=1}^{s} \delta_{ij}\dot{v}_i\dot{v}_j
\]

(587)

\[
= \frac{1}{2} \sum_{i=1}^{s} \dot{v}_i^2
\]

(588)

As an aside, we note that leaving the Kronecker delta invariant is one way to define an orthogonal transformation. An orthogonal transformation (in an \( s \)-dimensional Euclidean space) is one that preserves the lengths of all vectors, where the length (in Cartesian coordinates) is given by

\[
l^2 = \sum_{m,n=1}^{s} \delta_{mn}x_m x_n
\]

(589)

Now, a general linear transformation of \( x_m \) is given by

\[
y_m = \sum_{n=1}^{s} A_{mn}x_n
\]

(590)
where $A_{mn}$ is any invertible matrix. If we require the length to be unchanged, we have

$$\sum_{m,n=1}^s \delta_{mn} x_m x_n = \sum_{i,j=1}^s \delta_{ij} y_i y_j$$

(591)

$$= \sum_{i,j=1}^s \delta_{ij} \left( \sum_{m=1}^s A_{im} x_m \right) \left( \sum_{n=1}^s A_{jn} x_n \right)$$

(592)

$$= \sum_{m=1}^s \sum_{n=1}^s \left( \sum_{i,j=1}^s \delta_{ij} A_{im} A_{jn} \right) x_m x_n$$

(593)

Since **all** lengths must be preserved, equality must hold for all $x_i$. The (symmetric) coefficient matrices must therefore be the same,

$$\delta_{mn} = \sum_{i,j=1}^s \delta_{ij} A_{im} A_{jn}$$

(594)

$$= \sum_{i=1}^s A_{jm} A_{jn}$$

(595)

$$= \sum_{i=1}^s A_{mj}^t A_{jn}$$

(596)

But this is precisely the condition

$$A^t = A^{-1}$$

(597)

that frequently defines an orthogonal transformation. Looking two lines back in this derivation, we see that this condition is equivalent to the preservation of the Kronecker delta:

$$\delta_{mn} = \sum_{i,j=1}^s A_{mi}^t \delta_{ij} A_{jn}$$

This is a special case of preserving an arbitrary nondegenerate symmetric matrix. Such a matrix is called a **metric** when it is used to define lengths of vectors.
Returning to our problem, we see that we can perform another orthogonal transformation. When we do so, the Lagrangian becomes

$$L = \frac{1}{2} \sum_{i=1}^{s} v_i^2 - \frac{1}{2} \sum_{m,n=1}^{s} \left( \sum_{i,j=1}^{s} O_{mi}^{n'} k_{ij} O_{jn}' \right) v_m v_n$$

(598)

and since $k''$ is symmetric we can choose $O'$ to diagonalize it. Moreover, recall that the eigenvalues of $k_{mn}$ are known to be positive. We can show that the first orthogonal transformation does not change the eigenvalues as follows. The eigenvalues are given by the characteristic equation,

$$\det (k_{ij} - \lambda I) = 0$$

and, similarly, the eigenvalues of $k'_{mn}$ by

$$\det (k'_{ij} - \lambda I) = 0$$

Substituting $k' = O' k O$ we have

$$0 = \det (k'_{ij} - \lambda I) = \det (O' k_{ij} O - \lambda I) = \det (O^{-1} k_{ij} O - \lambda O^{-1} O) = \det (O^{-1} (k_{ij} - \lambda I) O) = \det (O O^{-1} (k_{ij} - \lambda I)) = \det (k_{ij} - \lambda I)$$

and the eigenvalues are the same as before.

Since the eigenvalues of $m_{mn}$ are also all positive, a similar argument shows that the matrix

$$k''_{mn} = \frac{k'_{mn}}{\sqrt{m_m m_n}}$$

(599)

must also have positive eigenvalues. We may therefore write these eigenvalues in a manifestly positive way, as squares. Checking units,

$$k'' \left( \frac{\text{energy}}{\text{mass} \times \text{length}^2} \right) = k'' \left( \frac{1}{\text{time}^2} \right)$$

we see that they have the units of frequency squared:

$$k''_{mn} = \omega_m^2 \delta_{mn} = \begin{pmatrix} \omega_1^2 & \cdots \\ \cdots & \omega_s^2 \end{pmatrix}$$

(600)
and the Lagrangian takes the final form

\[ L = \sum_{m=1}^{s} \frac{1}{2} (\dot{v}_m^2 - \omega_m^2 v_m^2) \]  

(601)

The degrees of freedom are no longer coupled. The equations of motion are simply

\[ \ddot{v}_m + \omega_m^2 v_m = 0 \]  

(602)

for \( m = 1, 2, \ldots, s \). Solutions for each \( v_m \) are immediate:

\[ v_m = A_m e^{i\omega_m t} \]  

(603)

and the general motion of the system is now described.

### 5.4 Damped oscillations

Damping, for small oscillations, may be treated as a general velocity-dependent function that hampers the motion of the system. For a system with one degree of freedom, with equation of motion,

\[ \ddot{x} + \omega^2 x = 0 \]  

(604)

this amounts to adding a force \( f = f(v) \). If we expand \( f \) in a Taylor series,

\[ f(v) = f_0 - bx + \ldots \]  

(605)

we see that \( \alpha \) must be positive to oppose the motion, and the constant \( f_0 \) must vanish. Discarding the higher order terms we have \( \ddot{x} + \omega^2 x = -b\dot{x} \), or

\[ \ddot{x} + \frac{b}{m} \dot{x} + \omega_0^2 x = 0 \]  

(606)

This is easily solved by the substitution \( x = Ae^{i\omega t} \). The exponential factor cancels and \( \omega \) must satisfy

\[ -\gamma^2 + \frac{ib}{m} \gamma + \omega_0^2 = 0 \]  

(607)

Solving,

\[ \gamma_+ = \frac{ib}{2m} + \sqrt{\frac{\omega_0^2}{4m^2} - \frac{b^2}{4m^2}} \]  

(608)

\[ \gamma_- = \frac{ib}{2m} - \sqrt{\frac{\omega_0^2}{4m^2} - \frac{b^2}{4m^2}} \]  

(609)
so letting \( \lambda = \frac{b}{2m} \) and \( \omega = \sqrt{\omega_0^2 - \lambda^2} \), the general solution is

\[
x = e^{-\lambda t} \left( A_1 e^{i\omega t} + A_2 e^{-i\omega t} \right)
\]  

(610)

For small damping, the frequency of oscillation \( \omega \) is not far from \( \omega_0 \), and the motion is simply oscillation with a slow exponential decrease in amplitude. For large damping, the square root becomes imaginary. Writing

\[
\sqrt{\omega_0^2 - \lambda^2} = i\sqrt{\lambda^2 - \omega_0^2}
\]

so that the square root is again real, we see that the motion becomes a pure exponential decay:

\[
x = A_1 e^{-(\lambda + \sqrt{\lambda^2 - \omega_0^2})t} + A_2 e^{-(\lambda - \sqrt{\lambda^2 - \omega_0^2})t}
\]  

(611)

Notice that both exponentials are decreasing, because \( \lambda - \sqrt{\lambda^2 - \omega_0^2} > 0 \). Therefore, \( x \) approaches zero as the sum of two decreasing exponentials.

An exceptional case occurs when \( \omega = \sqrt{\omega_0^2 - \lambda^2} = 0 \) so that the two roots of the equation for \( \gamma \) are the same. Then one solution is \( x = c_1 e^{-\lambda t} \).

We could just write down the second independent solution (it’s \( c_2 t e^{-\lambda t} \)), but we ought to be able to find it by taking a limit as \( \omega \to 0 \). Expanding the general non-oscillating solution for small \( \omega \) we have

\[
x = A_1 e^{-(\lambda + \omega)t} + A_2 e^{(\lambda - \omega)t} \quad (612)
\]

\[
e^{-\lambda t} \left( A_1 e^{\omega t} + A_2 e^{\omega t} \right) \quad (613)
\]

\[
\approx e^{-\lambda t} \left( A_1 (1 - \omega t) + A_2 (1 + \omega t) \right) \quad (614)
\]

\[
= e^{-\lambda t} \left( (A_1 + A_2) + (A_2 - A_1) \omega t \right) \quad (615)
\]

Defining \( c_1 = A_1 + A_2 \) and \( c_2 = \frac{1}{\omega}(A_2 - A_1) \) and letting \( A_2 - A_1 \) go to zero along with \( \omega \) we get the solution

\[
x = (c_1 + c_2 t) e^{-\lambda t} \quad (616)
\]

which we can check by substitution into the original equation of motion.

For a system with many degrees of freedom, we can add a general linear function of all of the velocities,

\[
f_k = -\sum_{m=1}^{n} b_{km} \ddot{y}_m \quad (617)
\]

103
to the Lagrange equation of motion,
\[ \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_m} - \frac{\partial L}{\partial x_m} = f_m \]  \hspace{1cm} (618)

It can be shown from a statistical analysis that \( b_{km} \) is symmetric, so that we can write
\[ f_k = -\frac{\partial F}{\partial \dot{x}_m} \]  \hspace{1cm} (619)
\[ F = \frac{1}{2} \sum_{m,n=1}^{s} b_{mn} \dot{x}_m \dot{x}_n \]  \hspace{1cm} (620)

and the Lagrange equation becomes
\[ \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_m} - \frac{\partial L}{\partial x_m} = -\frac{\partial F}{\partial \dot{x}_m} \]  \hspace{1cm} (621)

Finally, it is easy to see that
\[ \sum_{m=1}^{s} \dot{x}_m \frac{\partial F}{\partial \dot{x}_m} = 2F \]  \hspace{1cm} (622)

and this lets us express the rate of energy loss of the system in terms of \( F \). As before, we define \( E(t) \) as
\[ E(t) = \sum_{m=1}^{s} \dot{x}_m \frac{\partial L}{\partial \dot{x}_m} - L \]  \hspace{1cm} (623)

so
\[ \frac{dE}{dt} = \frac{d}{dt} \left( \sum_{m=1}^{s} \dot{x}_m \frac{\partial L}{\partial \dot{x}_m} - L \right) \]  \hspace{1cm} (624)
\[ = \frac{d}{dt} \left( \sum_{m=1}^{s} \dot{x}_m \frac{\partial L}{\partial \dot{x}_m} \right) - \frac{dL}{dt} \]  \hspace{1cm} (625)

Let’s look at the first term first. Using the total time derivative of \( L \),
\[ \frac{dL}{dt} = \sum_{m=1}^{s} \dot{x}_m \frac{\partial L}{\partial \dot{x}_m} + \sum_{m=1}^{s} \dot{x}_m \frac{\partial L}{\partial x_m} \]  \hspace{1cm} (626)
we write
\[
\sum_{m=1}^{s} \ddot{x}_m \frac{\partial L}{\partial \dot{x}_m} = \frac{dL}{dt} - \sum_{m=1}^{s} \dot{x}_m \frac{\partial L}{\partial x_m}
\] (627)
so substituting into the rate of change of energy we find
\[
\frac{dE}{dt} = \frac{d}{dt} \left( \sum_{m=1}^{s} \dot{x}_m \frac{\partial L}{\partial \dot{x}_m} - L \right)
\] (628)
\[
= \frac{dL}{dt} - \sum_{m=1}^{s} \ddot{x}_m \frac{\partial L}{\partial x_m} + \sum_{m=1}^{s} \dot{x}_m \frac{d}{dt} \frac{\partial L}{\partial x_m} - \frac{dL}{dt}
\] (629)
\[
= \sum_{m=1}^{s} \ddot{x}_m \left( \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_m} - \frac{\partial L}{\partial x_m} \right)
\] (630)
\[
= \sum_{m=1}^{s} \ddot{x}_m \left( - \frac{\partial F}{\partial \dot{x}_m} \right)
\] (631)
and therefore using the homogeneity expression,
\[
\frac{dE}{dt} = -2F
\] (632)

5.5 Forced oscillations with friction

Now we combine a driving force with damping, in the case of one degree of freedom. For a periodic driving force as before, we have
\[
\ddot{x} + \frac{b}{m} \dot{x} + \omega_0^2 x = f \cos \gamma t
\] (633)
Letting \( x = \text{Re}(Ae^{i\gamma t}) \), we can work with the complex form,
\[
\ddot{x} + \frac{b}{m} \dot{x} + \omega_0^2 x = \frac{f}{m} e^{i\gamma t}
\] (634)
We already have the homogeneous solution above, so all we need is a particular solution. With \( x = Ae^{i\gamma t} \) we have
\[
\left( -\gamma^2 + \frac{ib\gamma}{m} + \omega_0^2 \right) Ae^{i\gamma t} = f e^{i\gamma t}
\] (635)
and therefore

\[ A = \frac{f}{-\gamma^2 + \frac{ib\gamma}{m} + \omega_0^2} \quad (636) \]

\[ = \frac{f}{-\gamma^2 + \frac{ib\gamma}{m} + \omega_0^2} \left( -\gamma^2 - \frac{ib\gamma}{m} + \omega_0^2 \right) \quad (637) \]

\[ = \frac{f \left( (\omega_0^2 - \gamma^2) - i\frac{b\gamma}{m} \right)}{(\omega_0^2 - \gamma^2)^2 + \left( \frac{b\gamma}{m} \right)^2} \quad (638) \]

We can write \( A \) in polar form,

\[ A = re^{i\delta} \quad (639) \]

\[ = r \cos \delta + ir \sin \delta \quad (640) \]

Comparing the two expressions for \( A \) we see that

\[ r = \frac{f}{\sqrt{(\omega_0^2 - \gamma^2)^2 + \left( \frac{b\gamma}{m} \right)^2}} \quad (641) \]

\[ \cos \delta = \frac{f(\omega_0^2 - \gamma^2)}{(\omega_0^2 - \gamma^2)^2 + \left( \frac{b\gamma}{m} \right)^2} \quad (642) \]

\[ \sin \delta = \frac{f \left( - \frac{b\gamma}{m} \right)}{(\omega_0^2 - \gamma^2)^2 + \left( \frac{b\gamma}{m} \right)^2} \quad (643) \]

so more simply,

\[ \tan \delta = \frac{\sin \delta}{\cos \delta} \quad (644) \]

\[ = - \frac{b\gamma}{m(\omega_0^2 - \gamma^2)} \quad (645) \]

Notice that \( \delta \) is negative. This means that the response of the system to the force lags behind the force in phase:

\[ x = \frac{f}{\sqrt{(\omega_0^2 - \gamma^2)^2 + \left( \frac{b\gamma}{m} \right)^2}} e^{i(\gamma t - |\delta|)} \quad (646) \]

The amplitude with friction present is more realistic than without. Instead of diverging when the system is driven at its resonant frequency \( \omega_0 \),
the displacement \( x \) smoothly reaches a maximum. The amplitude at resonance is limited to

\[
x = \frac{mf}{b\gamma}
\]  

(647)

which can be very large if the damping is small. This ratio characterizes the strength of the resonance. Often, we want this response to be as strong as possible, and the resonance peak to be as narrow as possible. For example, a radio needs to be tunable to a specific frequency, and should respond strongly when that frequency is set. The quality, or \( Q \), of a system is defined as the ratio of the maximum response \( mf/b\gamma \) to the full width at half maximum (fwhm). The fwhm is given by setting

\[
\frac{f}{\sqrt{(\omega_0^2 - \gamma^2)^2 + \left(\frac{b\gamma}{m}\right)^2}} = \frac{mf}{2b\gamma}
\]  

(648)

\[
\sqrt{(\omega_0^2 - \gamma^2)^2 + \left(\frac{b\gamma}{m}\right)^2} = \frac{2b\gamma}{m}
\]  

(649)

\[
(\omega_0^2 - \gamma^2)^2 + \left(\frac{b\gamma}{m}\right)^2 = \left(\frac{2b\gamma}{m}\right)^2
\]  

(650)

\[
(\omega_0^2 - \gamma^2)^2 = 3\left(\frac{b\gamma}{m}\right)^2
\]  

(651)

\[
\omega_0^2 - \gamma^2 = \pm \frac{\sqrt{3}b\gamma}{m}
\]  

(652)

The positive driving frequencies satisfying this are

\[
\gamma_+ = \sqrt{\omega_0^2 + \frac{\sqrt{3}b\gamma}{m}}
\]  

(653)

\[
\gamma_- = \sqrt{\omega_0^2 - \frac{\sqrt{3}b\gamma}{m}}
\]  

(654)

so the full width at half maximum is

\[
\gamma_+ - \gamma_- = \sqrt{\omega_0^2 + \frac{\sqrt{3}b\gamma}{m}} - \sqrt{\omega_0^2 - \frac{\sqrt{3}b\gamma}{m}}
\]  

(655)
For weak damping, we can expand the square roots:

\[
\gamma_+ - \gamma_- = \omega_0 \left( \sqrt{1 + \frac{\sqrt{3}b\gamma}{m\omega_0^2}} - \sqrt{1 - \frac{\sqrt{3}b\gamma}{m\omega_0^2}} \right)
\]

\[
\approx \omega_0 \left( 1 + \frac{\sqrt{3}b\gamma}{2m\omega_0^2} - 1 + \frac{\sqrt{3}b\gamma}{2m\omega_0^2} \right)
\]

\[
= \frac{\sqrt{3}b\gamma}{m\omega_0}
\]

and the \(Q\) value is given by

\[
Q = \frac{m_f}{b\gamma} \times \frac{m\omega_0}{\sqrt{3}b\gamma}
\]

\[
= \frac{m^2f\omega_0}{\sqrt{3}b^2\gamma^2}
\]

or since \(\gamma\) is close to \(\omega_0\),

\[
Q \approx \frac{m^2f}{\sqrt{3}b^2\omega_0}
\]

### 5.6 Parametric resonance

There is a wide class of interesting but difficult problems in which the parameters of a system are time dependent, due the the action of external forces. Even the simple harmonic oscillator becomes complicated for general time dependence.

In general, we can write the equation of motion for the harmonic oscillator as

\[
\frac{d(m\dot{x})}{dt} + kx = 0
\]

This form applies even if the mass is changing. Now suppose the parameters \(-m, k\) depend on time. Then we can still reduce the number of parameters to one, as for the simple harmonic oscillator, by defining a new time variable. What we need is a new variable \(\tau\) such that

\[
\frac{df}{d\tau} = m(t) \frac{df}{dt}
\]
for any function $f$, because then we can write

$$\frac{d(mx)}{dt} = \frac{d}{dt} \left( m \frac{dx}{dt} \right)$$

$$= \frac{1}{m(t)} \left( \frac{d}{d\tau} \right) \left( m \frac{dx}{d\tau} \right)$$

$$= \frac{1}{m(t)} \left( \frac{d^2x}{d\tau^2} \right)$$

(664)

(665)

(666)

(667)

We find an expression for $\tau$ using the chain rule:

$$\frac{df}{d\tau} = m(t) \frac{df}{dt}$$

$$= m(t) \frac{d\tau}{dt} \frac{df}{d\tau}$$

(668)

(669)

Therefore, since $\frac{df}{d\tau}$ is arbitrary, we have

$$1 = m(t) \frac{d\tau}{dt}$$

(670)

or

$$\frac{dt}{m(t)} = d\tau$$

$$\tau = \int_t^\tau \frac{dt}{m(t)}$$

(671)

(672)

Using this new variable, we can write the equation of motion as

$$\frac{1}{m(t)} \frac{d^2x}{d\tau^2} + k(t)x = 0$$

(673)

$$\frac{d^2x}{d\tau^2} + m(t(\tau))k(t(\tau))x = 0$$

(674)

and can even make the (somewhat weird) definition $\omega^2(t) = m(t)k(t)$ to write

$$\frac{d^2x}{d\tau^2} + \omega^2(\tau)x = 0$$

(675)
This seems like great progress, but it is still a complicated equation. If we try to integrate it using standard techniques, they don’t work. For example, the energy theorem fails—if we multiply by \( \dot{x} \) to try and integrate, the \( \omega \) gets in the way:

\[
\frac{dx}{d\tau} \frac{d^2x}{d\tau^2} + \frac{dx}{d\tau} \omega^2(\tau) x = 0
\]

(676)

\[
\frac{dx}{d\tau} \frac{d}{d\tau} \left( \frac{dx}{d\tau} \right) = -\frac{dx}{d\tau} \omega^2(\tau) x
\]

(677)

\[
\dot{x} \ddot{x} = -\omega^2(\tau) x \dot{x}
\]

(678)

The problem requires other methods, which depend on the form of the function \( \omega^2(\tau) \).

We will consider the example of a periodic (not harmonic) function. Let the equation of motion be

\[
\frac{d^2x}{dt^2} + f(t) x = 0
\]

(679)

where at any time \( t \), \( f(t) \) satisfies the periodic condition

\[
f(t + T) = f(t)
\]

(680)

When this is the case, we can draw some conclusions about the stability of the system. First notice that if \( x(t) \) is a solution of the problem, then so is \( x(t + T) \), because we change the variable from \( t \) to \( t = t' + T \) in the equation of motion,

\[
\frac{d^2x(t)}{dt^2} + f(t) x(t) = 0
\]

(681)

\[
\frac{d^2x(t' + T)}{dt'^2} + f(t' + T) x(t' + T) = 0
\]

(682)

Replacing \( f(t' + T) = f(t') \), and dropping the primes, we have

\[
\frac{d^2x(t + T)}{dt^2} + f(t) x(t + T) = 0
\]

(683)

showing that \( x(t + T) \) is a solution.

Next, because this is a second order, homogeneous, ordinary differential equation, there will be two independent solutions. Suppose we have two
independent solutions, \( x_1(t) \) and \( x_2(t) \). Then every solution to the equation of motion is an arbitrary linear combination of them:

\[
x(t) = a_1 x_1(t) + a_2 x_2(t)
\]

(684)

for constants \( a_1 \) and \( a_2 \). Now, as shown above, since \( x_1(t) \) and \( x_2(t) \) are solutions, so are \( x_1(t + T) \) and \( x_2(t + T) \). But this means that \( x_1(t + T) \) and \( x_2(t + T) \) can each be expressed as some linear combination of \( x_1(t) \) and \( x_2(t) \) since every solution has this property. Thus, we can write

\[
\begin{align*}
  x_1(t + T) &= a x_1(t) + b x_2(t) \\
  x_2(t + T) &= c x_1(t) + d x_2(t)
\end{align*}
\]

(685)  

(686)

for some constants \( a, b, c \) and \( d \). Therefore, while solutions aren’t necessarily periodic, they are nearly so – their value at time \( t + T \) is given in terms of both basis functions at time \( t \).

Now let’s work with these expressions. If we let \( x_i = (x_1, x_2) \) we can write the matrix equation

\[
\begin{align*}
  x_i(t + T) &= \sum_{j=1}^{2} A_{ij} x_j(t) \\
  x(t + T) &= A x(t)
\end{align*}
\]

(687)  

(688)

where the matrix \( A \) has components

\[
A_{ij} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}
\]

(689)

This matrix is nondegenerate, since \( x_1 \) and \( x_2 \) are independent. By defining a new basis,

\[
y = Tx
\]

(690)

we can write

\[
\begin{align*}
  Tx(t + T) &= TAx(t) \\
  Tx(t + T) &= TAT^{-1}Tx(t) \\
  y(t + T) &= (TAT^{-1})y(t)
\end{align*}
\]

(691)  

(692)  

(693)

and in the new basis \( A \) is replaced by \( B = TAT^{-1} \). As long as \( A \) commutes with its transpose,

\[
[A, A^T] = 0
\]

(694)
(which holds if $A$ is symmetric) we can choose $TAT^{-1}$ to be diagonal. Let us assume that this is possible.

With $B = \left( \begin{array}{cc} \mu_1 \\ \mu_2 \end{array} \right)$ diagonal, we have a new basis with

$$y(t + T) = \left( \begin{array}{c} \mu_1 \\ \mu_2 \end{array} \right) y(t) \quad (695)$$

which simply says that

$$y_1(t + T) = \mu_1 y_1(t) \quad (696)$$
$$y_2(t + T) = \mu_2 y_2(t) \quad (697)$$

We now seek the most general analytic solution to these equations. Consider

$$y_1(t + T) = \mu_1 y_1(t) \quad (698)$$

Let’s transform $y_1$ to a new function,

$$\Pi_1(t) = (\mu_1)^{-t/T} y_1(t) \quad (699)$$

This function is analytic, and moreover,

$$\Pi_1(t + T) = (\mu_1)^{-(t+T)/T} y_1(t + T) \quad (700)$$
$$= \frac{1}{\mu_1} (\mu_1)^{-t/T} (\mu_1 y_1(t)) \quad (701)$$
$$= (\mu_1)^{-t/T} y_1(t) \quad (702)$$
$$= \Pi_1(t) \quad (703)$$

Therefore, we can invert to find $y_1$:

$$y_1(t) = (\mu_1)^{t/T} \Pi_1(t) \quad (704)$$

where $\Pi_1(t)$ is periodic. This result is unique, for if $y'$ is any other function satisfying

$$y'(t + T) = \mu_1 y'(t) \quad (705)$$

we consider the ratio

$$\Pi' = \frac{y'}{y_1} \quad (706)$$
which now satisfies

\[ \Pi'(t + T) = \frac{y'(t + T)}{y_1(t + T)} \]  
\[ = \frac{\mu_1 y'(t)}{\mu_1 y_1(t)} \]  
\[ = \frac{y'(t)}{y_1(t)} \]  
\[ = \Pi'(t) \]

and is therefore periodic. From the last two lines we have

\[ y'(t) = \Pi'(t)y_1(t) \]  
\[ = \Pi'(t)(\mu_1)^{t/T} \Pi_1(t) \]  
\[ = (\mu_1)^{t/T} \Pi''(t) \]

where \( \Pi''(t) = \Pi'(t)\Pi_1(t) \) is periodic. Therefore, we have found all solutions.

Returning to our problem, we have now shown that we can find two independent solutions of the form

\[ y_1(t) = (\mu_1)^{t/T} \Pi_1(t) \]  
\[ y_2(t) = (\mu_2)^{t/T} \Pi_2(t) \]

The two functions \( y_i \) satisfy the differential equation

\[ \frac{d^2 y_i}{dt^2} + f(t)y_i = 0 \]

Multiply the equation for \( y_1 \) by \( y_2 \) and vice versa to obtain:

\[ \frac{d^2 y_1}{dt^2} y_2 + f(t)y_1y_2 = 0 \]  
\[ y_1 \frac{d^2 y_2}{dt^2} + f(t)y_1y_2 = 0 \]

Now subtract:

\[ \frac{d^2 y_1}{dt^2} y_2 - y_1 \frac{d^2 y_2}{dt^2} = 0 \]
But this means that
\[
\frac{d}{dt}\left(\frac{dy_1}{dt}y_2 - y_1\frac{dy_2}{dt}\right) = \left(\frac{d^2y_1}{dt^2}y_2 + \frac{dy_1}{dt}\frac{dy_2}{dt} - y_1\frac{d^2y_2}{dt^2} - y_1\frac{dy_2}{dt}\right)
\]
\[
= \left(\frac{d^2y_1}{dt^2}y_2 - y_1\frac{d^2y_2}{dt^2}\right)
\]
\[
= 0
\]
so that
\[
\frac{dy_1}{dt}y_2 - y_1\frac{dy_2}{dt} = C = \text{const.}
\]
Now, differentiate the equation
\[
y_1(t + T) = \mu_1y_1(t)
\]
we find
\[
\frac{dy_1}{dt}(t + T) = \mu_1\frac{dy_1}{dt}(t)
\]
We can use this to evaluate the expression for \(C\) at time \(t + T\):
\[
C = \frac{dy_1(t + T)}{dt}y_2(t + T) - y_1(t + T)\frac{dy_2(t + T)}{dt}
\]
\[
= \mu_1\frac{dy_1(t)}{dt}y_2(t) - \mu_1y_1(t)\mu_2\frac{dy_2(t)}{dt}
\]
\[
= \mu_1\mu_2\left(\frac{dy_1(t)}{dt}y_2(t) - y_1(t)\frac{dy_2(t)}{dt}\right)
\]
\[
= \mu_1\mu_2C
\]
and therefore
\[
\mu_1\mu_2 = 1
\]
\[
\mu = \mu_1 = \frac{1}{\mu_2}
\]
Finally, the two independent solutions take the form
\[
y_1(t) = \mu^{t/T}\Pi_1(t)
\]
\[
y_2(t) = \frac{1}{\mu^{t/T}}\Pi_2(t)
\]
Unless \(\mu = 1\), either \(\mu\) or \(\frac{1}{\mu}\) is greater than one and the corresponding solution diverges. It is this divergence that is called parametric resonance.
5.6.1 Example

We can work out an example to see this divergence if we use a bit of what is known about parametric resonance. Suppose the function \( f(t) \) is given by

\[
f(t) = \omega_0^2 (1 + h \cos \gamma t)
\]

where \( h \ll 1 \). It is known that for this case, parametric resonance occurs when \( \gamma \) is near \( 2\omega_0 \). To show that we could expand \( x(t) \) in a somewhat complicated Fourier series, but it’s pretty tricky. The equation to solve is

\[
\frac{d^2 x}{dt^2} + \omega_0^2 (1 + h \cos \gamma t) x = 0
\]

and it’s possible to see that something interesting could happen near certain multiples of \( \omega_0 \). To guess what the behavior might be, we can think of the equation as a simple harmonic oscillator with a force,

\[
\frac{d^2 x}{dt^2} + \omega_0^2 x = hx \cos \gamma t
\]

where the force depends on \( x(t) \). In general, any disturbance will make \( x \) varying with the fundamental frequency \( \omega_0 \). Then the \( x \cos \gamma t \) term can be combined to give sum and difference frequencies, \( \gamma + \omega_0 \) and \( \gamma - \omega_0 \). In general, the doesn’t give a resonance. But suppose (for example) \( \gamma = 2\omega_0 \). Then the sum and difference terms have frequencies \( 3\omega_0 \) and \( \omega_0 \). The \( 3\omega_0 \) frequency just acts as a non-resonant force, but the frequency \( \omega_0 \) part of the right side acts as a driving force at the resonant frequency. So driving the system at \( 2\omega_0 \) actually produces a resonance, no matter how small the constant \( h \) is. Notice that, compared to the resonant response at frequency \( \omega_0 \), the motion of the system at frequency \( 3\omega_0 \) will be quite small. Let’s look at this particular resonance in detail.

Let \( \gamma \) be near \( 2\omega_0 \),

\[
\gamma = 2\omega_0 + \varepsilon
\]

where \( \varepsilon \) is small, and let

\[
\lambda = \frac{\gamma}{2} = \omega_0 + \frac{\varepsilon}{2}
\]

Then (using some insight from the known solution) we can expand \( x \) as

\[
x(t) = a(t) \cos \lambda t + b(t) \sin \lambda t
\]
and try and solve for \(a(t)\) and \(b(t)\). Differentiating, we have

\[
\ddot{x}(t) = \dot{a} \cos \lambda t - 2\dot{a} \lambda \sin \lambda t - a\lambda^2 \cos \lambda t + \dot{b} \sin \lambda t + 2\dot{b} \lambda \cos \lambda t - b\lambda^2 \sin \lambda t
\]  
(740)

so substituting,

\[
0 = \dot{a} \cos \lambda t - 2\dot{a} \lambda \sin \lambda t - a\lambda^2 \cos \lambda t - \dot{b} \sin \lambda t + 2\dot{b} \lambda \cos \lambda t - b\lambda^2 \sin \lambda t + \omega_0^2 (1 + h \cos 2\lambda t) (a \cos \lambda t + b \sin \lambda t)
\]  
(741)

Now we expand products of trigonometric functions into sums:

\[
\cos 2\lambda t \cos \lambda t = \frac{1}{2} (\cos 3\lambda t + \cos \lambda t)
\]  
(742)

\[
\cos 2\lambda t \sin \lambda t = \frac{1}{2} (\sin 3\lambda t - \sin \lambda t)
\]  
(743)

so collecting terms we have

\[
0 = \left( \ddot{a} - a\lambda^2 + 2\dot{b} \lambda + a\omega_0^2 + \frac{1}{2} a\omega_0^2 \right) \cos \lambda t + \left( \ddot{b} - 2\dot{a} \lambda - b\lambda^2 + b\omega_0^2 - \frac{1}{2} b\omega_0^2 \right) \sin \lambda t + \frac{1}{2} a\omega_0^2 (\cos 3\lambda t) + \frac{1}{2} b\omega_0^2 \sin 3\lambda t
\]  
(744)

Now we make an assumption which is only justified by doing a much longer calculation (this same calculation, but to higher order) that \(\dot{a} \sim \varepsilon a\) and therefore \(\ddot{a} \sim \varepsilon^2 a\) and is negligible. Also, the coefficient of the terms \(\cos 3\lambda t\) and \(\sin 3\lambda t\) turn out to be of second order, so we drop them (recall the comments above).

With these assumptions, we have

\[
0 = \left( -a\lambda^2 + 2\dot{b} \lambda + a\omega_0^2 + \frac{1}{2} a\omega_0^2 \right) \cos \lambda t + \left( -2\dot{a} \lambda - b\lambda^2 + b\omega_0^2 - \frac{1}{2} b\omega_0^2 \right) \sin \lambda t
\]  
(745)

116
Expanding \( \lambda = \omega_0 + \frac{\varepsilon}{2} \) and setting the coefficients of \( \cos \lambda t \) and \( \sin \lambda t \) separately to zero, we have

\[
0 = -a \left( \omega_0^2 + \omega_0\varepsilon + \frac{\varepsilon^2}{4} \right) + 2 \left( \omega_0 + \frac{\varepsilon}{2} \right) \dot{b} + a\omega_0^2 + \frac{1}{2} ah\omega_0^2 \quad (755)
\]

\[
0 = -2\dot{a} \left( \omega_0 + \frac{\varepsilon}{2} \right) - b \left( \omega_0^2 + \omega_0\varepsilon + \frac{\varepsilon^2}{4} \right) + b\omega_0^2 - \frac{1}{2} bh\omega_0^2 \quad (756)
\]

so dropping terms of second and higher order (for example, we regard \( \varepsilon b \) as second order), we get

\[
0 = -a\omega_0^2 - a\omega_0\varepsilon + 2\omega_0\dot{b} + a\omega_0^2 + \frac{1}{2} ah\omega_0^2 \quad (757)
\]

\[
= \left( -a\varepsilon + 2\dot{b} + \frac{1}{2} ah\omega_0 \right) \omega_0 \quad (758)
\]

\[
0 = -2\dot{a}\omega_0 - b\omega_0^2 - b\omega_0\varepsilon + b\omega_0^2 - \frac{1}{2} bh\omega_0^2 \quad (759)
\]

\[
= \left( -2\dot{a} - b\varepsilon - \frac{1}{2} bh\omega_0 \right) \omega_0 \quad (760)
\]

The final equations,

\[
2\dot{b} - a\varepsilon + \frac{1}{2} ah\omega_0 = 0 \quad (761)
\]

\[
2\dot{a} + b\varepsilon + \frac{1}{2} bh\omega_0 = 0 \quad (762)
\]

are now easy to solve with the substitutions

\[
a = a_0e^{st} \quad (763)
\]

\[
b = b_0e^{st} \quad (764)
\]

which lead to

\[
2b_0s - a_0\varepsilon + \frac{1}{2} h\omega_0a_0 = 0 \quad (765)
\]

\[
2a_0s + b_0\varepsilon + \frac{1}{2} h\omega_0b_0 = 0 \quad (766)
\]

Solving both of these for \( s \), we have:

\[
s = \frac{a_0}{2b_0} \left( \varepsilon - \frac{1}{2} h\omega_0 \right) \quad (767)
\]

\[
s = -\frac{b_0}{2a_0} \left( \varepsilon + \frac{1}{2} h\omega_0 \right) \quad (768)
\]

117
The two expressions for $s$ must be equal:

\[
\frac{a_0}{2b_0} \left( \varepsilon - \frac{1}{2} \hbar \omega_0 \right) = -\frac{b_0}{2a_0} \left( \varepsilon + \frac{1}{2} \hbar \omega_0 \right) \quad (769)
\]
\[
\left( \frac{a_0}{b_0} \right)^2 = \frac{\hbar \omega_0 + 2\varepsilon}{\hbar \omega_0 - 2\varepsilon} \quad (770)
\]
\[
\frac{a_0}{b_0} = \pm \frac{\hbar \omega_0 + 2\varepsilon}{\hbar \omega_0 - 2\varepsilon} \quad (771)
\]

The solutions for $a(t)$ and $b(t)$, and therefore for $x(t)$, will be real only if

\[
\frac{\hbar \omega_0 + 2\varepsilon}{\hbar \omega_0 - 2\varepsilon} > 0 \quad (772)
\]
\[
-h\omega_0 < 2\varepsilon < h\omega_0 \quad (773)
\]

Now, substituting back to find the exponent $s$, we have

\[
s = \frac{a_0}{2b_0} \left( \varepsilon - \frac{1}{2} \hbar \omega_0 \right) \quad (774)
\]
\[
= + \frac{1}{4} \sqrt{\frac{\hbar \omega_0 + 2\varepsilon}{\hbar \omega_0 - 2\varepsilon} (2\varepsilon - \hbar \omega_0)} \quad (775)
\]
\[
= - \frac{1}{4} \sqrt{\left( \frac{1}{2} \hbar \omega_0 \right)^2 - \varepsilon^2} \quad (776)
\]

Finally, we have two independent (and approximate!) solutions for $x(t)$:

\[
x_1(t) = a_0 e^{\frac{1}{2} \sqrt{\left( \frac{1}{2} \hbar \omega_0 \right)^2 - \varepsilon^2} t} \left( \cos \lambda t - \sqrt{\frac{\hbar \omega_0 + 2\varepsilon}{\hbar \omega_0 - 2\varepsilon} \sin \lambda t} \right) \quad (777)
\]
\[
x_2(t) = a_0 e^{-\frac{1}{2} \sqrt{\left( \frac{1}{2} \hbar \omega_0 \right)^2 - \varepsilon^2} t} \left( \cos \lambda t + \sqrt{\frac{\hbar \omega_0 + 2\varepsilon}{\hbar \omega_0 - 2\varepsilon} \sin \lambda t} \right) \quad (778)
\]

The sine and cosine terms are periodic with period $T = \frac{2\pi}{\lambda}$, while the exponential term gives us the value of the constant $\mu$:

\[
\mu = e^{\frac{1}{2} T \sqrt{\left( \frac{1}{2} \hbar \omega_0 \right)^2 - \varepsilon^2}} = e^{\frac{1}{2} \sqrt{\left( \frac{1}{2} \hbar \omega_0 \right)^2 - \varepsilon^2}} \quad (779)
\]

This is just what our general analysis predicted. One solution, $x_1$, diverges while $x_2$ converges.
5.7 Anharmonic oscillations

See text.

5.8 Resonance in nonlinear oscillations

See text.

5.9 Motion in a rapidly oscillating field

Suppose a particle moves in a potential $U(x)$ and a time dependent harmonic force,

$$f(x,t) = f_1(x) \cos \omega t + f_2(x) \sin \omega t$$

(780)

where the frequency $\omega$ is high compared to any frequency of oscillation of the system. We will assume that the response of the system to $f$ is small in comparison to its response to the potential $U(x)$. This is natural because of the high frequency – the direction of the force changes before the system (with its much lower natural frequency) can respond.

For simplicity, consider a particle with a single degree of freedom. Then we have

$$m \ddot{x} = -\frac{dU}{dx} + f(x,t)$$

(781)

Now we suppose that we can treat the effect of the high frequency part perturbatively, so that $x(t)$ is given by the sum

$$x(t) = X(t) + \xi(t)$$

(782)

where $X(t)$ is a solution to the equation without $f(x,t)$ acting:

$$m \ddot{X} = -\frac{dU}{dx}$$

(783)

Since the average of $f(x,t)$ over a period is zero we can average the equation for $x$ over $T = \frac{2\pi}{\omega}$. Defining

$$\langle x(t) \rangle = \frac{\omega}{2\pi} \int_{t}^{t+\frac{2\pi}{\omega}} x(t')dt'$$

(784)
we have

\[
\langle m\ddot{x} \rangle = - \left( \frac{dU}{dx} \right) + \langle f(t) \rangle
\]

(785)

\[
= - \frac{dU}{dx}
\]

(786)

since \( U(x) \) is independent of time. Therefore

\[
\langle x(t) \rangle = X(t)
\]

(787)

Now we can expand the equation

\[
m\ddot{x} = - \frac{dU}{dx} + f(x, t)
\]

(788)

in powers of \( \xi \). First, compute the Taylor series of \( \frac{dU}{dx}(x) \):

\[
\frac{dU}{dx} = \frac{dU}{dx}(X + \xi)
\]

(789)

\[
= \frac{dU}{dx}(X) + \frac{d^2U}{dx^2} \xi + \frac{1}{2} \frac{d^3U}{dx^3} \xi^2 + \ldots
\]

(790)

We'll actually look at the first two terms. We're also going to need the dependence of \( f(x, t) \) on \( X \),

\[
f(x, t) = f(X, t) + \frac{\partial f}{\partial x} \xi + \ldots
\]

(791)

Now to first order we have \( \frac{\partial f}{\partial x} \xi = \frac{\partial f}{\partial X} \xi \) and therefore

\[
m\ddot{X} + m\ddot{\xi} = - \frac{dU}{dx}(X) + \frac{d^2U}{dx^2} \xi + f(X, t) + \frac{\partial f}{\partial X} \xi
\]

(792)

and since

\[
m\ddot{X} = - \frac{dU}{dx}(X)
\]

(793)

we conclude

\[
m\ddot{\xi} = - \frac{d^2U}{dx^2} \xi + f(X, t) + \frac{\partial f}{\partial X} \xi
\]

(794)
This simplifies further, however, because \( \ddot{\xi} \gg \xi \) because it includes a factor \( \omega^2 \gg 1 \). Also, \( f(X, t) \) need not be small. So we can simply take

\[
\begin{align*}
m\ddot{\xi} &= f(X, t) \quad (795) \\
\xi &= \frac{f(X, t)}{m\omega^2} \quad (796)
\end{align*}
\]

This solution for \( \xi \) lets us consider the average behavior of the system. Consider the average:

\[
\begin{align*}
m\langle \ddot{X} \rangle + m\langle \ddot{\xi} \rangle &= -\left< \frac{dU}{dx} \right> - \left< \frac{d^2U}{dx^2} \xi \right> + \langle f(X, t) \rangle + \left< \frac{\partial f}{\partial X} \xi \right> \quad (797) \\
m\langle \ddot{X} \rangle &= -\frac{dU}{dx} - \frac{d^2U}{dx^2} \langle \xi \rangle + \left< \frac{\partial f}{\partial X} \xi \right> \\
&= -\frac{dU}{dx} + \left< \frac{\partial f}{\partial X} \xi \right> \quad (799) \\
&= -\frac{d}{dx} \langle U \rangle - \left< \frac{\partial f}{\partial X} \frac{f}{m\omega^2} \right> \quad (800) \\
&= -\frac{d}{dx} \langle U \rangle - \frac{1}{2m\omega^2} \left< \frac{\partial f^2}{\partial X} \right> \quad (801) \\
&= -\frac{d}{dx} \left( U + \frac{1}{2m\omega^2} \left< f^2 \right> \right) \quad (802)
\end{align*}
\]

Now, since the time average of \( \sin^2 \) or \( \cos^2 \) is \( \frac{1}{2} \), this is simply

\[
\begin{align*}
m\langle \ddot{X} \rangle &= -\frac{d}{dx} \left( U + \frac{f_1^2 + f_2^2}{2m\omega^2} \right) = -\frac{dU_{\text{eff}}}{dx} \quad (803) \\
U_{\text{eff}} &= U + \frac{f_1^2 + f_2^2}{2m\omega^2} \quad (804)
\end{align*}
\]

Notice that the additional term in the effective potential is just the average kinetic energy added by \( f(t) \),

\[
U_{\text{eff}} = U + \frac{\langle \dot{\xi}^2 \rangle}{2m\omega^2} \quad (805)
\]
6 Motion of a rigid body

A rigid body is defined as one in which the distances between any pair of particles comprising the body remain constant. It doesn’t actually matter whether we regard the body as made up of many small particles or as a continuum of material – the essential features remain the same.

6.0.1 Degrees of freedom of a rigid body

First we want to show that the complete description of a rigid body involves 6 degrees of freedom. Three of these degrees of freedom are given by the three components of the center of mass position. We must now show that the condition of being rigid reduces the remaining $3N - 3$ degrees of freedom to just three. To do this we will first prove that there exists an orthonormal frame of reference in which all particles of the body remain at rest. Of course, this frame itself is moving, but describing the motion of the body reduces to describing the motion of this special orthonormal rest frame. Then, we will show that describing the position of the frame requires three degrees of freedom.

We now construct the rest frame of the body. Pick four non-coplanar points in the body, at positions $x_0, x_1, x_2, x_3$. Then the three vectors

\[
\begin{align*}
\mathbf{v}_1 &= x_1 - x_0 \\
\mathbf{v}_2 &= x_2 - x_0 \\
\mathbf{v}_3 &= x_3 - x_0 
\end{align*}
\]

(806)

are independent and connect points in the body. From these, we can construct the three orthonormal unit vectors using the Gramm-Schmidt orthogonalization procedure. Pick one of the vectors and normalize it to unit length:

\[
\mathbf{e}_1 = \frac{\mathbf{v}_1}{|\mathbf{v}_1|}
\]

Now pick a second vector and subtract its projection along $\mathbf{e}_1$:

\[
\mathbf{v}_2' = \mathbf{v}_2 - \mathbf{e}_1 (\mathbf{e}_1 \cdot \mathbf{v}_2)
\]

(807)

Notice that $\mathbf{v}_2' \cdot \mathbf{e}_1 = 0$. Normalize $\mathbf{v}_2'$ to unit length to give our second unit vector

\[
\mathbf{e}_2 = \frac{\mathbf{v}_2'}{|\mathbf{v}_2'|}
\]

122
Finally, find a third vector orthogonal to \( \mathbf{e}_1 \) and \( \mathbf{e}_2 \) by subtracting the \( \mathbf{e}_1 \) and \( \mathbf{e}_2 \) projections from \( \mathbf{v}_3 \) to give \( \mathbf{v}_3 - \mathbf{e}_2 (\mathbf{e}_2 \cdot \mathbf{v}_3) - \mathbf{e}_1 (\mathbf{e}_1 \cdot \mathbf{v}_3) \), then normalize it to length one:

\[
\mathbf{e}_3 = \frac{\mathbf{v}_3 - \mathbf{e}_2 (\mathbf{e}_2 \cdot \mathbf{v}_3) - \mathbf{e}_1 (\mathbf{e}_1 \cdot \mathbf{v}_3)}{|\mathbf{v}_3 - \mathbf{e}_2 (\mathbf{e}_2 \cdot \mathbf{v}_3) - \mathbf{e}_1 (\mathbf{e}_1 \cdot \mathbf{v}_3)|}
\]  

Collecting all three vectors we may write them as \( \mathbf{e}_i = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \). For convenience we write these definitions and their inverses as

\[
\mathbf{e}_i = \sum_j \alpha_{ij} \mathbf{v}_j
\]

\[
\mathbf{v}_i = \sum_j \beta_{ij} \mathbf{e}_j
\]

Next we show that if \( \mathbf{w}_1 \) is the vector connecting points \( A \) and \( B \) of the body, and \( \mathbf{w}_2 \) is the vectors connecting points \( C \) and \( D \) in the body, then their dot product is constant. First, observe that the difference \( \mathbf{w}_2 - \mathbf{w}_1 \) between \( \mathbf{w}_1 \) and \( \mathbf{w}_2 \) is equivalent to the vector connecting \( A \) and \( C \). Therefore, the lengths of \( \mathbf{w}_1, \mathbf{w}_2 \) and \( \mathbf{w}_2 - \mathbf{w}_1 \) all remain constant by our definition of a rigid body. But the dot product of \( \mathbf{w}_1 \) and \( \mathbf{w}_2 \) may be written in terms of these constant lengths as

\[
\mathbf{w}_1 \cdot \mathbf{w}_2 = \frac{1}{2} \left( |\mathbf{w}_1|^2 + |\mathbf{w}_2|^2 - |\mathbf{w}_2 - \mathbf{w}_1|^2 \right)
\]  

so the dot product remains constant. Note also that linear combinations of constant vectors are constant.

Since the basis, \( \mathbf{e}_i \), is build from linear combinations and dot products of constant vectors, the basis is constant. Writing the orthonormality relations for the \( \mathbf{e}_i \) in one equation,

\[
\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}
\]  

we see that these orthonormality relations hold for all time, regardless of changes in the original vectors relative to an inertial frame.

To complete the proof, let \( \mathbf{x} \) be the position of an arbitrary point in the body. Because the \( \mathbf{e}_i \) are independent, we can express \( \mathbf{x} - \mathbf{x}_0 \) in terms of them:

\[
\mathbf{x} - \mathbf{x}_0 = a_i \mathbf{e}_i
\]  

123
Using the orthonormality relation, we have

\[
\begin{align*}
    a_j &= e_j \cdot (a_i e_i) \\
    &= e_j \cdot (x - x_0) \\
    &= \sum_k \alpha_{jk} v_k \cdot (x - x_0) \\
    &= \text{const}.
\end{align*}
\] (813)

The conclusion follows because all of the dot products in the final expression are constant, as are the coefficients. Thus all three of the components of \( x - x_1 \) are constant in the \( e_i \) basis, that is, \( x - x_1 \) is a constant vector in the \( e_i \) frame. Since the vector \( x \) was arbitrary, we have established a coordinate system for the body in which every point of the body is simultaneously at rest.

As a consequence of the rigidity of the body, we can now specify the position of the body completely once we specify the position of the center of mass and the orientations of the initial three unit vectors, \( e_1, e_2, e_3 \). We show below that this requires three numbers. Intuitively, we can see the three parameters as follows. Pick a coordinate system and begin by specifying the center of mass position of the body, \( R \). This requires three degrees of freedom. Now, specify a direction for \( e_1 \) from \( R \). This requires the specification of two angles, say \( \theta \) and \( \varphi \). Next, specify \( e_2 \). Since we already have the direction of \( e_1 \), the direction of \( e_2 \) is already restricted to a plane and we only need to specify an angle in this plane. Once \( e_1 \) and \( e_2 \) are specified, their cross product determines \( e_3 \). We therefore require three center of mass coordinates and three angles to completely specify the position of a rigid body.

### 6.1 Motion of a rigid body

Now we want to describe the motion of a rigid body. As shown above, the motion may be divided into the center of mass motion together with the specification of the orientation of the body relative to the center of mass. We begin by writing the position of an arbitrary point in the body as

\[
x = R + r
\] (814)

with \( R \) the center of mass and \( r \) the vector from the center of mass to the point. The motion of the center of mass has been described in previous sections. We now consider the motion due to changing \( r \).
As shown above, once we know the orientation of the orthonormal frame $\mathbf{e}_i$, we have specified the position of the entire body. Let’s study the types of transformation that change the orthonormal frame $\mathbf{e}_i$ into another orthonormal frame, $\mathbf{e}_i'$. These are precisely the orthogonal transformations (or equivalently, the rotations). To see this, suppose $\Lambda$ is the linear transformation connecting $\mathbf{e}_i$ to an arbitrary $\mathbf{e}_i'$. (Notice that the transformation must be linear in order to take vectors to vectors.) Then we may write

$$\mathbf{e}_i' = \Lambda_{ij} \mathbf{e}_j \quad (815)$$

But we require the new basis to be orthonormal, so we have

$$\delta_{ij} = \mathbf{e}_i' \cdot \mathbf{e}_j' = (\Lambda_{ik} \mathbf{e}_k) \cdot (\Lambda_{jm} \mathbf{e}_m) = \Lambda_{ik} \Lambda_{jm} \delta_{km} = \Lambda_{im} \Lambda_{jm} = \Lambda_{im} \Lambda_{mj} \quad (816)$$

or simply $\Lambda \Lambda^t = 1$. This is the defining relation of an orthogonal transformation. Equivalently,

$$\Lambda^t = \Lambda^{-1} \quad (817)$$

The complex version of the same relationship, $U^\dagger = U^{-1}$, gives the definition of a unitary transformation, familiar from quantum mechanics.

One further definition of the orthogonal transformations is

$$\delta_{ij} = \Lambda_{ik} \Lambda_{jm} \delta_{km} \quad (818)$$

If we think of $\delta_{km}$ as the metric of Euclidean space, then we see that orthogonal transformations are those that preserve the Euclidean inner product.

Now the motion of a rigid body can be described by specifying the motion of the frame $\mathbf{e}_i$ fixed in the body. To an outside inertial observer, however, the frame $\mathbf{e}_i$ moves. Suppose we define an inertial frame of reference centered on the center of mass, chosen to coincide with $\mathbf{e}_i$ at time $t = 0$. As time progresses, the two frames, $\mathbf{e}_i(0)$ (the inertial frame) and $\mathbf{e}_i(t)$ (the body frame, moving with the rigid body) will change relative to one another. However, both remain orthonormal frames, so at any time $t$ they must be related by an orthogonal transformation, $\Lambda(t)$:

$$\mathbf{e}_i(t) = \Lambda_{ij}(t) \mathbf{e}_j(0) \quad (819)$$
The body remains at rest relative to \( \mathbf{e}_i(t) \), but moves relative to \( \mathbf{e}_i(0) \). Now we can write the vector \( \mathbf{r} \) in either frame:

\[
\mathbf{r} = \mathbf{x}_i \mathbf{e}_i(0) = x_i \mathbf{e}_i(t)
\]  

(820)

In general, the components \( \mathbf{x}_i \) are functions of time, but in the body frame the components \( x_i \) are constant.

Equivalently, we may expand the body basis as:

\[
\mathbf{r} = x_i \mathbf{e}_i(t) = x_i (\Lambda_{ij}(t)\mathbf{e}_j(0)) = (x_i \Lambda_{ij}(t)) \mathbf{e}_j(0)
\]

(821)

so we see that the components of \( \mathbf{r} \) with respect to the inertial frame are \( \mathbf{x}_i(t) = x_i \Lambda_{ij}(t) \).

Now we can compute the velocity. The derivative of \( \mathbf{r} \) is

\[
\frac{d\mathbf{r}(t)}{dt} = \lim_{\Delta t \to 0} \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{x_i \mathbf{e}_i(t + \Delta t) - x_i \mathbf{e}_i(t)}{\Delta t}
\]

(822)

To find \( \mathbf{e}_i(t + \Delta t) \) we note that \( \mathbf{e}_i(t) \) and \( \mathbf{e}_j(t + \Delta t) \) must differ by an infinitesimal rotation, \( \Lambda_{ij}(\Delta t) \). That is,

\[
\mathbf{e}_i(t + \Delta t) = \Lambda_{ij}(\Delta t) \mathbf{e}_i(t)
\]

Therefore, we expand

\[
\Lambda_{ij}(\Delta t) = \delta_{ij} + w_{ij} \Delta t
\]

where \( \Delta t \) is infinitesimal and the components \( w_{ij} \) are of order unity. Then the condition on \( w_{ij} \) that makes \( \Lambda_{ij} \) a rotation is

\[
\Lambda_{ij}(t) \Lambda_{jk}^t(t) = \delta_{ik}
\]

(823)

so to linear order in \( \Delta t \),

\[
(\delta_{ij} + w_{ij} \Delta t) (\delta_{jk} + w_{jk}^t \Delta t) = \delta_{ik}
\]

\[
\delta_{ik} + \Delta t \left( w_{ik}^t + w_{ik} \right) = \delta_{ik}
\]

126
Cancelling the identity matrix and $\Delta t$, we see that $w_{ij}$ must be antisymmetric:

$$w'_{ik} + w_{ik} = 0$$

$$w'_{ik} = w_{ki} = -w_{ik} \quad (824)$$

In three dimensions, there is a clever way to write an antisymmetric matrix using the totally antisymmetric symbol $\varepsilon_{ijk}$ defined by

$$\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1$$

$$\varepsilon_{132} = \varepsilon_{213} = \varepsilon_{321} = -1 \quad (825)$$

with all other components zero. Let $\Omega_k$ be three arbitrary numbers, and set

$$w_{ij} = \varepsilon_{ijk} \Omega_k \quad (826)$$

For example,

$$w_{12} = \varepsilon_{12k} \Omega_k$$

$$= \varepsilon_{121} \Omega_1 + \varepsilon_{122} \Omega_2 + \varepsilon_{123} \Omega_3$$

$$= 0 \cdot \Omega_1 + 0 \cdot \Omega_2 + 1 \cdot \Omega_3$$

$$= \Omega_3$$

This gives a general antisymmetric matrix:

$$w_{ij} = \varepsilon_{ijk} \Omega_k$$

$$= \begin{pmatrix} 0 & \Omega_3 & -\Omega_2 \\ -\Omega_3 & 0 & \Omega_1 \\ \Omega_2 & -\Omega_1 & 0 \end{pmatrix} \quad (827)$$

We have now shown that an infinitesimal change of frame may be written in terms of a vector $\Omega_k$ as

$$\mathbf{e}_i(t + \Delta t) = (\delta_{ij} + \Delta t \varepsilon_{ijk} \Omega_k) \mathbf{e}_j(t) \quad (828)$$

Now we can continue with our derivative:

$$\frac{d\mathbf{r}}{dt} = x_i \lim_{\Delta t \to 0} \frac{(\mathbf{e}_i(t + \Delta t) - \mathbf{e}_i(t))}{\Delta t}$$

$$= x_i \lim_{\Delta t \to 0} \frac{((\delta_{ij} + \Delta t \varepsilon_{ijk} \Omega_k) \mathbf{e}_j(t) - \mathbf{e}_i(t))}{\Delta t}$$

$$= x_i \lim_{\Delta t \to 0} \frac{\Delta t \varepsilon_{ijk} \Omega_k \mathbf{e}_j(t)}{\Delta t}$$

$$= x_i \varepsilon_{ijk} \Omega_k \mathbf{e}_j(t) \quad (829)$$

127
The components of the cross product may be written using $\varepsilon_{ijk}$ as

$$(u \times v)_i = \varepsilon_{ijk} u_j v_k$$

or

$$u \times v = (\varepsilon_{ijk} u_j v_k) e_i$$

For example, setting $i = 1$ and summing over all values of $j$ and $k$ for which $\varepsilon_{1jk}$ is nonvanishing, we have

$$(u \times v)_1 = \varepsilon_{1jk} u_j v_k$$
$$= \varepsilon_{123} u_2 v_3 + \varepsilon_{132} u_3 v_2$$
$$= u_2 v_3 - u_3 v_2$$

(830)

Therefore, we can write the components of $\frac{d\mathbf{r}}{dt}$ as

$$\frac{d\mathbf{r}}{dt} = x_i \varepsilon_{ijk} \Omega_k e_j(t)$$

(831)

$$= \varepsilon_{jki} \Omega_k x_i e_j(t)$$

(832)

$$= (\mathbf{\Omega} \times \mathbf{r})_i e_j(t)$$

(833)

or simply

$$\frac{d\mathbf{r}}{dt} = \mathbf{\Omega} \times \mathbf{r}$$

(834)

This may be interpreted as the velocity, relative to the inertial frame, of a particle at position $\mathbf{r} = x_i e_i(t)$ in the body frame.

Adding the center of mass velocity, the total velocity of a particle at position $\mathbf{r}$ in a rigid body is therefore

$$\mathbf{v} = \mathbf{V}_{cm} + \mathbf{\Omega} \times \mathbf{r}$$

(835)

The vector $\mathbf{\Omega}$ is called the *angular velocity* of the rigid body. It depends only on the motion of the frame, and is therefore independent of which point of the body we consider. It is also independent of where we center the frame, since it characterizes only the rotation from one orthonormal frame to another.

Since the velocity of $\mathbf{r}$ is often “derived” in a few lines from a diagram, we review what has been shown. We started only from the assumption that any two particles in a rigid body remain at a fixed distance from one another. It is conceivable, (but not the case), that some complicated
motions of the particles are possible – we know from our definition that each particle must stay on a sphere around each other particle, but this is a long way from our intuition about rigidity. We showed, however, that there exists a frame of reference moving with the body, such that every particle of the body remains at rest in that frame. This implies that a rigid body is described by 6 degrees of freedom: the 3 center of mass coordinates and 3 angles required to specify the orientation of an orthonormal frame.

Next, we derived the class of transformations that preserve the set of orthonormal frames. These are the rotations, or orthogonal transformations. We showed that an infinitesimal rotation is described by an antisymmetric $3 \times 3$ matrix, and hence by a (pseudo-)vector, $\Omega$. From this we could compute directly the rate of change of the frame and therefore the rate of change of position of any particle in the body. These derivations do not rely on our intuitions about rigid bodies, but rather demonstrate the correctness of those intuitions.

Now we can construct the Lagrangian for a rigid body. This begins with the introduction of the moment of inertia tensor.

### 6.2 The inertia tensor

The kinetic energy of a rotating body is found by adding together the kinetic energies of its constituent particles. It is often most convenient to treat the matter making up a macroscopic body as a continuum of material of density $\rho(r)$ and integrate over all points of the body. We can still think of this as a sum, where the infinitesimal volume $d^3r$ at position $r$ has mass $dm = \rho(r)d^3r$ and kinetic energy $dT = \frac{1}{2}(dm)\mathbf{v}^2(r) = \frac{1}{2}\rho(r)\mathbf{v}^2(r)d^3r$. The sum becomes an integral over the volume of the rigid body:

$$T = \frac{1}{2} \int \rho(r)\mathbf{v}^2(r)d^3r \tag{836}$$

Substituting our expression for the velocity and expanding,

$$T = \frac{1}{2} \int \rho(r) (V_{cm} + \Omega \times r)^2 d^3r$$

$$= \frac{1}{2} \int \rho(r) \left( V_{cm}^2 + 2V_{cm} \cdot (\Omega \times r) + (\Omega \times r) \cdot (\Omega \times r) \right) d^3r \tag{837}$$
We evaluate the resulting terms one at a time. First, since the center of mass velocity is constant (for any isolated system!),

\[
\frac{1}{2} \int \rho(r) V_{cm}^2 \, d^3 r = \frac{1}{2} V_{cm}^2 \int \rho(r) \, d^3 r = \frac{1}{2} MV_{cm}^2 \tag{838}
\]

where \( M \) is the total mass of the rigid body. Next, we rewrite the triple product of the second term by cycling the vectors:

\[
V_{cm} \cdot (\Omega \times r) = r \cdot (V_{cm} \times \Omega) \tag{839}
\]

(By the way, this property is easy to see using the \( \varepsilon_{ijk} \) symbol. The left side is \( V_i \varepsilon_{ijk} \Omega_j r_k \) while the right side is \( r_i \varepsilon_{ijk} V_j \Omega_k = r_k \varepsilon_{kij} V_i \Omega_j \), and \( \varepsilon_{ijk} = \varepsilon_{kij} \).

Then, since \( V_{cm} \times \Omega \) is independent of \( r \), we can bring it out of the integral:

\[
\int \rho(r) V_{cm} \cdot (\Omega \times r) \, d^3 r = (V_{cm} \times \Omega) \cdot \int r \rho(r) \, d^3 r \tag{840}
\]

Here we recognize the integral version of the definition of the center of mass, with \( m_i \rightarrow \rho(r) \, d^3 r \) and \( \sum \rightarrow \int \):

\[
R_{cm} = \frac{1}{M} \sum_{i=1}^{N} m_i r_i \tag{841}
\]

\[
R_{cm} = \frac{1}{M} \int r \rho(r) \, d^3 r \tag{842}
\]

The middle term therefore becomes

\[
\int \rho(r) V_{cm} \cdot (\Omega \times r) \, d^3 r = (V_{cm} \times \Omega) \cdot R_{cm} \tag{843}
\]

but we can drop it altogether if we choose the origin of our coordinate system at the center of mass.

Finally, the third term,

\[
\frac{1}{2} \int \rho(r) (\Omega \times r) \cdot (\Omega \times r) \, d^3 r \tag{844}
\]

may be simplified using the identity

\[
\varepsilon_{ijk} \varepsilon_{imn} = \delta_{jm} \delta_{kn} - \delta_{jn} \delta_{km} \tag{845}
\]
It is not hard to see why this is true. Since there is a sum on \( i \), \( i \) must take the same value in both \( \varepsilon \)-symbols. Suppose \( i = 1 \). Then \( jk \) is either 23 or 32, and the same is true of \( mn \). This means that either \( j = m \) and \( k = n \), or \( j = n \) and \( k = m \).

The dot product in the integral may be written in components as

\[
(\boldsymbol{\Omega} \times \mathbf{r}) \cdot (\boldsymbol{\Omega} \times \mathbf{r}) = (\varepsilon_{ijk}\Omega_j r_k) (\varepsilon_{imn}\Omega_m r_n)
\]

\[
= (\varepsilon_{ijk}\varepsilon_{imn})(\Omega_j r_k \Omega_m r_n)
\]

\[
= (\delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km})(\Omega_j r_k \Omega_m r_n)
\]

\[
= \delta_{jm}\delta_{kn}\Omega_j r_k \Omega_m r_n - \delta_{jn}\delta_{km}\Omega_j r_k \Omega_m r_n
\]

\[
= \Omega_m r_n \Omega_m r_n - \Omega_n r_m \Omega_m r_n
\]

\[
= \Omega^2 r^2 - (\Omega \cdot r)^2
\]  

(846)

We can now rearrange the integral into a more useful form.

\[
J_3 = \frac{1}{2} \int \rho(\mathbf{r}) (\boldsymbol{\Omega} \times \mathbf{r}) \cdot (\boldsymbol{\Omega} \times \mathbf{r}) \, d^3 r
\]

\[
= \frac{1}{2} \int \rho(\mathbf{r}) (\Omega^2 r^2 - (\Omega \cdot r)^2) \, d^3 r
\]

\[
= \frac{1}{2} \int \rho(\mathbf{r}) (\Omega_m \Omega_n \delta_{mn} r^2 - \Omega_m \Omega_n r_m r_n) \, d^3 r
\]

\[
= \frac{1}{2} \int \rho(\mathbf{r})\Omega_m \Omega_n (\delta_{mn} r^2 - r_m r_n) \, d^3 r
\]

\[
= \frac{1}{2} \Omega_m \Omega_n \int \rho(\mathbf{r}) (\delta_{mn} r^2 - r_m r_n) \, d^3 r
\]  

(847)

In the last step we extract the constant angular velocity vectors from the integral. This neatly separates the motion from the specific nature of the rigid body, because the remaining integral

\[
I_{mn} = \int \rho(\mathbf{r}) (\delta_{mn} r^2 - r_m r_n) \, d^3 r
\]

(848)

depends only on the way material is distributed within the body. The matrix \( I_{mn} \) is called the *moment of inertia tensor*. Once we have computed \( I_{mn} \) for a given object, we no longer need to consider its shape or how mass is spread within it – its motion depends only on \( I_{mn} \).
The kinetic energy is now given (in center of mass coordinates) by the remarkably simple form

\[
T = \frac{1}{2} M V_{cm}^2 + \frac{1}{2} \Omega_m I_{mn} \Omega_n \tag{849}
\]

and the Lagrangian is

\[
L = \frac{1}{2} M V_{cm}^2 + \frac{1}{2} \Omega_m (\alpha, \beta, \gamma) I_{mn} \Omega_n (\alpha, \beta, \gamma) - U (\mathbf{R}, \alpha, \beta, \gamma) \tag{850}
\]

where \( \mathbf{R} \) is the vector position of the center of mass and \( \alpha, \beta, \gamma \) are three coordinates (generally angles and often absent from \( U \)) specifying the orientation of the frame \( e_i \). Once we introduce a convenient set of angles (the Euler angles), we will use \( L \) to find the equations of motion of a rigid body. For the remainder of this section, we will examine some properties of the moment of inertia tensor.

### 6.2.1 Properties of \( I \)

The moment of inertia tensor, \( I \), has components

\[
I_{mn} = \int \rho(r) (\delta_{mn} r^2 - r_m r_n) \, d^3 r
\]

\[
= \begin{pmatrix}
\int \rho (y^2 + z^2) & -\int \rho xy & -\int \rho xz \\
-\int \rho xy & \int \rho (x^2 + z^2) & -\int \rho yz \\
-\int \rho xz & -\int \rho yz & \int \rho (x^2 + y^2)
\end{pmatrix} \tag{851}
\]

Since the inertia tensor is symmetric, \( I_{mn} = I_{nm} \), it is always possible to find an orthogonal transformation that will diagonalize it. Since an orthogonal transformation leaves an orthonormal frame orthonormal, we can always choose the rest frame \( e_i(t) \) to be the frame in which \( I \) takes diagonal form,

\[
I_{mn} = \begin{pmatrix}
I_1 & & \\
& I_2 & \\
& & I_3
\end{pmatrix} \tag{852}
\]

The eigenvalues, \( I_1, I_2 \) and \( I_3 \) are called the principal moments of inertia. The basis \( e_i(t) \) in which \( I \) is diagonal are called the principal axes of inertia.

Example 1: Find the moment of inertia of a uniform thin rod of length \( l \).
Let the rod lie along the $x$ axis with its center at the origin. The density of the rod is constant in the $x$ direction for $|x| < l$, and is given by Dirac delta functions in the $y$ and $z$ directions. Using the step function

$$
\Theta(x) = \begin{cases} 
0 & x < 0 \\
1 & x \geq 0 
\end{cases}
$$

we therefore have

$$
\rho(x, y, z) = \lambda \Theta\left(\frac{l}{2} - |x|\right) \delta(y) \delta(z)
$$

with the constant $\lambda$ determined by the condition that the total mass is the integral of the density over all space:

$$
M = \int \rho(x, y, z) dx dy dz = \lambda \int_{-\infty}^{l/2} \Theta\left(\frac{l}{2} - |x|\right) \delta(y) \delta(z) dx dy dz = \lambda \int_{-l/2}^{l/2} dx = \lambda l
$$

Therefore, $\lambda = M/l$. By symmetry, the moment of inertia about the $y$ and $z$ axes will be equal, so we can just compute $I_{22}$:

$$
I_{22} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \rho(x, y, z) (x^2 + y^2) d^3 x = \frac{M}{l} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Theta\left(\frac{l}{2} - |x|\right) \delta(y) \delta(z) (x^2 + y^2) d^3 x = \frac{M}{l} \int_{-\infty}^{\infty} \Theta\left(\frac{l}{2} - |x|\right) \delta(y) (x^2 + y^2) d^2 x
$$

$$
= \frac{M}{l} \int_{-\infty}^{\infty} \Theta\left(\frac{l}{2} - |x|\right) x^2 dx = \frac{M}{l} \int_{-l/2}^{l/2} x^2 dx
$$

133
\[
\begin{align*}
I_3 &= \frac{M}{3l} x^3 \bigg|^{1/2}_{-1/2} \\
 &= \frac{M}{3l} \left( \frac{l^3}{8} - \left( -\frac{l^3}{8} \right) \right) \\
 &= \frac{3l^2}{12} 
\end{align*}
\]

We have used the obvious symmetry of the rod to choose our axes to be the principal axes, so we immediately have \( I_3 = I_2 = I_{22} = \frac{Ml^2}{12}, \ I_1 = 0, \) and all other components vanish. Thus

\[
I_{mn} = \begin{pmatrix}
0 & \frac{Ml^2}{12} \\
\frac{Ml^2}{12} & \frac{Ml^2}{12}
\end{pmatrix}
\]

Example 2: Find the moment of inertia tensor of a right circular cylinder of length \( L \) and radius \( R \). Once again we may take the principal axes to be the Cartesian axes, with the cylinder along the \( z \) axis and the center of mass at the origin. The density, however, is now given in cylindrical coordinates by by

\[
\rho(x, y, z) = \frac{M}{\pi R^2 L} \Theta \left( \frac{L}{2} - |z| \right) \Theta (R - r)
\]

For the moment, \( I_3 \), about the \( z \) axis, we have

\[
I_3 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} r^2 \rho(r, \varphi, z) \ r d\varphi \ dr \ dz \\
= \frac{M}{\pi R^2 L} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Theta \left( \frac{L}{2} - |z| \right) \Theta (R - r) r^3 d\varphi \ dr \ dz \\
= \frac{M}{\pi R^2 L} 2\pi \int_{-L/2}^{L/2} \int_{0}^{R} r^3 dr \ dz \\
= \frac{MR^2}{2}
\]

The remaining moments are equal to one another, \( I_1 = I_2 = I \) and equal to the moment about any axis lying in the \( xy \) plane and passing through the
origin. We choose the $x$ axis, so $I$ is given by

\[
I = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (r^2 \cos^2 \varphi + z^2) \rho(r, \varphi, z) \, rd\varphi \, dr \, dz
\]

\[
= \frac{M}{\pi R^2 L} \int_{-L/2}^{L/2} \int_{0}^{R} \int_{0}^{2\pi} \left( \frac{L}{2} - |z| \right) \left( R - r \right) (r^3 \cos^2 \varphi + rz^2) \, d\varphi \, dr \, dz
\]

\[
= \frac{M}{\pi R^2 L} \int_{-L/2}^{L/2} \int_{0}^{R} \int_{0}^{2\pi} \left( \frac{L}{2} - |z| \right) \left( R - r \right) (r^3 \cos^2 \varphi + rz^2) \, d\varphi \, dr \, dz
\]

\[
= \frac{M}{\pi R^2 L} \int_{-L/2}^{L/2} \int_{0}^{R} \int_{0}^{2\pi} \left( \frac{L}{2} - |z| \right) \left( R - r \right) (r^3 \cos^2 \varphi + rz^2) \, d\varphi \, dr \, dz
\]

\[
+ \frac{M}{\pi R^2 L} \int_{-L/2}^{L/2} \int_{0}^{R} \int_{0}^{2\pi} rz^2 d\varphi \, dr \, dz
\]

\[
= \frac{MR^4 L}{4\pi R^2 L} \int_{0}^{2\pi} \cos^2 \varphi \, d\varphi \, dz + \frac{2\pi MR^2}{2\pi R^2 L} \int_{-L/2}^{L/2} z^2 \, dz
\]

\[
= \frac{MR^2}{4\pi} \int_{0}^{2\pi} \cos^2 \varphi \, d\varphi \, dz + \frac{M}{L} \left[ \frac{z^3}{3} \right]_{-L/2}^{L/2}
\]

\[
= \frac{MR^2}{4} + \frac{M}{L} \frac{2L^3}{24}
\]

\[
= \frac{M}{12} (3R^2 + L^2)
\]

(860)

Notice that if $R$ is small compared to the length we have $I_3 < I$, but we can have $I < I_3$ provided

\[
L < \sqrt{3R}
\]

### 6.3 Angular momentum of a rigid body

While we have found the form of the kinetic energy for a rigid body, and established that it has 6 degrees of freedom, we have written the result for $T$ in terms of the angular velocity vector $\Omega$. But $\Omega$ is not necessarily given as the time derivative of a set of angular coordinates. The full Lagrangian treatment of the problem must wait a bit, until we establish a relevant set of angles. In the meantime, however, we can use what we already know to write both the angular momentum and a set of equations of motion.

Recall that the angular momentum has been found for any free body to
be given by

\[ \mathbf{M} = \sum_{i=1}^{N} m_i \mathbf{x}_i \times \mathbf{v}_i \]

or equivalently, for a continuous body,

\[ \mathbf{M} = \int \rho(\mathbf{x}) \ \mathbf{x} \times \mathbf{v}(\mathbf{x}) \ d^3r \]

Working again with the continuous expression, we substitute for \( \mathbf{v} \), using center of mass coordinates, \( \mathbf{R}_{cm} = 0 \). Since

\[ \mathbf{x} = \mathbf{R}_{cm} + \mathbf{r} = \mathbf{r} \]
\[ \mathbf{v} = \mathbf{V}_{cm} + \mathbf{\Omega} \times \mathbf{r} = \mathbf{\Omega} \times \mathbf{r} \]

we have

\[ \mathbf{M} = \int \rho(\mathbf{r}) \ \mathbf{r} \times (\mathbf{\Omega} \times \mathbf{r}) \ d^3r \]
\[ = \int \rho(\mathbf{r}) \ (\mathbf{\Omega} r^2 - \mathbf{r} (\mathbf{\Omega} \cdot \mathbf{r})) \ d^3r \]

Working in components this becomes

\[ M_i = \int \rho(\mathbf{r}) \ (\Omega_i r^2 - r_i (\Omega_j r_j)) \ d^3r \]
\[ = \int \rho(\mathbf{r}) \ (\delta_{ij} \Omega_j r^2 - r_i (\Omega_j r_j)) \ d^3r \]
\[ = \Omega_j \int \rho(\mathbf{r}) \ (\delta_{ij} r_j^2 - r_i r_j) \ d^3r \]
\[ = \Omega_j I_{ij} \]

or in vectors again,

\[ \mathbf{M} = \mathbf{I} \mathbf{\Omega} \]

Therefore, the angular momentum vector \( \mathbf{M} \) is found by acting on the angular velocity vector with the moment of inertia tensor.

It is important to recognize that since the moment of inertia tensor \( \mathbf{I} \) is not generally proportional to the identity, \( \mathbf{M} \) is not in the same direction as \( \mathbf{\Omega} \). For this reason, a free rotating body will appear to \textit{precess}. To see how this occurs, consider a free symmetric top, that is, any body that has one
symmetry axis, and therefore has two of its principal moments of inertia equal. Choosing the body frame $\mathbf{e}_i(t)$ to diagonalize the moment of inertia tensor, we let

$$\mathbf{I} = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I' \end{pmatrix}$$

with $I' \neq I$. Then the conserved angular momentum is

$$\mathbf{M} = \begin{pmatrix} M_1 \\ M_2 \\ M_3 \end{pmatrix} = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I' \end{pmatrix} \begin{pmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \end{pmatrix}$$

Consider the general case, in which the angular velocity vector points in an arbitrary direction. Now because the $I_1 = I_2 = I$, rotations in the $xy$-plane will not change the form of $\mathbf{I}$. Therefore, without loss of generality, we can rotate our axes so that the angular velocity lies in the $xz$ plane, so that $\Omega_2 = 0$. The angular momentum is therefore given by

$$M_1 = I\Omega_1$$
$$M_2 = 0$$
$$M_3 = I'\Omega_3$$

Since the top is free (that is, no forces or torques act), the angular momentum is constant, and we make it our reference direction. Let $\mathbf{M}$ make an angle $\theta$ with the $z$ axis. Then $\mathbf{M} = M(\sin \theta, 0, \cos \theta)$ so we have

$$\Omega_1 = \frac{M \sin \theta}{I}$$
$$\Omega_3 = \frac{M \cos \theta}{I'}$$

This solves the problem of the body’s motion, but now we would like to figure out how this appears in terms of “spin” and “precession”. By the spin velocity, $\Omega_{sp}$, we mean the part of the angular velocity about the symmetry axis of the body. The precession velocity, $\Omega_{pr}$, is observed as the rotation of the symmetry axis about some fixed direction in space. That fixed direction is, of course, the direction of the angular momentum. So the precession vector is the part of $\Omega$ parallel to $\mathbf{M}$. With this understanding, we write:

$$\Omega = \Omega_{sp} + \Omega_{pr}$$

137
The equations of motion of a rigid body

As long as we know that all forces in a given problem come from potentials, the Euler-Lagrange equations are equivalent to Newton’s laws. We are therefore justified in finding the equations of motion for a rigid body using Newton’s second law. For the $i^{th}$ particle this reads:

$$\sum f_i = \frac{dp_i}{dt}$$

Here, the sum is over all forces, $f_i$, acting on the $i^{th}$ particle. We can break this sum into separate parts. First, we separate out all external forces
acting on $m_i$, $f_{ext,i}$. The remainder of $f_i$ is due to interactions within the rigid body, which we may regard as a sum over all particles, $j$, of the force $f_{j-i}$ of particle $j$ on particle $i$:

$$\sum_{j \neq i} f_{j-i} + \sum_{j \neq i} f_{j-i} = \frac{d\mathbf{p}_i}{dt}$$

Now we know, since the rigid body has only 6 degrees of freedom, that there are far more equations here than we need. We simplify to the essential by summing over all particles $i$:

$$\sum_{i=1}^{N} \sum_{j \neq i} f_{i,ext} + \sum_{i=1}^{N} \sum_{j \neq i} f_{j-i} = \sum_{i=1}^{N} \frac{d\mathbf{p}_i}{dt}$$

This simplifies considerably. The first term is just the total external force on the body

$$\sum_{i=1}^{N} \sum_{j \neq i} f_{i,ext} = \sum_{i=1}^{N} \left( \sum_{i=1}^{N} f_{i,ext} \right) = F_{ext}$$

while the rightmost term is just the rate of change of the total momentum,

$$\sum_{i=1}^{N} \frac{d\mathbf{p}_i}{dt} = \frac{d}{dt} \sum_{i=1}^{N} \mathbf{p}_i = \frac{d\mathbf{P}}{dt}$$

Finally, we can use Newton’s third law, requiring equal but opposite reactions to applied forces, to observe that the force that the $j^{th}$ particle exerts on the $i^{th}$ is just the negative of the force the $i^{th}$ exerts on the $j^{th}$:

$$f_{j-i} = -f_{i-j}.$$ Rewriting the double sum of the middle term to allow these forces to combine, we find:

$$\sum_{i=1}^{N} \sum_{j \neq i} f_{j-i} = \sum_{i=1}^{N} \sum_{j < i} (f_{j-i} + f_{i-j}) = 0$$

Therefore,

$$F_{ext} = \frac{d\mathbf{P}}{dt}$$

This provides 3 of our 6 equations of motion.
To find three more, we consider the evolution of the total angular momentum. Again writing

$$\sum r_{i,ext} + \sum_{j \neq i} f_{j \rightarrow i} = \frac{dp_i}{dt}$$

and working in the center of mass frame, we construct the contribution of the \( i^{th} \) particle to \( \mathbf{M} \):

$$\sum r_i \times f_{i,ext} + \sum_{j \neq i} r_i \times f_{j \rightarrow i} = r_i \times \frac{dp_i}{dt}$$

and then sum over all particles in the rigid body,

$$\sum_{i=1}^{N} \sum_{j \neq i} r_i \times f_{i,ext} + \sum_{i=1}^{N} \sum_{j \neq i} r_i \times f_{j \rightarrow i} = \sum_{i=1}^{N} r_i \times \frac{dp_i}{dt}$$

Again we simplify each term. The first term is just the sum of all externally produced torques on the body

$$\mathbf{K} = \sum_{i=1}^{N} \sum_{j \neq i} r_i \times f_{i,ext}$$

that is, the total torque. \( \mathbf{K} \) is due solely to external forces. For the second term we again use Newton’s third law:

$$\sum_{i=1}^{N} \sum_{j \neq i} r_i \times f_{j \rightarrow i} = \sum_{i=1}^{N} \sum_{j=1}^{i} (r_i \times f_{j \rightarrow i} + r_j \times f_{i \rightarrow j})$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{i} (r_i - r_j) \times f_{j \rightarrow i}$$

Since the force between the \( i^{th} \) and \( j^{th} \) particles acts along the line between the particles, this cross product vanishes and the sum is zero term by term. Notice that if the force did not act along the line joining the particles, then a pair of particles could spontaneously develop arbitrarily large angular velocity. We never observe this.
Finally, we may rewrite the right hand side as

\[ \sum_{i=1}^{N} \mathbf{r}_i \times \frac{d\mathbf{p}_i}{dt} = \sum_{i=1}^{N} \left( \frac{d}{dt} (\mathbf{r}_i \times \mathbf{p}_i) - \frac{d}{dt} (\mathbf{r}_i \times \mathbf{p}_i) \right) \]

\[ = \frac{d}{dt} \sum_{i=1}^{N} \mathbf{r}_i \times \mathbf{p}_i - \sum_{i=1}^{N} \mathbf{v}_i \times m_i \mathbf{v}_i \]

\[ = \frac{d\mathbf{M}}{dt} \]

This means that the equation gives the rate of change of the angular momentum in terms of the total torque of the system. The final equations of motion are therefore

\[ \mathbf{F}_{\text{ext}} = \frac{d\mathbf{P}}{dt} \]

\[ \mathbf{K} = \frac{d\mathbf{M}}{dt} \]

These equations are particularly useful for the field of statics, in which there is no motion of the system. When this is the case, the external forces and torques must exactly cancel:

\[ \mathbf{F}_{\text{ext}} = 0 \]
\[ \mathbf{K} = 0 \]

As a simple example, consider a ladder leaning against a wall at an angle \( \theta \) from the vertical. There are three external forces: (1) gravity, \( mg \) acting downward, (2) the force of the wall pushing on the top of the ladder, and (3) the force of the floor pushing on the foot of the ladder. Friction with the floor allows this forces to have a component in both the tangential and normal directions. Taking the ladder to lie in the \( xz \) plane, we have tangential and normal forces

\[ N_{\text{wall}}, f_{x,\text{floor}}, N_{\text{floor}} \]

Because the system is static, the total of all forces in each direction must be zero, so

\[ N_{\text{floor}} - Mg = 0 \]
\[ f_{x,\text{floor}} - N_{\text{wall}} = 0 \]
Therefore,

\[ f_{x,\text{floor}} = N_{\text{wall}} \]
\[ N_{\text{floor}} = Mg \]

This eliminates two of the three unknowns, but leaves the third undetermined. Thus, we cannot solve the problem by considering the sums of forces alone. We must also look at the torques. If the length of the ladder is \( L \), then the total clockwise torque is

\[ \frac{L}{2} (N_{\text{wall}} \cos \theta + f_{x,\text{floor}} \cos \theta - N_{\text{floor}} \sin \theta) = 0 \]

so combining the equations,

\[ N_{\text{wall}} = \frac{1}{2} Mg \tan \theta \]
\[ f_{x,\text{floor}} = \frac{1}{2} Mg \tan \theta \]
\[ N_{\text{floor}} = Mg \]

### 6.5 Euler angles

A full treatment of rotating rigid bodies requires a clear specification of a set of coordinates. We choose a set by explicitly relating our moving frame, \( \mathbf{e}_i(t) \), to the inertial frame, \( \mathbf{e}_i(0) \). The transformation that takes one frame to the other is, as we have shown, an orthogonal transformation and is characterized by three parameters. We now choose these parameters in a convenient way as follows. Beginning with the body and inertial frames aligned we perform three rotations:

1. Rotate the body frame about the common \( z \) axis by an angle \( \phi \).
2. Rotate the body frame about its new \( x \) axis by an angle \( \theta \).
3. Rotate the body frame about its new \( z \) axis by an angle \( \psi \)

The three angles \((\theta, \phi, \psi)\) are the Euler angles. It is useful to think of the first two rotations as positioning the direction of the \( z \) axis of the body frame, and the final rotation as spinning about that orientation. (See figure 47).
The first two rotations are closely related to the usual spherical angles $(\theta, \varphi)$, but with one small difference: instead of using the usual azimuthal angle $\varphi$, we have $\phi = \varphi + \frac{\pi}{2}$ because $\phi$ refers to the position of the $x$ axis while $\varphi$ refers to the $xy$ projection of the $z$ axis.

It is easy, if a bit messy, to construct the full specification of $\mathbf{e}_i(t)$ in terms of $\mathbf{e}_i(0)$ because each rotation is a simple rotation of the new position. The result is that the total rotation is just a compounding of three simple rotations. Let the fixed frame be

$$\mathbf{e}_i(0) = (i, j, k)$$

We now implement the three rotations.

First, we rotate by an angle $\phi$ about the $k$ axis:

$$\begin{pmatrix}
\mathbf{e}_1' \\
\mathbf{e}_2' \\
\mathbf{e}_3'
\end{pmatrix} = \begin{pmatrix}
\cos \phi & -\sin \phi \\
\sin \phi & \cos \phi
\end{pmatrix}
\begin{pmatrix}
i \\
j \\
k
\end{pmatrix}
= \begin{pmatrix}
i \cos \phi - j \sin \phi \\
i \sin \phi + j \cos \phi \\
k
\end{pmatrix}$$

Next, rotate by $\theta$ about the new $x$ axis. This leaves the new $x$ axis fixed, so

$$\begin{pmatrix}
\mathbf{e}_1'' \\
\mathbf{e}_2'' \\
\mathbf{e}_3''
\end{pmatrix} = \begin{pmatrix}
1 & 0 & 0 \\
0 & \cos \theta & -\sin \theta \\
0 & \sin \theta & \cos \theta
\end{pmatrix}
\begin{pmatrix}
i \cos \phi - j \sin \phi \\
i \sin \phi + j \cos \phi \\
k
\end{pmatrix}
= \begin{pmatrix}
i \cos \phi - j \sin \phi \\
i \sin \phi \cos \theta + j \cos \phi \cos \theta - k \sin \theta \\
i \sin \phi \sin \theta + j \cos \phi \sin \theta + k \cos \theta
\end{pmatrix}$$

Finally, we rotate by $\psi$ about the $\mathbf{e}_3''$ axis:

$$\begin{pmatrix}
\mathbf{e}_1(t) \\
\mathbf{e}_2(t) \\
\mathbf{e}_3(t)
\end{pmatrix} = \begin{pmatrix}
\cos \psi & -\sin \psi \\
\sin \psi & \cos \psi
\end{pmatrix}
\begin{pmatrix}
i \cos \phi \cos \psi - j \sin \phi \cos \psi - i \sin \phi \cos \theta \sin \psi \\
i \cos \psi \cos \phi \cos \theta + j \sin \psi \cos \phi \cos \theta - k \cos \psi \cos \theta - i \sin \phi \cos \theta \cos \psi - k \sin \theta \sin \psi
\end{pmatrix}$$

The final result is

$$\mathbf{e}_1(t) = \begin{pmatrix}
i \cos \phi \cos \psi - j \sin \phi \cos \psi - i \sin \phi \cos \theta \sin \psi
\end{pmatrix}$$
\[-j \cos \phi \cos \theta \sin \psi + k \sin \theta \sin \psi\]

\[\mathbf{e}_2(t) = i \cos \phi \sin \psi - j \sin \phi \sin \psi + i \sin \phi \cos \theta \cos \psi + j \cos \phi \cos \theta \cos \psi - k \sin \theta \cos \psi\]

\[\mathbf{e}_3(t) = i \sin \phi \sin \theta + j \cos \phi \sin \theta + k \cos \theta\]

Carrying out the final multiplication we find \(\mathbf{e}_i(t)\) given by

\[\mathbf{e}_1(t) = i (\cos \phi \cos \psi - \sin \phi \cos \theta \sin \psi) - j (\sin \phi \cos \psi + \cos \phi \cos \theta \sin \psi) + k \sin \theta \sin \psi\]

\[\mathbf{e}_2(t) = i (\cos \phi \sin \psi + \sin \phi \cos \theta \cos \psi) - j (\sin \phi \sin \psi - \cos \phi \cos \theta \cos \psi) - k \sin \theta \cos \psi\]

\[\mathbf{e}_3(t) = i \sin \phi \sin \theta + j \cos \phi \sin \theta + k \cos \theta\]

Now we could let each of the Euler angles be a function of time, and compute the angular velocity vector by differentiating and using

\[\frac{d\mathbf{e}_i(t)}{dt} = \sum_{\jmath \neq k} \epsilon_{ijk} \mathbf{e}_j(t) \Omega_k\]

to read off the components \(\Omega_k\). However, the calculation is tedious, and we can get the result much more readily by writing \(\Omega\) as the sum of three rotations:

\[\Omega = \dot{\phi} \mathbf{n}_\phi + \dot{\theta} \mathbf{n}_\theta + \dot{\psi} \mathbf{n}_\psi\]

where the directions of the three rotations can be determined from Figure 47. Changing \(\phi\) is a rotation about the original, fixed \(z\) axis. Projecting this direction into the moving frame, we find

\[\mathbf{n}_\phi = \mathbf{e}_3(t) \cos \theta + \mathbf{\rho} \sin \theta\]

where

\[\mathbf{\rho} = \mathbf{e}_1(t) \sin \psi + \mathbf{e}_2(t) \cos \psi\]

Changing the angle \(\theta\) constitutes a rotation about the line of nodes, so

\[\mathbf{n}_\theta = \mathbf{e}_1(t) \cos \psi - \mathbf{e}_2(t) \sin \psi\]
Finally, changing $\psi$ is a rotation about the body $z$ axis, so

$$n_\psi = e_3(t)$$

(864)

Collecting everything together, we have

$$\Omega = \dot{\phi} \left( e_3(t) \cos \theta + e_1(t) \sin \psi \sin \theta + e_2(t) \cos \psi \sin \theta \right) + \dot{\theta} \left( e_1(t) \cos \psi - e_2(t) \sin \psi \right) + \dot{\psi} e_3(t)$$

$$= \left( \dot{\theta} \cos \psi + \dot{\phi} \sin \psi \sin \theta \right) e_1(t) + \left( \dot{\phi} \cos \psi \sin \theta - \dot{\theta} \sin \psi \right) e_2(t) + \left( \dot{\phi} \cos \theta + \dot{\psi} \right) e_3(t)$$

(865)

Now that the angular velocity is expressed in terms of the Euler angles, we can write the kinetic energy in terms of coordinates. The result simplifies considerably for a top if we take $e_3(t)$ as the symmetry axis. Setting $I_1 = I_2 = I$ the kinetic energy for a top is

$$T = \frac{1}{2} I \left( \dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta \right) + \frac{1}{2} I_3 \left( \dot{\phi} \cos \theta + \dot{\psi} \right)^2$$

We next use this form of $T$ to study the motion of a top in a gravitational field with one point fixed.

### 6.6 Rigid bodies in contact

This section deals with the elementary subject of statics, but also contains an important discussion of constraints. There are two types: holonomic and non-holonomic, depending on whether the constraints are imposed as relations between the coordinates (holonomic) or are more complicated functions of velocities, forces, etc. (non-holonomic). In general, non-holonomic constraints are not integrable. Holonomic constraints may be worked into the Euler-Lagrange equations in a systematic way. See Goldstein for a detailed treatment.

Briefly, holonomic constraints may be written as relationships between the otherwise unconstrained coordinates of a system. Generally the constraint may be written as the vanishing of some functions of the coordinates:

$$c^{(a)}(q_i) = 0$$

145
for \( a = 1, \ldots, k \). In this case, we modify the Lagrangian by adding the constraints, multiplied by arbitrary parameters \( \lambda_a \) called Lagrange multipliers. The new action is then

\[
S = \int \left[ L(q_i, \dot{q}_i) + \lambda_a c_i^{(a)}(q_i) \right] dt
\]

This expression is treated as a new system with an increased number of degrees of freedom – we vary not only the coordinates but also the parameters \( \lambda_a \). The equations of motion following from the \( \lambda_a \) variations are just the constraint equations, while the \( \lambda_a \) are determined from the remaining equations. If the original system had \( N \) degrees of freedom, the new system has \( N + k \), with \( k \) of them being the constraint equations, another \( k \) allowing us to solve for the \( \lambda_a \), and the final \( N - k \) describing the remaining unconstrained degrees of freedom of the system. The solution for the \( \lambda_a \) can be shown to give the forces required to impose the constraints.

For non-holonomic systems, only certain cases can be handled. One result is that, for a set of \( k \) constraints (with \( a = 1, 2, \ldots, k \)) of the form

\[
\sum_{i=1}^{N} c_i^{(a)}(q) \dot{q}_i = 0
\]

we may solve

\[
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_i} \right) - \left( \frac{\partial L}{\partial q_i} \right) = \sum_{a=1}^{k} \lambda_a c_i^{(a)}(q)
\]

together with the constraints. This is a system of \( N + k \) equations for the \( N + k \) unknown functions \((q_i, \lambda^{(a)})\).

### 6.7 Lagrangian for the symmetric top

As an example, we consider the case of a symmetric top, with one point fixed. Let \( e_3(t) \) be the symmetry axis. This guarantees that the Lagrangian is independent of the angle \( \psi \) (but not \( \dot{\psi} \)). Then (with our usual convention that the body frame is chosen so that the moment of inertia tensor is diagonal), the moment of inertia takes the form \( I_1 = I_2 = I \). We will assume \( I > I_3 \). Since the choice of \( \psi \) cannot change \( L, T \) or \( U \), we can set \( \dot{\psi} = 0 \), so the initial angular velocity reduces to

\[
\Omega = \dot{\theta} e_1(t) + (\dot{\phi} \sin \theta) e_2(t) + (\dot{\phi} \cos \theta + \dot{\psi}) e_3(t)
\]
Taking the center of mass at the origin, the kinetic energy is
\[
T = \frac{1}{2} M \dot{\mathbf{R}}^2 + \frac{1}{2} I \left( \dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta \right) + \frac{1}{2} I_3 \left( \dot{\phi} \cos \theta + \dot{\psi} \right)^2 \tag{869}
\]
Suppose the top has one point fixed on a horizontal plane, while gravity acts at the center of mass. If the distance from the fixed tip of top to the center of mass is \( a \), then the potential is given by
\[
U = mga \cos \theta \tag{870}
\]
The velocity \( \dot{\mathbf{R}} \) of the center of mass is simply \( \dot{\mathbf{R}} = \Omega \times a \mathbf{e}_3 (t) \) where \( a \mathbf{e}_3 (t) \) is the position of the center of mass relative to the fixed point. The Lagrangian is therefore
\[
L = \frac{1}{2} M (\Omega \times a \mathbf{e}_3 (t))^2 + \frac{1}{2} I \left( \dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta \right) + \frac{1}{2} I_3 \left( \dot{\phi} \cos \theta + \dot{\psi} \right)^2 - mga \cos \theta \tag{871}
\]
To simplify the kinetic energy due to center of mass motion, we see that the terms involving the center of mass become
\[
\frac{1}{2} M (\Omega \times a \mathbf{e}_3 (t))^2 = \frac{1}{2} M (\Omega \times a \mathbf{e}_3 (t)) \cdot (\Omega \times a \mathbf{e}_3 (t)) = \frac{1}{2} M a^2 \mathbf{e}_3 (t) \cdot ((\Omega \times \mathbf{e}_3 (t)) \times \Omega) = \frac{1}{2} M a^2 \mathbf{e}_3 (t) \cdot (\mathbf{e}_3 (t) \Omega^2 - \Omega (\mathbf{e}_3 (t) \cdot \Omega)) = \frac{1}{2} M a^2 (\Omega^2 - (\mathbf{e}_3 (t) \cdot \Omega)^2) = \frac{1}{2} M a^2 \left( \dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta \right)
\]
The Lagrangian may then be rewritten as:
\[
L = \frac{1}{2} (I + Ma^2) \left( \dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta \right) + \frac{1}{2} I_3 \left( \dot{\phi} \cos \theta + \dot{\psi} \right)^2 - mga \cos \theta
\]
\[
= \frac{1}{2} I' \left( \dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta \right) + \frac{1}{2} I_3 \left( \dot{\phi} \cos \theta + \dot{\psi} \right)^2 - mga \cos \theta \tag{872}
\]
The center of mass motion, governed by the constraint of the fixed point, simply alters the effective value of the moment of inertia orthogonal to the spin.
6.8 Tops

We immediately see that both $\psi$ and $\phi$ are cyclic, so their conjugate momenta are constant:

\[
p_{\phi} = \frac{\partial L}{\partial \dot{\phi}} = I'\phi \sin^2 \theta + I_3 \left( \dot{\phi} \cos \theta + \dot{\psi} \right) \cos \theta = \text{const.} \quad (873)
\]

\[
p_{\psi} = \frac{\partial L}{\partial \dot{\psi}} = I_3 \left( \dot{\phi} \cos \theta + \dot{\psi} \right) = \text{const.} \quad (874)
\]

Using the second, the first may be written as

\[
p_{\phi} = I'\dot{\phi} \sin^2 \theta + p_{\psi} \cos \theta
\]

This may be solved for $\dot{\phi}:

\[
I'\dot{\phi} \sin^2 \theta = p_{\phi} - p_{\psi} \cos \theta
\]

\[
\dot{\phi} = \frac{p_{\phi} - p_{\psi} \cos \theta}{I' \sin^2 \theta} \quad (875)
\]

The final equation of motion is

\[
0 = \frac{d}{dt} I' \dot{\theta} - \frac{\partial}{\partial \theta} \left( \frac{1}{2} I' \left( \dot{\theta}^2 + \dot{\phi}^2 \sin^2 \theta \right) + \frac{1}{2} I_3 \left( \dot{\phi} \cos \theta + \dot{\psi} \right)^2 - mga \cos \theta \right)
\]

\[
= I' \ddot{\theta} - I' \dot{\phi}^2 \sin \theta \cos \theta + I_3 p_{\psi} \dot{\theta} \phi \sin \theta - mga \sin \theta \quad (877)
\]

It is instructive to compare these equations with the equations of motion derived previously. Since the torque is

\[
K = (mga \sin \theta) e_3(t) \times n_\theta
\]

\[
= (mga \sin \theta) e_3(t) \times (e_1(t) \cos \psi - e_2(t) \sin \psi)
\]

\[
= mga \sin \theta (e_2(t) \cos \psi + e_1(t) \sin \psi)
\]

where we have used the fact that rotating about $\theta$ occurs about the line of nodes,

\[
n_\theta = e_1(t) \cos \psi - e_2(t) \sin \psi \quad (878)
\]

Then writing the torque equation we have:

\[
K = \frac{dM}{dt}
\]
where

\[ K = mga \sin \theta \left( e_2(t) \cos \psi + e_1(t) \sin \psi \right) \]

and

\[
\begin{align*}
M &= \mathbf{I} \Omega \\
&= I \dot{\mathbf{e}}_1(t) + I \left( \dot{\phi} \sin \theta \right) \mathbf{e}_2(t) + I_3 \left( \dot{\phi} \cos \theta + \dot{\psi} \right) \mathbf{e}_3(t) \\
&= I \dot{\mathbf{e}}_1(t) + I \left( \dot{\phi} \sin \theta \right) \mathbf{e}_2(t) + I_3 p_\psi \mathbf{e}_3(t) \quad (879)
\end{align*}
\]

This is rather ugly if we substitute our expression for \( \dot{\phi}(\theta) \), but we can get a lot of information out by treating it perturbatively. Since we know that a top can be made to hold a fixed angle while precessing about a vertical axis, we can seek a consistent solution with \( \dot{\theta} = \theta_0 = \text{const.} \). Substituting into the solution for \( \dot{\phi} \), we find that \( \dot{\phi} = \frac{p_\phi - p_\psi \cos \theta_0}{I \sin^2 \theta_0} \) is constant, \( \dot{\phi}_0 \). Then the equation for \( \dot{\theta} \) becomes

\[
\begin{align*}
0 &= I \dot{\phi}_0^2 \sin \theta_0 \cos \theta_0 - p_\psi \dot{\phi}_0 \sin \theta_0 + mga \sin \theta_0 \quad (880) \\
0 &= I \dot{\phi}_0^2 \cos \theta_0 - p_\psi \dot{\phi}_0 + mga \\
\cos \theta_0 &= \frac{p_\psi \dot{\phi}_0 - mga}{I \dot{\phi}_0^2} \quad (881) \\
\end{align*}
\]

Therefore, solutions with \( \theta_0 < \frac{\pi}{2} \) exist as long as

\[ \dot{\phi}_0 p_\psi > mga \quad (883) \]

that is, the product of the precession rate \( \dot{\phi}_0 \) and the momentum \( p_\psi \) about the axis must be greater than the gravitational potential. This can be insured, for example, by spinning the top fast enough.

Now let’s look at what happens if we perturb this smooth motion a little bit. This lets us describe a wider class of solutions, while at the same time checking the stability. Let \( \theta = \theta_0 + \varepsilon \), with \( \varepsilon \ll \theta_0 \). Then we can expand the equation of motion for \( \dot{\theta} \) to first order in \( \varepsilon \) :

\[
\begin{align*}
0 &= I \ddot{\theta} - I \dot{\phi}_0^2 \sin \theta \cos \theta + p_\psi \dot{\phi}_0 \sin \theta - mga \sin \theta \quad (884) \\
&= I \ddot{\varepsilon} - I \dot{\phi}_0^2 \sin(\theta_0 + \varepsilon) \cos(\theta_0 + \varepsilon) \\
&\quad + p_\psi \dot{\phi}_0 \sin(\theta_0 + \varepsilon) - mga \sin(\theta_0 + \varepsilon) \quad (885)
\end{align*}
\]
and
\[
\dot{\phi} = \frac{p_\phi - p_\psi \cos (\theta_0 + \varepsilon)}{I \sin^2 (\theta_0 + \varepsilon)}
\]
(887)

When expanded to first order, the trig functions are:
\[
\sin (\theta_0 + \varepsilon) = \sin \theta_0 \cos \varepsilon + \cos \theta_0 \sin \varepsilon 
\approx \sin \theta_0 + \varepsilon \cos \theta_0
\]
(888)
\[
\cos (\theta_0 + \varepsilon) = \cos \theta_0 \cos \varepsilon - \sin \theta_0 \sin \varepsilon 
\approx \cos \theta_0 - \varepsilon \sin \theta_0
\]
(890)

so to first order,
\[
0 = I \ddot{\varepsilon} - I \dot{\phi}^2 \left( \sin \theta_0 + \varepsilon \cos \theta_0 \right) \left( \cos \theta_0 - \varepsilon \sin \theta_0 \right) 
+ p_\psi \dot{\phi} \left( \sin \theta_0 + \varepsilon \cos \theta_0 \right) - mga \left( \sin \theta_0 + \varepsilon \cos \theta_0 \right)
= I \ddot{\varepsilon} - I \dot{\phi}^2 \left( \sin \theta_0 \cos \theta_0 - \varepsilon \sin^2 \theta_0 + \varepsilon \cos^2 \theta_0 \right) 
+ \left( p_\psi \dot{\phi} \sin \theta_0 + \varepsilon p_\psi \dot{\phi} \cos \theta_0 \right) - mga \sin \theta_0 - mga \varepsilon \cos \theta_0
= I \ddot{\varepsilon} + \varepsilon \left( I \dot{\phi}^2 \left( \sin^2 \theta_0 - \cos^2 \theta_0 \right) + p_\psi \dot{\phi} \cos \theta_0 - mga \cos \theta_0 \right)
- I \dot{\phi}^2 \sin \theta_0 \cos \theta_0 + p_\psi \dot{\phi} \sin \theta_0 - mga \sin \theta_0
\]
(892)

We must still expand \( \dot{\phi} \):
\[
\dot{\phi} = \frac{p_\phi - p_\psi \cos (\theta_0 + \varepsilon)}{I \sin^2 (\theta_0 + \varepsilon)}
\approx \frac{p_\phi - p_\psi \cos \theta_0 - \varepsilon \sin \theta_0}{I \left( \sin \theta_0 + \varepsilon \cos \theta_0 \right)^2}
\approx \frac{\left( p_\phi - p_\psi \cos \theta_0 \right) \left( 1 + \varepsilon \left( \frac{p_\psi \sin \theta_0}{p_\phi - p_\psi \cos \theta_0} \right) \right)}{I \sin^2 \theta_0 \left( 1 + \varepsilon \cot \theta_0 \right)^2}
\approx \frac{\left( p_\phi - p_\psi \cos \theta_0 \right) \left( 1 + \varepsilon \left( \frac{p_\psi \sin \theta_0}{p_\phi - p_\psi \cos \theta_0} \right) \right)}{I \sin^2 \theta_0 \left( 1 + \varepsilon \cot \theta_0 \right)^2} \left( 1 - 2 \varepsilon \cot \theta_0 \right)
\approx \dot{\phi}_0 + \varepsilon \left( \frac{\dot{\phi}_0 p_\psi \sin \theta_0}{p_\phi - p_\psi \cos \theta_0} - 2 \dot{\phi}_0 \cot \theta_0 \right)
= \dot{\phi}_0 + \varepsilon \left( \frac{p_\psi}{I \sin \theta_0} - \frac{2 I \dot{\phi}_0 \cos \theta_0}{I \sin \theta_0} \right)
\]

150
\[ = \dot{\phi}_0 + \varepsilon \left( \frac{p_\Psi - 2I\dot{\phi}_0 \cos \theta_0}{I \sin \theta_0} \right) \quad (893) \]

Now combining the two expansions, dropping terms of second order and higher

\[ 0 = I \ddot{\varepsilon} + \varepsilon \left( I \dot{\phi}_0^2 \left( \sin^2 \theta_0 - \cos^2 \theta_0 \right) + p_\Psi \dot{\phi}_0 \cos \theta_0 - mga \cos \theta_0 \right) \]

\[ -I \dot{\phi}_0^2 \left( 1 + 2\varepsilon \left( \frac{p_\Psi - 2I\dot{\phi}_0 \cos \theta_0}{\dot{\phi}_0 I \sin \theta_0} \right) \right) \sin \theta_0 \cos \theta_0 \]

\[ +p_\Psi \dot{\phi}_0 \left( 1 + 2\varepsilon \left( \frac{p_\Psi - 2I\dot{\phi}_0 \cos \theta_0}{\dot{\phi}_0 I \sin \theta_0} \right) \right) \sin \theta_0 - mga \sin \theta_0 \]

\[ = I \ddot{\varepsilon} + \varepsilon \left( I \dot{\phi}_0^2 \left( \sin^2 \theta_0 - \cos^2 \theta_0 \right) + p_\Psi \dot{\phi}_0 \cos \theta_0 - mga \cos \theta_0 \right) \]

\[ +\frac{2p_\Psi}{I} \varepsilon \left( p_\Psi - 2I\dot{\phi}_0 \cos \theta_0 \right) - 2\dot{\phi}_0 \varepsilon \left( p_\Psi - 2I\dot{\phi}_0 \cos \theta_0 \right) \cos \theta_0 \]

\[ -I \dot{\phi}_0^2 \sin \theta_0 \cos \theta_0 + p_\Psi \dot{\phi}_0 \sin \theta_0 - mga \sin \theta_0 \quad (894) \]

and since we have \( I \dot{\phi}_0^2 \sin \theta_0 \cos \theta_0 - p_\Psi \dot{\phi}_0 \sin \theta_0 + mga \sin \theta_0 = 0 \) from the unperturbed solution the equation for the perturbation takes the simple form

\[ 0 = I \ddot{\varepsilon} + k\varepsilon \quad (895) \]

where \( k \) takes the not-so-simple form

\[ k = I \dot{\phi}_0^2 \left( \sin^2 \theta_0 - \cos^2 \theta_0 \right) + p_\Psi \dot{\phi}_0 \cos \theta_0 - mga \cos \theta_0 \]

\[ +\frac{2p_\Psi}{I} \left( p_\Psi - 2I\dot{\phi}_0 \cos \theta_0 \right) - 3p_\Psi \dot{\phi}_0 \cos \theta_0 + 3I \dot{\phi}_0^2 \cos^2 \theta_0 \]

\[ = I \dot{\phi}_0^2 \sin^2 \theta_0 - mga \cos \theta_0 + \frac{2p_\Psi^2}{I} - 3p_\Psi \dot{\phi}_0 \cos \theta_0 + 3I \dot{\phi}_0^2 \cos^2 \theta_0 \]

\[ = \left( \frac{2p_\Psi^2}{I} + I \dot{\phi}_0^2 + 2I \dot{\phi}_0^2 \cos^2 \theta_0 \right) - \left( mga \cos \theta_0 + 3p_\Psi \dot{\phi}_0 \right) \cos \theta_0 \]

\[ = \left( \frac{2p_\Psi^2}{I} + I \dot{\phi}_0^2 + \left( 2I \dot{\phi}_0^2 - mga \right) \cos^2 \theta_0 \right) - 3p_\Psi \dot{\phi}_0 \cos \theta_0 \quad (896) \]

The solutions of the equation for \( \varepsilon \) are oscillatory if \( k \) is positive and exponential if \( k \) is negative. The oscillating case implies stability, while exponential behavior will destroy the initial state, indicating instability. Therefore,
let’s examine the sign of \( k \) in detail. From the equilibrium condition we have

\[
\cos \theta_0 = \frac{p_\psi \dot{\phi}_0 - mga}{I \dot{\phi}_0^2}
\]  

(897)

where we required

\[
\dot{\phi}_0 p_\psi > mga
\]  

(898)

Substituting the expression for the cosine, \( k \) becomes

\[
k = \left( \frac{2p_\psi^2}{I} + I \dot{\phi}_0^2 + \left( 2I \dot{\phi}_0^2 - mga \right) \left( \frac{p_\psi \dot{\phi}_0 - mga}{I \dot{\phi}_0^2} \right)^2 \right)
\]

\[
-3p_\psi \dot{\phi}_0 \left( \frac{p_\psi \dot{\phi}_0 - mga}{I \dot{\phi}_0^2} \right)
\]

\[
= \frac{2p_\psi^2}{I} + I \dot{\phi}_0^2 + \left( 2I \dot{\phi}_0^2 - mga \right) \left( \frac{p_\psi \dot{\phi}_0 - mga}{I \dot{\phi}_0^2} \right)^2
\]

\[
-3p_\psi \dot{\phi}_0 + \frac{3p_\psi mga}{I \dot{\phi}_0}
\]

\[
= I \dot{\phi}_0^2 - \frac{p_\psi^2}{I} + \left( 2I \dot{\phi}_0^2 - mga \right) \left( \frac{p_\psi \dot{\phi}_0 - mga}{I \dot{\phi}_0^2} \right)^2 + \frac{3p_\psi mga}{I \dot{\phi}_0}
\]

\[
\frac{k}{I \dot{\phi}_0^2} = 1 - \left( \frac{p_\psi \dot{\phi}_0}{I \dot{\phi}_0^2} \right)^2 + \left( 2 - \frac{mga}{I \dot{\phi}_0^2} \right) \left( \frac{p_\psi \dot{\phi}_0 - mga}{I \dot{\phi}_0^2} \right)^2
\]

\[
+3 \left( \frac{p_\psi \dot{\phi}_0}{I \dot{\phi}_0^2} \right) \left( \frac{mga}{I \dot{\phi}_0^2} \right)
\]

(899)

(900)

In the last line we have expressed \( k \) in terms of the positive definite ratios

\[
\alpha = \frac{p_\psi \dot{\phi}_0}{I \dot{\phi}_0^2}
\]

(901)

\[
\beta = \frac{mga}{I \dot{\phi}_0^2} < \alpha
\]

(902)

so we have

\[
k = I \dot{\phi}_0^2 \left( 1 - \alpha^2 + (2 - \beta) (\alpha - \beta)^2 + 3\alpha\beta \right)
\]
\[
\begin{align*}
&= I_{\phi_0}^2 (1 + \alpha^2 - \alpha^2 \beta - \alpha \beta + 2 \alpha \beta^2 + 2 \beta^2 - \beta^3) \\
&= I_{\phi_0}^2 (1 + \alpha^2 - \alpha^2 \beta - \alpha^2 \beta + \alpha \beta^2 + \beta^2 + \alpha \beta^2 - \beta^3) \\
&= I_{\phi_0}^2 (1 + \alpha (\alpha - \beta) - \alpha \beta (\alpha - \beta) + \beta^2 + \beta^2 (\alpha - \beta)) \\
&= I_{\phi_0}^2 (1 + (\alpha - \alpha \beta)(\alpha - \beta) + \beta^2 + \beta^2 (\alpha - \beta))
\end{align*}
\]

Try the ratios another way. Let $\gamma = \beta / \alpha < 1$. Then

\[
\begin{align*}
k &= I_{\phi_0}^2 (1 + \alpha^2 \left( (2 - \beta)(1 - \gamma)^2 - 1 + 3\gamma \right) ) \\
&= I_{\phi_0}^2 (1 + \alpha^2 \left( (2 - 4\gamma + 2\gamma^2 - \beta + 2\beta \gamma - \beta \gamma^2 - 1 + 3\gamma) \right) ) \\
&= I_{\phi_0}^2 (1 + \alpha^2 \left( ((1 - \gamma + 2\gamma^2) - \alpha (\gamma - 2\gamma^2 + \gamma^3)) \right) )
\end{align*}
\]

We can get an easy estimate by approximating

\[
\begin{align*}
k &= I_{\phi_0}^2 \left( (1 - \gamma + 2\gamma^2) - \alpha (\gamma - 2\gamma^2 + \gamma^3) \right) \\
&= I_{\phi_0}^2 \left( (1 - 2\gamma + \gamma^2 + \gamma + \gamma^2) - \alpha \gamma (1 - 2\gamma + \gamma^2) \right) \\
&> I_{\phi_0}^2 \left( (1 - \gamma)^2 - \alpha \gamma (1 - \gamma)^2 \right) \\
&= I_{\phi_0}^2 \left( 1 + \alpha^2 (1 - \gamma)^2 (1 - \alpha \gamma) \right)
\end{align*}
\]

Therefore, $k$ is positive at least when

\[
1 - \alpha \gamma > 0
\]

\[
\beta = \frac{mga}{I_{\phi_0}^2} = \alpha \gamma < 1
\]

\[
\frac{mga}{I_{\phi_0}^2}
\]

A complete answer is given by the condition

\[
\begin{align*}
1 + \alpha^2 (1 - \gamma + 2\gamma^2) &> \alpha^2 \beta (1 - \gamma)^2 \\
1 + \alpha^2 (1 - \gamma)^2 + \alpha \beta (1 + \gamma) &> \alpha^2 \beta (1 - \gamma)^2 \\
\frac{1}{\alpha^2} + (1 - \gamma)^2 + \gamma (1 + \gamma) &> \beta (1 - \gamma)^2 \\
\beta &< 1 + \frac{\gamma}{1 - \gamma} + \frac{1}{\alpha^2 (1 - \gamma)^2}
\end{align*}
\]

where each term on the right side is positive.
6.9 Euler’s equations

We have found the equations of motion in the laboratory frame:

\[ \mathbf{F}_{\text{ext}} = \frac{d\mathbf{P}}{dt}, \]

\[ \mathbf{K} = \frac{d\mathbf{M}}{dt}. \]

However, since the rigid body is moving, even if we choose our initial frame of reference so that the moment of inertia tensor is diagonal, it immediately becomes non-diagonal. Only in a frame fixed to the body can the inertia tensor remain diagonal throughout the motion. Since the form of the angular momentum \( \mathbf{M} \) depends on the moment of inertia tensor, the general problem of rigid body motion is much simpler if we transform the equations of motion to the moving frame.

The transformation to the moving frame is typically accomplished by replacing all time derivatives, \( \frac{d}{dt} \), in the center of mass frame, by the combined operator

\[ \frac{d}{dt} = \left( \frac{d}{dt} \right)' + \Omega \times \]

where \( \left( \frac{d}{dt} \right)' \) is the time derivative in the moving frame. This trick is possible only because the problem is a simple one. What is, in fact, being done is a time-dependent change of reference frame. We show this as follows.

Let an arbitrary vector \( \mathbf{A} \) be written in both frames of reference (the lab frame \( \mathbf{e}_i(0) \) and the moving frame, \( \mathbf{e}_i(t) \)). Remember that the vector \( \mathbf{A} \) is unchanged by using a different basis, but its components will be different. We therefore write

\[ \mathbf{A} = \tilde{A}_i \mathbf{e}_i(0) = A_i \mathbf{e}_i(t) \]

Furthermore, we know that the two frames are related by a rotation,

\[ \mathbf{e}_i(t) = \Lambda_{ij}(t) \mathbf{e}_j(0) \]

and we know that the rate of change of the moving frame with respect to the fixed frame is given by

\[ \frac{d\mathbf{e}_i(t)}{dt} = \varepsilon_{ijk} \mathbf{e}_j(t) \Omega_k \]
Therefore, if we take the time derivative of \( \mathbf{A} \), we have two expressions:

\[
\frac{d\mathbf{A}}{dt} = \frac{d}{dt} (\tilde{A}_i \mathbf{e}_i(0)) = \frac{d}{dt} (A_i \mathbf{e}_i(t)) \tag{912}
\]

Now, the fixed frame is static: \( \frac{d}{dt} (\mathbf{e}_i(0)) = 0 \). Therefore, regardless of how the components \( \tilde{A}_i \) depend on time, we may write

\[
\left( \frac{d\tilde{A}_i}{dt} \right) \mathbf{e}_i(0) = \frac{dA_i}{dt} \mathbf{e}_i(t) + \tilde{A}_i \frac{de_i(t)}{dt} \tag{913}
\]
\[
= \frac{dA_i}{dt} \mathbf{e}_i(t) + A_i \varepsilon_{ijk} e_j(t) \Omega_k \tag{914}
\]
\[
= \frac{dA_i}{dt} \mathbf{e}_i(t) + A_i \varepsilon_{kij} e_j(t) \Omega_k \tag{915}
\]
\[
= \mathbf{e}_i(t) \left( \frac{dA_i}{dt} + (\Omega \times \mathbf{A})_i \right) \tag{916}
\]

This is the correct relationship. Now the popular trick is to think of the left side of the equation as the time derivative of \( \mathbf{A} \) with respect to the fixed frame while thinking of the right hand side as the time derivative of \( \mathbf{A} \) with respect to the moving frame, by writing

\[
\frac{d}{dt} \mathbf{A} = \left( \frac{d\mathbf{A}}{dt} \right)' + \Omega \times \mathbf{A} \tag{917}
\]

But what we’re really doing is differentiating the different components of \( \mathbf{A} \) in the two frames. Doing the change of frames correctly reminds us that after completing our calculation in the moving frame, we may want to change back to the fixed frame by writing

\[
\left( \frac{d\tilde{A}_j}{dt} \right) \mathbf{e}_j(0) = \mathbf{e}_j(t) \left( \frac{dA_i}{dt} + (\Omega \times \mathbf{A})_i \right) \tag{918}
\]
\[
= \Lambda_{ij}(t) \mathbf{e}_j(0) \left( \frac{dA_i}{dt} + (\Omega \times \mathbf{A})_i \right) \tag{919}
\]

so that the components are related by:

\[
\frac{d\tilde{A}_j}{dt} = \Lambda_{ij}(t) \left( \frac{dA_i}{dt} + (\Omega \times \mathbf{A})_i \right) \tag{920}
\]
\[
= \Lambda_{ji}^{-1}(t) \left( \frac{dA_i}{dt} + (\Omega \times \mathbf{A})_i \right) \tag{921}
\]
Finding the equations of motion in the moving frame is now easy. For the linear momentum:

\[
\mathbf{F}_{\text{ext}} = \frac{d\mathbf{P}}{dt}
\]

\[
(F_{\text{ext}})_i \hat{e}_i(t) = \hat{e}_i(t) \left( \frac{dA_i}{dt} + (\Omega \times \mathbf{A})_i \right)
\]

\[
(F_{\text{ext}})_i = \frac{dP_i}{dt} + (\Omega \times \mathbf{P})_i
\]

where all components are in the moving frame. For the angular momentum,

\[
\mathbf{K} = \frac{d\mathbf{M}}{dt}
\]

\[
K_i = \frac{dM_i}{dt} + (\Omega \times \mathbf{M})_i
\]

Replacing \( M_i = I_{ij} \Omega_j \) and remembering that we choose the frame so that \( I_{ij} \) is diagonal, we have

\[
K_i = \frac{d(I_{ij} \Omega_j)}{dt} + \epsilon_{ijk} \Omega_j I_{kl} \Omega_l
\]

Because of the diagonal form of the moment of inertia tensor, this equation is simpler when written out in components. For example, setting \( i = 1 \) and performing the sums,

\[
K_1 = I_{1j} \frac{d\Omega_j}{dt} + \epsilon_{1jk} \Omega_j I_{kl} \Omega_l
\]

\[
= I_{11} \frac{d\Omega_1}{dt} + \epsilon_{123} \Omega_2 I_{3l} \Omega_l + \epsilon_{132} \Omega_3 I_{2l} \Omega_l
\]

\[
= I_1 \frac{d\Omega_1}{dt} + \Omega_2 I_{33} \Omega_3 - \Omega_3 I_{22} \Omega_2
\]

\[
= I_1 \frac{d\Omega_1}{dt} + (I_3 - I_2) \Omega_2 \Omega_3
\]

\[
K_i = \frac{d(I_{ij} \Omega_j)}{dt} + \epsilon_{ijk} \Omega_j I_{kl} \Omega_l
\]

The other two components are found the same way, giving

\[
K_1 = I_1 \frac{d\Omega_1}{dt} + (I_3 - I_2) \Omega_2 \Omega_3
\]
\[ K_2 = I_2 \frac{d\Omega_2}{dt} + (I_1 - I_3) \Omega_3 \Omega_1 \]
\[ K_3 = I_3 \frac{d\Omega_3}{dt} + (I_2 - I_1) \Omega_1 \Omega_2 \]

We consider two examples. First, suppose a constant torque \( K \) is applied about the \( z \) axis of a uniform sphere which is initially at rest. Then, since all three principal moments are equal, \( I_1 = I_2 = I_3 = I \), we have just

\[
K = I \frac{d\Omega_1}{dt} \\
0 = I \frac{d\Omega_2}{dt} \\
0 = I \frac{d\Omega_3}{dt}
\]

The solution is immediate:

\[
\Omega = \left( \Omega_{10} + \frac{K}{I} t, \Omega_{20}, \Omega_{30} \right) \\
= \left( \frac{K}{I} t, 0, 0 \right)
\]

Our second example is that of a free symmetric top. In this case we set \( I_1 = I_2 = I > I_3 \) and \( \mathbf{F}_{ext} = \mathbf{K} = 0 \). Then

\[
0 = I \frac{d\Omega_1}{dt} + (I_3 - I) \Omega_2 \Omega_3 \\
0 = I \frac{d\Omega_2}{dt} + (I - I_3) \Omega_3 \Omega_1 \\
0 = I_3 \frac{d\Omega_3}{dt}
\]

Since the third equation immediately implies

\[ \Omega_3 = \Omega_{30} = const. \]

we may replace \( \Omega_3 \) in the first two. Defining \( \beta = \frac{(I - I_3)\Omega_{30}}{I} \) we have

\[
0 = \frac{d\Omega_1}{dt} - \beta \Omega_2 \\
0 = \frac{d\Omega_2}{dt} + \beta \Omega_1
\]
We could multiply the first by $\Omega_1$ and the second by $\Omega_2$, then subtract. Integration then recovers the conservation of energy, and a second integration yields the full solution. However, for equations of this form it is easier to differentiate a second time:

\[
0 = \frac{d^2 \Omega_1}{dt^2} - \beta \frac{d \Omega_2}{dt} = \frac{d^2 \Omega_1}{dt^2} + \beta^2 \Omega_1
\]

We see immediately that $\Omega_1$ varies sinusoidally, and we find $\Omega_2$ from $\Omega_2 = \frac{1}{\beta} \frac{d \Omega_1}{dt}$.

6.10 The asymmetrical top

Read on your own.

6.11 Motion in a rotating frame of reference

The moving frame of reference $\mathbf{e}_i(t)$ is an example of a non-inertial frame. There is no reason not to use such frames if they are convenient, as they are in the case of rigid bodies. However, they do mean that Newton’s second law does not hold in its usual form. Consider the case of a system with its center of mass at rest, so that $\mathbf{e}_i(t)$ and $\mathbf{e}_i(0)$ remain fixed at the same point, but $\mathbf{e}_i(t)$ rotates with respect to $\mathbf{e}_i(0)$. We wish to describe the motion of a particle with respect to each of these systems. Since $\mathbf{e}_i(0)$ is taken to be an inertial frame, we know that Newton’s second law holds:

\[
\mathbf{F} = m \frac{d^2 \mathbf{x}}{dt^2}
\]

or more explicitly,

\[
\bar{F}_i \mathbf{e}_i(0) = m \frac{d^2 (\bar{x}_i \mathbf{e}_i(0))}{dt^2} = m \mathbf{e}_i(0) \frac{d^2 \bar{x}}{dt^2}
\]

But we may also expand $\mathbf{x}$ and $\mathbf{F}$ in the non-inertial frame. Let

\[
\mathbf{x} = \bar{x}_i \mathbf{e}_i(0) = x_i \mathbf{e}_i(t)
\]

\[
\mathbf{F} = \bar{F}_i \mathbf{e}_i(0) = F_i \mathbf{e}_i(t)
\]
Then
\[ F_i e_i(0) = m \frac{d^2(x_i e_i(t))}{dt^2} = m \frac{d}{dt} \left( \frac{dx_i}{dt} e_i(t) + x_i \frac{de_i(t)}{dt} \right) = m \frac{d}{dt} \left( \left( \frac{dx_i}{dt} + (\Omega \times x)_i \right) e_i(t) \right) \] (926)

where we have used
\[ \frac{de_i(t)}{dt} = \varepsilon_{ijk} e_j(t) \Omega_k \] (927)
as before. Continuing in the same way with the second time derivative,
\[ F_i e_i(t) = m \frac{d}{dt} \left( \left( \frac{dx_i}{dt} + (\Omega \times x)_i \right) e_i(t) \right) \]
\[ = m \left( \frac{d}{dt} \left( \frac{dx_i}{dt} + (\Omega \times x)_i \right) \right) e_i(t) + \left( \Omega \times \left( \frac{dx}{dt} + \Omega \times x \right) \right)_i e_i(t) \]
\[ = m \left( \frac{d^2 x_i}{dt^2} + \left( \Omega \times \frac{dx}{dt} \right)_i + \left( \frac{d\Omega}{dt} \times x \right)_i \right) e_i(t) \]
\[ + m \left( \Omega \times \frac{dx}{dt} \right)_i + \Omega \times (\Omega \times x)_i \right) e_i(t) \] (928)
or simply,
\[ \mathbf{F} = m \left( \frac{d^2 x}{dt^2} + 2\Omega \times \mathbf{v} + \dot{\Omega} \times \mathbf{x} + \Omega \times (\Omega \times \mathbf{x}) \right) \] (929)

Where we must remember that all quantities are now referred to the moving frame. Clearly, Newton's law does not hold in this frame! We can rearrange this to give the acceleration
\[ m \frac{d^2 \mathbf{x}}{dt^2} = \mathbf{F} + 2mv \times \Omega + mx \times \dot{\Omega} - m\Omega \times (\Omega \times \mathbf{x}) \] (930)

This is the Coriolis theorem. Generically, the extra terms that appear on the right are called apparent forces. The final term is, of course, the usual centrifugal force. The term \(2mv \times \Omega\) is called the Coriolis force.
7 The Canonical Equations

Perhaps the most beautiful formulation of classical mechanics, and the one which ties most closely to quantum mechanics, is the canonical formulation. In this approach, the position and velocity variables of Lagrangian mechanics are replaced by the position and conjugate momentum, \( p_i \equiv \frac{\partial L}{\partial \dot{q}_i} \). It turns out that by doing this the coordinates and momenta are put on an equal footing, giving the equations of motion a much larger symmetry.

To make the change of variables, we use a Legendre transformation. This may be familiar from thermodynamics, where the internal energy, Gibb’s energy, free energy and enthalpy are related to one another by making different choices of the independent variables. Thus, for example, if we begin with

\[
dU = TdS - PdV
\]

where \( T \) and \( P \) are regarded as functions of \( S \) and \( V \), we can set

\[
H = U + VP
\]

and compute

\[
dH = dU + PdV + VdP
\]

\[
= TdS - PdV + PdV + VdP
\]

\[
= TdS + VdP
\]

(933)

to achieve a formulation in which \( T \) and \( V \) are treated as functions of \( S \) and \( P \).

The same technique works here. We have the Lagrangian, \( L(q_i, \dot{q}_i) \) and wish to find a function \( H(q_i, p_i) \). The differential of \( L \) is

\[
dL = \sum_{i=1}^{N} \frac{\partial L}{\partial q_i} dq_i + \sum_{i=1}^{N} \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i
\]

\[
= \sum_{i=1}^{N} \dot{p}_idq_i + \sum_{i=1}^{N} p_i d\dot{q}_i
\]

(934)

where the second line follows by using the equations of motion and the definition of the conjugate momentum. Therefore, set

\[
H(q_i, p_i) = \sum_{i=1}^{N} p_i \dot{q}_i - L
\]

(935)
so that

\[ dH = \sum_{i=1}^{N} dp_i \dot{q}_i + \sum_{i=1}^{N} p_i dq_i - dL \]

\[ = \sum_{i=1}^{N} dp_i \dot{q}_i + \sum_{i=1}^{N} p_i \dot{q}_i - \sum_{i=1}^{N} \dot{p}_i dq_i - \sum_{i=1}^{N} p_i dq_i \]

\[ = \sum_{i=1}^{N} \dot{q}_i dp_i - \sum_{i=1}^{N} \dot{p}_i dq_i \]

(936)

As expected from the Legendre transformation, \( q_i \) and \( p_i \) are now the independent variables.

Notice that, as it happens, \( H \) is of the same form as the energy. When the Lagrangian is independent of time, the Hamiltonian is equal to the conserved energy. But even if the Lagrangian depends on time so that there is no conserved energy, the Hamiltonian is still given by eq.(935).

Clearly, \( H \) is a function of the momenta. To see that we have really eliminated the dependence on velocity we may compute directly,

\[ \frac{\partial H}{\partial \dot{q}_j} = \frac{\partial}{\partial \dot{q}_j} \left( \sum_{i=1}^{N} p_i \dot{q}_i - L(q_i, \dot{q}_i) \right) \]

\[ = \sum_{i=1}^{N} p_i \delta_{ij} - \frac{\partial L}{\partial \dot{q}_j} \]

\[ = p_j - \frac{\partial L}{\partial \dot{q}_j} \]

\[ = 0 \]

(937)

Hamilton’s equations of motions now follow immediately. Notice that the Euler-Lagrange equations are already build into eq.(936) for \( dH \). Since the differential of \( H \) may always be written as

\[ dH = \sum_{i=1}^{N} \frac{\partial H}{\partial q_j} dq_i + \sum_{i=1}^{N} \frac{\partial H}{\partial p_j} dp_i \]

(938)

we can simply equate the two expressions:

\[ dH = \sum_{i=1}^{N} dp_i \dot{q}_i - \sum_{i=1}^{N} \dot{p}_i dq_i = \sum_{i=1}^{N} \frac{\partial H}{\partial q_i} dq_i + \sum_{i=1}^{N} \frac{\partial H}{\partial p_i} dp_i \]

(939)
Since the differentials $dq_i$ and $dp_i$ are all independent, we can equate their coefficients,

\[
\dot{p}_i = -\frac{\partial H}{\partial q_i} \\
\dot{q}_i = \frac{\partial H}{\partial p_j}
\]  

(940)

These are Hamilton’s equations.

Notice that we assumed that the sum in the definition of $H$ runs over all coordinates, $i = 1, 2, \ldots, N$. However, there is no reason that we can’t restrict the sum to fewer coordinates. This substitutes momenta for only some of the velocities, so that some degrees or freedom are described by Hamilton’s equations while the remainder still satisfy the Euler-Lagrange equations. The function, $R(q_i, p_o, \dot{q}_m)$, produced in this way is called the Routhian.

This shows that Hamilton’s equations follow from the Euler-Lagrange equations. Conversely, observe that using Hamilton’s equations and our definitions of momentum and the Hamiltonian, we have

\[
\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = \dot{p}_i \\
= -\frac{\partial H}{\partial q_i} \\
= -\frac{\partial}{\partial q_i} \left( \sum_{i=1}^{N} p_i \dot{q}_i - L \right) \\
= \frac{\partial L}{\partial q_i}
\]  

(941)

The Euler-Lagrange equations therefore follow from Hamilton’s equations.
7.1 Phase space and the symplectic form

Our next goal is to confirm the advantages we claimed for the Hamiltonian formulation. We expect more symmetry in the new system, and we would like to see any further advantages the system has.

To begin, some definitions. By configuration space, we mean the space of all possible positions of the particles comprising the system, or the complete set of possible values of the degrees of freedom of the problem. Thus, configuration space is the $N$-dimensional space of all values of $q_i$. By momentum space, we mean the $N$-dimensional space of all possible values of all of the conjugate momenta.

Finally, we introduce phase space - the $2N$-dimensional space of all possible values of both position and momentum. Let

$$\xi^A = (q^i, p_j)$$

where $A = 1, \ldots, 2N$. Then the $2N$ variables $\xi^A$ provide a set of coordinates for phase space. We would like to write Hamilton’s equations in terms of these, thereby treating all $2N$ directions on an equal footing.

For example, the phase space of a one dimensional simple harmonic oscillator is the two dimensional space spanned by $x$ and $p$. The solutions for the motion of the oscillator are curves in this space. Thus, since the motion is given by

$$x = A \cos(\omega t + \varphi)$$

$$p = -m\omega A \sin(\omega t + \varphi)$$

the curves satisfy

$$x^2 + \frac{p^2}{m^2\omega^2} = A^2$$

and therefore are ellipses in the $xp$ plane. The choice of initial phase $\varphi$ does not change the ellipse, but changing $A$ changes the size of the ellipse. Ranging over all values of $A$, we find a nested series of ellipses filling the $xp$ plane.

Certain conserved quantities are evident in this example. For example, the energy

$$E = \frac{1}{2}m\omega^2 x^2 + \frac{p^2}{2m} = \frac{1}{2}m\omega^2 A^2$$

is immediately seen to be constant. We also have a trivial example of the Liouville theorem which we will prove below. The theorem states that
volumes in phase space are preserved by the classical motion. By a volume, we mean a subset of initial conditions occupying a definite volume in phase space. This subset of initial conditions represents an ensemble of identical physical systems started with a range of possible initial conditions. As we trace the motion of such an ensemble in phase space, one system at a time, we see that the region occupied moves smoothly about an elliptical path without distorting. It is therefore obvious that the area in the $xp$ plane occupied by the ensemble is the same at any time. Even for more complicated systems in which the shape of the region changes radically with time, this total area is preserved.

Returning to write Hamilton’s equations in terms of $\xi^A$,
\[
\frac{d\xi^A}{dt} = \left( \begin{array}{c} \dot{q}^i \\ \dot{p}_j \end{array} \right) = \left( \begin{array}{c} \frac{\partial H}{\partial p_j} \\ -\frac{\partial H}{\partial q^i} \end{array} \right) = \Omega^{AB} \frac{\partial H}{\partial \xi^B}
\]
where the existence of $\Omega^{AB}$ in the last step follows because of the linearity of the right hand side in the partial derivatives of $H$. It is easy to see that $\Omega^{AB}$ is given by
\[
\Omega^{AB} = \left( \begin{array}{cc} 0 & \delta^i_j \\ -\delta^j_i & 0 \end{array} \right)
\]
Note that $\Omega^{AB}$ is antisymmetric. Hamilton’s equations are now
\[
\frac{d\xi^A}{dt} = \Omega^{AB} \frac{\partial H}{\partial \xi^B}
\]  
(942)

The tensor $\Omega^{AB}$ is called the symplectic tensor, and its inverse, written as $\Omega_{AB}$ is called the symplectic form. The essential structure of Hamilton dynamics is embodied in this object. To see why, consider what happens to Hamilton’s equations if we want to change to a new set of coordinates, $\chi^A = \chi^A (\xi)$, Let the inverse transformation be $\xi^A (\chi)$. The time derivatives become
\[
\frac{d\xi^A}{dt} = \frac{\partial \xi^A}{\partial \chi^B} \frac{d\chi^B}{dt}
\]

164
while the right side becomes

\[
\Omega_{AB} \frac{\partial H}{\partial \xi^B} = \Omega_{AB} \frac{\partial \chi^C}{\partial \xi^B} \frac{\partial H}{\partial \chi^C}
\]

Equating these expressions,

\[
\frac{\partial \xi^A}{\partial \chi^B} \frac{d\chi^B}{dt} = \Omega_{AB} \frac{\partial \chi^D}{\partial \xi^B} \frac{\partial H}{\partial \chi^D}
\]

we multiply by the Jacobian matrix, \( \frac{\partial \xi^A}{\partial \chi^B} \) to get

\[
\frac{\partial \chi^C}{\partial \xi^A} \frac{\partial \xi^A}{\partial \chi^B} \frac{d\chi^B}{dt} = \frac{\partial \chi^C}{\partial \xi^A} \Omega_{AB} \frac{\partial \chi^D}{\partial \xi^B} \frac{\partial H}{\partial \chi^D}
\]

and finally

\[
\frac{d\chi^C}{dt} = \frac{\partial \chi^C}{\partial \xi^A} \Omega_{AB} \frac{\partial \chi^D}{\partial \xi^B} \frac{\partial H}{\partial \chi^D}
\]

Defining the symplectic form in the new coordinate system,

\[
\tilde{\Omega}^{CD} = \frac{\partial \chi^C}{\partial \xi^A} \Omega_{AB} \frac{\partial \chi^D}{\partial \xi^B}
\]

we see that Hamilton’s equations are entirely the same if the transformation leaves the symplectic form invariant,

\[
\tilde{\Omega}^{CD} = \Omega^{CD}
\]

Any linear transformation \( M^A \xi^B \) leaving the symplectic form invariant,

\[
\Omega^{AB} \equiv M^A \xi^B \frac{d\Omega^{CD}}{d\xi^D}
\]

is called a symplectic transformation. Coordinate transformations which are symplectic transformations at each point are called canonical. Therefore those functions \( \chi^A (\xi) \) satisfying

\[
\Omega^{CD} \equiv \frac{\partial \chi^C}{\partial \xi^A} \Omega_{AB} \frac{\partial \chi^D}{\partial \xi^B}
\]

are canonical transformations. Canonical transformations preserve Hamilton’s equations.
7.2 Poisson brackets

The symplectic tensor allows us to define a product on functions on phase space. The importance of this product, called the Poisson bracket, is that it too is preserved by canonical transformations.

Let $\xi^A$ be any set of phase space coordinates in which Hamilton’s equations take the form of eq.(942), and let $f$ and $g$ be any two functions of these phase space coordinates, $\xi^A$. Such coordinates are called canonical, while functions of canonical variables are called dynamical variables. Then we define the Poisson bracket,

$$\{f, g\} = \Omega^{AB} \frac{\partial f}{\partial \xi^A} \frac{\partial g}{\partial \xi^B}$$

In a different set of coordinates, $\chi^A(\xi)$, we have

$$\{f, g\}' = \Omega^{AB} \frac{\partial f}{\partial \chi^A} \frac{\partial g}{\partial \chi^B}$$

$$= \Omega^{AB} \left( \frac{\partial \xi^C}{\partial \chi^A} \frac{\partial f}{\partial \xi^C} \right) \left( \frac{\partial \xi^D}{\partial \chi^B} \frac{\partial g}{\partial \xi^D} \right)$$

$$= \left( \frac{\partial \xi^C}{\partial \chi^A} \Omega^{CD} \frac{\partial \xi^D}{\partial \chi^B} \right) \frac{\partial f}{\partial \xi^C} \frac{\partial g}{\partial \xi^D}$$

Therefore, if the coordinate transformation is canonical so that

$$\frac{\partial \xi^C}{\partial \chi^A} \Omega^{AB} \frac{\partial \xi^D}{\partial \chi^B} = \Omega^{CD}$$

then we have

$$\{f, g\}' = \Omega^{AB} \frac{\partial f}{\partial \xi^A} \frac{\partial g}{\partial \xi^B} = \{f, g\}$$

and the Poisson bracket is unchanged. We conclude that canonical transformations preserve all Poisson brackets.

An important special case of the Poisson bracket occurs when one of the functions is the Hamiltonian. In that case, we have

$$\{f, H\} = \Omega^{AB} \frac{\partial f}{\partial \xi^A} \frac{\partial H}{\partial \xi^B}$$

$$= \frac{\partial f}{\partial x^i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p^i} \frac{\partial H}{\partial x_i}$$
\[
\frac{df}{dt} = \frac{\partial f}{\partial x^i} \frac{dx^i}{dt} - \frac{\partial f}{\partial p_i} \left( - \frac{dp_i}{dt} \right)
\]

or simply,
\[
\frac{df}{dt} = \{ f, H \} + \frac{\partial f}{\partial t}
\]

This shows that as the system evolves classically, the total time rate of change of any dynamical variable is the sum of the Poisson bracket with the Hamiltonian and the partial time derivative. If a dynamical variable has no explicit time dependence, then \( \frac{\partial f}{\partial t} = 0 \) and the total time derivative is just the Poisson bracket with the Hamiltonian.

The coordinates now provide a special case. Since neither \( x^i \) nor \( p_i \) has any explicit time dependence, with have
\[
\begin{align*}
\frac{dx^i}{dt} &= \{ H, x^i \} \\
\frac{dp_i}{dt} &= \{ H, p_i \}
\end{align*}
\]

and we can check this directly:
\[
\begin{align*}
\frac{dq_i}{dt} &= \{ H, x^i \} \\
&= \sum_{j=1}^{N} \left( \frac{\partial x^i}{\partial x^j} \frac{\partial H}{\partial p_j} - \frac{\partial x^i}{\partial p_j} \frac{\partial H}{\partial x^j} \right) \\
&= \sum_{j=1}^{N} \delta_{ij} \frac{\partial H}{\partial p_j} \\
&= \frac{\partial H}{\partial p_i}
\end{align*}
\]

and
\[
\begin{align*}
\frac{dp_i}{dt} &= \{ H, p_i \} \\
&= \sum_{j=1}^{N} \left( \frac{\partial p_i}{\partial q_j} \frac{\partial H}{\partial p_j} - \frac{\partial p_i}{\partial p_j} \frac{\partial H}{\partial q_j} \right) \\
&= - \frac{\partial H}{\partial q_i}
\end{align*}
\]

167
Notice that since $q_i, p_i$ and are all independent, and do not depend explicitly on time, $\frac{\partial q_i}{\partial p_j} = \frac{\partial p_i}{\partial q_j} = 0 = \frac{\partial q_i}{\partial t} = \frac{\partial p_i}{\partial t}$.

Finally, we define the fundamental Poisson brackets. Suppose $x^i$ and $p_j$ are a set of coordinates on phase space such that Hamilton’s equations hold in the either the form of eqs.(943) or of eqs.(940). Since they themselves are functions of $(x^m, p_n)$ they are dynamical variables and we may compute their Poisson brackets with one another. With $\xi^A = (x^m, p_n)$ we have

$$\{ x^i, x^j \}_\xi = \Omega^{AB} \frac{\partial x^i}{\partial \xi^A} \frac{\partial x^j}{\partial \xi^B}$$

$$= \sum_{m=1}^{N} \left( \frac{\partial x^i}{\partial x^m} \frac{\partial x^j}{\partial p_m} - \frac{\partial x^i}{\partial p_m} \frac{\partial x^j}{\partial x^m} \right)$$

$$= 0$$

for $x^i$ with $x^j$,

$$\{ x^i, p_j \}_\xi = \Omega^{AB} \frac{\partial x^i}{\partial \xi^A} \frac{\partial p_j}{\partial \xi^B}$$

$$= \sum_{m=1}^{N} \left( \frac{\partial x^i}{\partial x^m} \frac{\partial p_j}{\partial p_m} - \frac{\partial x^i}{\partial p_m} \frac{\partial p_j}{\partial x^m} \right)$$

$$= \sum_{m=1}^{N} \delta^i_m \delta^m_j$$

$$= \delta^i_j$$

for $x^i$ with $p_j$ and finally

$$\{ p_i, p_j \}_\xi = \Omega^{AB} \frac{\partial p_i}{\partial \xi^A} \frac{\partial p_j}{\partial \xi^B}$$

$$= \sum_{m=1}^{N} \left( \frac{\partial p_i}{\partial x^m} \frac{\partial p_j}{\partial p_m} - \frac{\partial p_i}{\partial p_m} \frac{\partial p_j}{\partial x^m} \right)$$

$$= 0$$

for $p_i$ with $p_j$. The subscript $\xi$ on the bracket indicates that the partial derivatives are taken with respect to the coordinates $\xi^A = (x^i, p_j)$. We summarize these relations as

$$\{ \xi^A, \xi^B \}_\xi = \Omega^{AB}$$

168
We end this subsection with a theorem: Let the coordinates $\xi^A$ be canonical. Then a transformation $\chi^A(\xi)$ is canonical if and only if it satisfies the fundamental bracket relation

$$\{\chi^A, \chi^B\}_\xi = \Omega^{AB}$$

But the bracket on the left is defined by

$$\{\chi^A, \chi^B\}_\xi = \Omega^{CD} \frac{\partial \chi^A}{\partial \xi^C} \frac{\partial \chi^B}{\partial \xi^D}$$

so in order for $\chi^A$ to satisfy the canonical bracket we must have

$$\Omega^{CD} \frac{\partial \chi^A}{\partial \xi^C} \frac{\partial \chi^B}{\partial \xi^D} = \Omega^{AB}$$

which is just the condition shown above for a coordinate transformation to be canonical. Conversely, suppose the transformation $\chi^A(\xi)$ is canonical and $\{\xi^A, \xi^B\}_\xi = \Omega^{AB}$. Then eq. (944) holds and we have

$$\{\chi^A, \chi^B\}_\xi = \Omega^{CD} \frac{\partial \chi^A}{\partial \xi^C} \frac{\partial \chi^B}{\partial \xi^D} = \Omega^{AB}$$

so $\chi^A$ satisfies the fundamental bracket relation.

In summary, each of the following statements is equivalent:

1. $\chi^A(\xi)$ is a canonical transformation.

2. $\chi^A(\xi)$ is a coordinate transformation of phase space that preserves Hamilton’s equations.

3. $\chi^A(\xi)$ preserves the symplectic form, according to

$$\Omega^{AB} \frac{\partial \xi^C}{\partial \chi^A} \frac{\partial \xi^D}{\partial \chi^B} = \Omega^{CD}$$

4. $\chi^A(\xi)$ satisfies the fundamental bracket relations

$$\{\chi^A, \chi^B\}_\xi = \Omega^{CD} \frac{\partial \chi^A}{\partial \xi^C} \frac{\partial \chi^B}{\partial \xi^D}$$

We end the section with three examples of canonical transformations.
7.2.1 Example 1: Coordinate transformations

Let the new configuration space variable, $q^i$, be and an arbitrary function of the spatial coordinates:

$$q^i = q^i(x^j)$$

and let $\pi_j$ be the momentum variables corresponding to $q^i$. Then $(q^i, \pi_j)$ satisfy the fundamental Poisson bracket relations iff:

$$\{q^i, q^j\}_{x,p} = 0$$
$$\{q^i, \pi_j\}_{x,p} = \delta^i_j$$
$$\{\pi_i, \pi_j\}_{x,p} = 0$$

Check each:

$$\{q^i, q^j\}_{x,p} = \sum_{m=1}^{N} \left( \frac{\partial q^i}{\partial x^m} \frac{\partial q^j}{\partial p_m} - \frac{\partial q^i}{\partial p_m} \frac{\partial q^j}{\partial x^m} \right)$$
$$= 0$$

since $\frac{\partial q^j}{\partial p_m} = 0$. For the second bracket,

$$\delta^i_j = \{q^i, \pi_j\}_{x,p}$$
$$= \sum_{m=1}^{N} \left( \frac{\partial q^i}{\partial x^m} \frac{\partial \pi_j}{\partial p_m} - \frac{\partial q^i}{\partial p_m} \frac{\partial \pi_j}{\partial x^m} \right)$$
$$= \sum_{m=1}^{N} \frac{\partial q^i}{\partial x^m} \frac{\partial \pi_j}{\partial p_m}$$

Since $q^i$ is independent of $p_m$, we can satisfy this only if

$$\frac{\partial \pi_j}{\partial p_m} = \frac{\partial x^m}{\partial q^i}$$

Integrating gives

$$\pi_j = \frac{\partial x^n}{\partial q^i} p_n + c_j$$

with $c_j$ an arbitrary constant. The presence of $c_i$ does not affect the value of the Poisson bracket. Choosing $c_j = 0$, we compute the final bracket:

$$\{\pi_i, \pi_j\}_{x,p} = \sum_{m=1}^{N} \left( \frac{\partial \pi_i}{\partial x^m} \frac{\partial \pi_j}{\partial p_m} - \frac{\partial \pi_i}{\partial p_m} \frac{\partial \pi_j}{\partial x^m} \right)$$

170
Therefore, the transformations
\[ q^j = q^j(x^i) \]
\[ \pi_j = \frac{\partial x^m}{\partial q^j} p_n + c_j \]
is a canonical transformation for any functions \( q^i(x) \). This means that the symmetry group of Hamilton's equations is at least as big as the symmetry group of the Euler-Lagrange equations.

### 7.2.2 Example 2: Interchange of \( x \) and \( p \).

The transformation
\[ q^i = p_i \]
\[ \pi_i = -x^i \]
is canonical. We easily check the fundamental brackets:
\[ \{ q^i, q^j \}_{x,p} = \{ p_i, p_j \}_{x,p} = 0 \]
\[ \{ q^i, \pi_j \}_{x,p} = \{ p_i, -x^j \}_{x,p} = - \{ p_i, x^j \}_{x,p} = + \{ x^j, p_i \}_{x,p} = \delta^i_j \]
\[ \{ \pi_i, \pi_j \}_{x,p} = \{-x^i, -x^j \}_{x,p} = 0 \]
Interchange of \( x^i \) and \( p_j \), with a sign, is therefore canonical. The use of generalized coordinates does not include such a possibility, so Hamiltonian dynamics has a larger symmetry group than Lagrangian dynamics.
For our last example, we first show that the composition of two canonical transformations is also canonical. Let \( \psi (\chi) \) and \( \chi (\xi) \) both be canonical. Defining the composition transformation, \( \psi (\xi) = \psi (\chi (\xi)) \), we compute

\[
\Omega^{CD} \frac{\partial \psi^A}{\partial \xi^C} \frac{\partial \psi^B}{\partial \xi^D} = \Omega^{CD} \left( \frac{\partial \psi^A}{\partial \chi^E} \frac{\partial \chi^E}{\partial \xi^C} \right) \left( \frac{\partial \psi^B}{\partial \chi^F} \frac{\partial \chi^F}{\partial \xi^D} \right)
\]

\[
= \frac{\partial \chi^E}{\partial \xi^C} \frac{\partial \chi^F}{\partial \xi^D} \Omega^{CD} \left( \frac{\partial \psi^A}{\partial \chi^E} \right) \left( \frac{\partial \psi^B}{\partial \chi^F} \right)
\]

\[
= \Omega^{EF} \left( \frac{\partial \psi^A}{\partial \chi^E} \right) \left( \frac{\partial \psi^B}{\partial \chi^F} \right)
\]

so that \( \psi (\chi (\xi)) \) is canonical.

**7.2.3 Example 3: Momentum transformations**

By the previous results, the composition of an arbitrary coordinate change with \( x, p \) interchanges is canonical. Consider the effect of composing (a) an interchange, (b) a coordinate transformation, and (c) an interchange.

For (a), let

\[
q_i^1 = p_i \quad \pi_i^1 = -x_i
\]

Then for (b) we choose an arbitrary function of \( q_i^1 \):

\[
Q^i = Q^i (q_i^1) = Q^i (p_j)
\]

\[
P_i = \frac{\partial q_i^1}{\partial Q^i} \pi_n = -\frac{\partial p_n}{\partial Q^i} x^n
\]

Finally, for (c), another interchange:

\[
q^i = P_i = -\frac{\partial p_n}{\partial Q^i} x^n
\]

\[
\pi_i = -Q^i = -Q^i (p_j)
\]

This establishes that replacing the momenta by any three independent functions of the momenta, preserves Hamilton’s equations.

We next find a way to characterize general canonical transformations.
7.3 Generating functions for canonical transformations

Working with the Hamiltonian formulation of classical mechanics, we are allowed more transformations of the variables than with the Newtonian or Lagrangian formulations. We are now free to redefine our coordinates according to

\[ q_i = q_i(x_i, p_i, t) \]  
\[ \pi_i = \pi_i(x_i, p_i, t) \]

as long as the transformation is canonical. We now find general conditions under which Hamilton’s equations are preserved. Given a system described by a Hamiltonian \( H(x, p) \), we seek a related Hamiltonian \( H'(q, \pi) \) such that the equations of motion have the same form, namely

\[ \frac{dx_i}{dt} = \frac{\partial H}{\partial p_i} \]  
\[ \frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i} \]

in the original system and

\[ \frac{dq_i}{dt} = \frac{\partial H'}{\partial \pi_i} \]  
\[ \frac{d\pi_i}{dt} = -\frac{\partial H'}{\partial q_i} \]

in the transformed variables.

We start with the principle of least action with the action in the form

\[ S = \int (p_i dx_i - H dt) \]

There will also be an action of the form

\[ S = \int (\pi_i dq_i - H' dt) \]

The integrands must differ by at most an overall constant multiple (\( \lambda \)) and the addition of a total differential, \( df = \frac{df}{dt} dt \), so

\[ p_i dx_i - H dt = \lambda \pi_i dq_i - \lambda H' dt + df \]
Any transformation satisfying this equation with $\lambda = 1$ is a canonical transformation. We can write this expression as a solution for $df$:

$$df = p_i dx_i - \lambda \pi_i dq_i + (\lambda H' - H) dt$$  \hfill (954)

From this we see that $f$ may depend independently on $x_i$, $q_i$, and $t$. Then since

$$df = \frac{\partial f}{\partial x_i} dx_i + \frac{\partial f}{\partial q_i} dq_i + \frac{\partial f}{\partial t} dt$$  \hfill (955)

we can match up terms to require

\begin{align*}
p_i &= \frac{\partial f}{\partial x_i} \quad \hfill (956) \\
\pi_i &= -\frac{1}{\lambda} \frac{\partial f}{\partial q_i} \quad \hfill (957) \\
H' &= \frac{1}{\lambda} \left( H + \frac{\partial f}{\partial t} \right) \quad \hfill (958)
\end{align*}

The first equation

$$p_i = \frac{\partial f(x_i, q_i, t)}{\partial x_i}$$  \hfill (959)

gives $q_i$ implicitly in terms of the original variables, while the second determines $\pi_i$. Notice that once we pick a function $q_i = q_i(p_i, x_i, t)$, the form of $\pi_i$ is fixed. The third equation gives the new Hamiltonian in terms of the old one.

While eqs. (956-958) are useful as they stand, some insight is gained by rewriting the results in a less implicit form. Suppose we wish let the new position coordinates be arbitrary functions of the old phase space coordinates:

$$q_i = g_i(x_j, p_j, t)$$

Then at each position $x_i$ and time $t$ we can invert for $p_i$:

$$p_i = g_i^{-1}(q_j, x_j, t)$$

Setting

$$f(q_j, x_j, t) = \int g_i^{-1}(q_j, x_j, t) \, dx_i$$

we satisfy the first requirement of a generating function:
\[ p_i = \frac{\partial f(x_i, q_i, t)}{\partial x_i} \]  

Then the new momentum and Hamiltonian follow as

\[ \pi_i = -\frac{\partial f}{\partial q_i} \]  
\[ H' = H + \frac{\partial f}{\partial t} \]

We see from this that we can replace the position coordinates by arbitrary functions of all of the phase space coordinates. The form of the conjugate momentum is then determined. This result demonstrates that there are far more canonical transformations, \( q_i = g_i(x_j, p_j, t) \), than simple coordinate transformations, \( q^i(x^j) \).

Sometimes it is more convenient to specify the new momentum

\[ \pi_i = \pi_i(p_i, x_i, t) \]

than the new coordinates \( q_i = q_i(p_i, x_i, t) \). A Legendre transformation accomplishes this. Just replace \( f = g - \lambda \pi_i q_i \). Then

\[ df = dg - d\pi_i q_i - \pi_i dq_i = p_i dx_i - \lambda \pi_i dq_i + (\lambda H' - H) \ dt \]  
\[ dg = p_i dx_i + \lambda q_i d\pi_i + (\lambda H' - H) \ dt \]

and we see that \( g = g(x_i, \pi_i, t) \). In this case, \( g \) satisfies

\[ p_i = \frac{\partial g}{\partial x_i} \]  
\[ q_i = \frac{1}{\lambda} \frac{\partial g}{\partial \pi_i} \]  
\[ H' = \frac{1}{\lambda} \left( H + \frac{\partial g}{\partial t} \right) \]

Since canonical transformations can interchange or mix up the roles of \( x \) and \( p \), they are called canonically conjugate. Within Hamilton’s framework, position and momentum lose their independent meaning, except that variables always come in conjugate pairs.
7.4 Properties of Poisson brackets

Let’s look at some general properties of the Poisson brackets.

Bracketing with a constant always gives zero

\[ \{ f, c \} = 0 \] (966)

The Poisson bracket is \textit{linear}

\[ \{ af_1 + bf_2, g \} = a \{ f_1, g \} + b \{ f_2, g \} \] (967)

and \textit{Leibnitz}

\[
\{ f_1 f_2, g \} = \Omega^{CD} \frac{\partial (f_1 f_2)}{\partial \xi^C} \frac{\partial g}{\partial \xi^D} \\
= \Omega^{CD} \left( f_2 \frac{\partial f_1}{\partial \xi^C} + f_1 \frac{\partial f_2}{\partial \xi^C} \right) \frac{\partial g}{\partial \xi^D} \\
= f_2 \Omega^{CD} \frac{\partial f_1}{\partial \xi^C} \frac{\partial g}{\partial \xi^D} + f_1 \Omega^{CD} \frac{\partial f_2}{\partial \xi^C} \frac{\partial g}{\partial \xi^D} \\
= \{ f_1, g \} f_2 + f_1 \{ f_2, g \} 
\] (968) (969) (970) (971)

These three properties are the defining properties of a \textit{derivation}, which is the formal generalization of differentiation. The action of the Poisson bracket with any given function \( f \) on the class of all functions, \([ f, \cdot ]\) is therefore a derivation.

If we take the time derivative of a bracket, we find

\[
\frac{\partial}{\partial t} \{ f, g \} = \Omega^{CD} \frac{\partial^2 f}{\partial t \partial \xi^C} \frac{\partial g}{\partial \xi^D} + \Omega^{CD} \frac{\partial f}{\partial \xi^C} \frac{\partial^2 g}{\partial t \partial \xi^D} \\
= \Omega^{CD} \frac{\partial}{\partial \xi^C} \left( \frac{\partial f}{\partial t} \right) \frac{\partial g}{\partial \xi^D} + \Omega^{CD} \frac{\partial f}{\partial \xi^C} \frac{\partial}{\partial \xi^D} \left( \frac{\partial g}{\partial t} \right) \\
= \left\{ \frac{\partial f}{\partial t}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} \right\} 
\] (972)

There are further properties. The bracket is antisymmetric

\[ \{ f, g \} = - \{ g, f \} \] (973)

and satisfies the Jacobi identity,

\[
\{ f, \{ g, h \} \} + \{ g, \{ h, f \} \} + \{ g, \{ h, f \} \} = 0 
\] (974)

176
for all functions $f$, $g$ and $h$. These properties are two of the three defining properties of a Lie algebra (the third defining property of a Lie algebra is that the set of objects considered, in this case the space of functions, be a vector space). The antisymmetry is obvious, but proving the Jacobi identity takes some work. Working first on one term,

\[
\{f, \{g, h\}\} = \left\{ f, \Omega^{AB} \frac{\partial g}{\partial \xi^A} \frac{\partial h}{\partial \xi^B} \right\} \\
= \Omega^{CD} \frac{\partial}{\partial \xi^C} \left( \Omega^{AB} \frac{\partial g}{\partial \xi^A} \frac{\partial h}{\partial \xi^B} \right) \\
= \Omega^{CD} \Omega^{AB} \frac{\partial f}{\partial \xi^C} \left( \frac{\partial^2 g}{\partial \xi^D \partial \xi^A} \frac{\partial h}{\partial \xi^B} + \frac{\partial g}{\partial \xi^A} \frac{\partial^2 h}{\partial \xi^D \partial \xi^B} \right)
\]

Now, combining all three,

\[
J = \{f, \{g, h\}\} + \{g, \{h, f\}\} + \{g, \{h, f\}\} = \Omega^{CD} \Omega^{AB} \frac{\partial f}{\partial \xi^C} \frac{\partial h}{\partial \xi^D} \left( \frac{\partial^2 g}{\partial \xi^E \partial \xi^A} \frac{\partial h}{\partial \xi^B} + \frac{\partial g}{\partial \xi^A} \frac{\partial^2 h}{\partial \xi^E \partial \xi^B} \right)
\]

Collecting corresponding second derivatives we have

\[
J = \Omega^{CD} \Omega^{AB} \frac{\partial f}{\partial \xi^C} \frac{\partial h}{\partial \xi^D} \frac{\partial^2 g}{\partial \xi^E \partial \xi^A} + \Omega^{CD} \Omega^{AB} \frac{\partial h}{\partial \xi^D} \frac{\partial^2 f}{\partial \xi^E \partial \xi^A} + \Omega^{CD} \Omega^{AB} \frac{\partial f}{\partial \xi^C} \frac{\partial g}{\partial \xi^D} \frac{\partial^2 h}{\partial \xi^E \partial \xi^A}
\]

Look at the first pair, renaming indices

\[
J_1 = \Omega^{CD} \Omega^{AB} \frac{\partial f}{\partial \xi^C} \frac{\partial h}{\partial \xi^B} \frac{\partial^2 g}{\partial \xi^D \partial \xi^A} + \Omega^{CD} \Omega^{AB} \frac{\partial h}{\partial \xi^D} \frac{\partial^2 f}{\partial \xi^B \partial \xi^A}
\]
\[ \begin{align*}
&= \Omega^{CD} \Omega_{AB} \frac{\partial f}{\partial \xi^C} \frac{\partial h}{\partial \xi^B} \frac{\partial^2 g}{\partial \xi^D \partial \xi^A} + \Omega^{CD'} \Omega_{B'A'} \frac{\partial f}{\partial \xi^C} \frac{\partial h}{\partial \xi^{B'}} \frac{\partial^2 g}{\partial \xi^{D'} \partial \xi^{A'}} \\
&= 0
\end{align*} \]

The terms cancel because they differ only by the interchange \( \Omega^B A' = -\Omega^A B' \). The remaining pairs cancel in the same way.

We conclude the section with Poisson’s theorem. Suppose \( f \) and \( g \) are constants of the motion. Then Poisson’s theorem states that their Poisson bracket, \( \{ f, g \} \), is also a constant of the motion. To prove the theorem, we start with \( f \) and \( g \) constant:

\[ 0 = \frac{df}{dt} = \{ f, H \} + \frac{\partial f}{\partial t} = 0 \]
\[ 0 = \frac{dg}{dt} = \{ g, H \} + \frac{\partial g}{\partial t} = 0 \]

Now consider the time dependence of the bracket:

\[ \frac{d}{dt} \{ f, g \} = \{ \{ f, g \}, H \} + \frac{\partial}{\partial t} \{ f, g \} \]

Using the Jacobi identity on the first term on the right, and the relation for time derivatives on the second term, we have

\[ \frac{d}{dt} \{ f, g \} = \{ \{ f, g \}, H \} + \frac{\partial}{\partial t} \{ f, g \} = -\{ \{ g, H \}, f \} - \{ \{ H, f \}, g \} + \left\{ \frac{\partial f}{\partial t}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} \right\} = 0 \]

\[ (975) \]
7.5 Hamilton’s principal function (The action as a function of the coordinates)

Landau and Lifshitz consider an object they call “the action as a function of the coordinates.” This is more commonly and correctly called Hamilton’s principal function. The action, of course, is a functional, not a function. That is, the action is a mapping of curves to the reals, rather than a mapping from points in space or phase space to the reals. We define Hamilton’s principal function \( S \) in the following way. Pick an initial point of space, \( x_0^i \), and an initial time \( t_0 \), and let \( S(x_f^i, t) \) be the value of the action evaluated along the actual path that a physical system would follow in going from \( x_0^i \) at time \( t_0 \) to \( x_f^i \) at time \( t_f \):

\[
S(x_f^i, t) = S_{\text{physical}} = \int_{t_0}^{t_f} L(x_{cl}^i, \dot{x}_{cl}^i, t) \, dt
\]

where \( x_{cl}^i(t) \) is the classical solution to the equations of motion and \( x_f^i \) is the final position at time \( t_f \). For the moment we assume that \( S(x_f^i, t) \) is a function, not a functional. The proof that \( S(x_f^i, t) \) is a function is given in the next section.

Dropping the subscript on \( x_f^i \), we now consider how \( S(x^i, t) \) changes with a general variation of the coordinates. First consider a general variation of the action:

\[
\delta S = \int_{t_0}^{t_f} \left( \frac{\partial L}{\partial x_i} \delta x_i + \frac{\partial L}{\partial \dot{x}_i} \delta \dot{x}_i \right) \, dt
\]

\[
= \left[ \frac{\partial L}{\partial x_i} \delta x_i \right]_{t_0}^{t_f} + \int_{t_0}^{t_f} \left( \frac{\partial L}{\partial x_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} \right) \delta x_i \, dt
\]

Now suppose require the path to satisfy the Euler-Lagrange equations, so that the second term vanishes. There are still some allowed variations, for example, variations of the final time. However, as we shall show, this restriction reduces \( S \) to \( S \). Then

\[
\frac{\partial L}{\partial x_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} = 0
\]

leaves us with

\[
\delta S_{\text{physical}} = \frac{\partial S}{\partial x^i} dx^i
\]
\[
\left[ \frac{\partial L}{\partial \dot{x}^i} \right]_0^t = \frac{\partial L}{\partial \dot{x}^i} \ dx^i (t) = p_i \ dx^i (t)
\]

or simply

\[
\frac{\partial S}{\partial x_i} = p_i \tag{978}
\]

To find the dependence of \( S \) on \( t \), we write

\[
dS = S \big|_{\text{physical}} = \int L \ dt
\]

as

\[
\frac{dS}{dt} = L \tag{979}
\]

But we also have

\[
\frac{dS}{dt} = \frac{\partial S}{\partial x_i} \dot{x}_i + \frac{\partial S}{\partial t} \tag{980}
\]

Equating these and using \( \frac{\partial S}{\partial x_i} = p_i \) gives

\[
L = \frac{\partial S}{\partial x_i} \dot{x}_i + \frac{\partial S}{\partial t} = p_i \dot{x}_i + \frac{\partial S}{\partial t} \tag{981}
\]

so that the partial of \( S \) with respect to \( t \) is

\[
\frac{\partial S}{\partial t} = L - p_i \dot{x}_i = -H \tag{982}
\]

Combining the results for the derivatives of \( S \) we may write

\[
dS = \frac{\partial S}{\partial x^i} dx^i + \frac{\partial S}{\partial t} dt = p_i dx^i - H dt \tag{983}
\]

This is a nontrivial condition on the solution of the classical problem. It means that form \( p_i dx_i - H dt \) must be a total differential, which cannot be true for arbitrary \( p_i \) and \( H \). In the next sub-section we find the conditions it implies.
7.5.1 Integrability of Hamilton’s principal function

Before moving to the proof that \( S \) is a function, we consider a slightly different variation of the action. Let \( \delta x^i \) result from a change in both the initial and final positions and times. Then we have

\[
\delta S = \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial x^i} \delta x_i + \frac{\partial L}{\partial \dot{x}_i} \delta \dot{x}_i \right) dt
\]

\[
= \left[ \frac{\partial L}{\partial x^i} \delta x_i \right]_{t_0}^{t_1} + \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial x^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} \right) \delta x_i dt
\]

(984)

Once again requiring the path to satisfy the Euler-Lagrange equations, so that the second term vanishes, we have

\[
\frac{\partial S}{\partial x^i} dx^i = \left[ \frac{\partial L}{\partial x^i} dx^i \right]_{t_0}^{t_1}
\]

\[
= p_i (t_1) dx^i (t_1) - p_i (t_0) dx^i (t_0)
\]

Holding \( x^i \) constant at \( t_0 \) leaves

\[
\frac{\partial S}{\partial x^i (t_0)} = p_i (t_1)
\]

while holding \( x^i \) constant at the upper limit \( t_1 \) while varying the lower limit \( t_0 \) gives

\[
\frac{\partial S}{\partial x^i (t_0)} = -p_i (t_0)
\]

The properties of \( S \) are easy to find using differential forms. The differential of Hilbert’s principal function,

\[
dS = p_i dx_i - H dt
\]

(985)

is a 1-form,

\[
dS = p_i dx_i - H dt
\]

(986)

Now, expanding the differential we have

\[
\frac{\partial S}{\partial x_i} dx_i + \frac{\partial S}{\partial p_i} dp_i + \frac{\partial S}{\partial t} dt = p_i dx_i - H dt
\]

(987)
so comparing term by term,

\[ \frac{\partial S}{\partial p_i} = 0 \]  \hspace{1cm} (988)
\[ \frac{\partial S}{\partial x_i} = p_i \]  \hspace{1cm} (989)
\[ \frac{\partial S}{\partial t} = -H \]  \hspace{1cm} (990)

These are the same relations we found in the last section by a different method.

We can do more than this, however. One of the most powerful uses of differential forms is for finding integrability conditions. Here’s how it works. Suppose we allow \( p_i \) to be arbitrary rather than constrained by the motion. Then we have a 1-form

\[ \omega = p_i \, dx_i - H \, dt \]  \hspace{1cm} (991)

which is not necessarily the differential of a function \( S \).

Now we can use the following theorem (the converse to the Poincaré lemma). Suppose (in a star-shaped region) that

\[ \int \omega = 0 \]  \hspace{1cm} (992)

for some \( p \)-form \( \theta \). The \( p \)-form is then said to be closed. It follows that

\[ \theta = d\sigma \]  \hspace{1cm} (993)

for some \( (p-1) \)-form \( \sigma \). If \( \theta \) can be written in this form, it is said to be exact. Thus, closed forms are exact. Conversely, if \( \theta = d\sigma \), then \( d\theta = 0 \) because

\[ d\theta = d^2\sigma = 0 \]  \hspace{1cm} (994)

because \( d^2 = 0 \) always, so exact forms are closed.

You are already familiar with certain applications of the converse to the Poincaré lemma. The simplest is for the differential of a one-form,:

\[ d\theta = 0 \]  \hspace{1cm} (995)

Then there exists a function \( f \) (a function is a zero-form) such that

\[ \theta = df \]  \hspace{1cm} (995)
This is the usual condition for the integrability of a conservative force. Think of \( f \) as the negative of the potential and \( \theta \) as the force, \( \theta = F_i dx^i \). Then the integrability condition
\[
d\theta = 0
\]is just
\[
0 = d\theta = dF_i \wedge dx^i = \frac{\partial F_i}{\partial x^j} dx^j \wedge dx^i
= \frac{1}{2} \left( \frac{\partial F_i}{\partial x^j} dx^j \wedge dx^i - \frac{\partial F_i}{\partial x^j} dx^i \wedge dx^j \right)
= \frac{1}{2} \left( \frac{\partial F_i}{\partial x^j} dx^j \wedge dx^i - \frac{\partial F_j}{\partial x^i} dx^j \wedge dx^i \right)
= \frac{1}{2} \left( \frac{\partial F_i}{\partial x^j} - \frac{\partial F_j}{\partial x^i} \right) dx^j \wedge dx^i
\]
This is equivalent to the vanishing of the curl of \( F \), the usual condition for the existence of a potential.

The Poincaré lemma and its converse give us a way to tell when a 1-form is the differential of a function. Given
\[
\omega = p_i dx_i - H dt
\]we have \( \omega = dS \) if and only if \( d\omega = 0 \). Therefore, demand
\[
0 = d\omega = dp_i \wedge dx_i - \frac{\partial H}{\partial x_i} dx_i \wedge dt - \frac{\partial H}{\partial p_i} dp_i \wedge dt - \frac{\partial H}{\partial t} dt \wedge dt
= dp_i \wedge dx_i - \frac{\partial H}{\partial x_i} dx_i \wedge dt - \frac{\partial H}{\partial p_i} dp_i \wedge dt
\]
where the last term vanishes because \( dt \wedge dt = 0 \). We can rearrange this result as
\[
0 = dp_i \wedge dx_i - \frac{\partial H}{\partial x_i} dx_i \wedge dt - \frac{\partial H}{\partial p_i} dp_i \wedge dt
= \left( dp_i + \frac{\partial H}{\partial x_i} dt \right) \wedge \left( dx_i - \frac{\partial H}{\partial p_i} dt \right)
\]183
where we again use $dt \wedge dt = 0$ and also use $dx_i \wedge dt = -dt \wedge dx_i$. Since each factor is one of Hamilton’s equations, this condition is clearly satisfied when the equations of motion are satisfied. Therefore, we have proved our earlier claim that Hamilton’s principal function is indeed a function.

Another way to see the same result is to try to integrate $\omega$ directly to find the function. This gives

$$S = \int \omega$$

which is a function (that is, single valued) if and only if the value of the integral is independent of the path. We therefore have the condition

$$\int_{C_1} \omega = \int_{C_2} \omega$$

for any two paths, $C_1$ and $C_2$ having the same endpoints. If we follow $C_1$ in one direction and $C_2$ in the other we have a closed path, $C_1 - C_2$ so that

$$\int_{C_1 - C_2} \omega = \int_{C_1} \omega - \int_{C_2} \omega = 0$$

Using Stoke’s theorem (which is actually just one case of the converse to the Poincaré lemma), we see that $\omega$ must be curl free.

We can also consider the converse to the integrability of $S$. However, notice that integrability of $\omega$ is not sufficient to guarantee that all of the equations of motion are satisfied since it is clearly possible to make one factor vanish without the other vanishing. We conclude that $d\omega = 0$ (and hence $\omega = dS$) is a necessary but not sufficient condition for the equations of motion to be satisfied.

### 7.5.2 Hamilton’s equations from $\omega$

It is possible to write the action in terms of $x_i$ and $p_i$ and vary these independently to arrive at Hamilton’s equations of motion. We have

$$S = \int \omega$$

(1000)

Notice that this is immediate if we work from $L$:

$$S = \int L dt$$

184
\[ \omega = \int (p_i \dot{x}_i - H) \, dt \]
\[ = \int (p_i \, dx_i - H \, dt) \]
\[ = \int \omega \]  \hspace{1cm} (1001) 

Since \( \omega \) depends on position and momentum (rather than position and velocity), it is these we vary. Thus:

\[ \delta S = \delta \int (p_i \dot{x}_i - H) \, dt \]
\[ = \int \left( \delta p_i \dot{x}_i + p_i \delta \dot{x}_i - \frac{\partial H}{\partial x_i} \delta x_i - \frac{\partial H}{\partial p_i} \delta p_i \right) \, dt \]
\[ = p_i \delta x_i \bigg|_{t_i}^{t_1} + \int \left( \delta p_i \dot{x}_i - \dot{p}_i \delta x_i - \frac{\partial H}{\partial x_i} \delta x_i - \frac{\partial H}{\partial p_i} \delta p_i \right) \, dt \]
\[ = \int \left( \left( \dot{x}_i \frac{\partial H}{\partial p_i} - \dot{p}_i \frac{\partial H}{\partial x_i} \right) \delta p_i - \left( \dot{p}_i + \frac{\partial H}{\partial x_i} \right) \delta x_i \right) \, dt \]  \hspace{1cm} (1002) 

and since \( \delta p_i \) and \( \delta x_i \) are independent we conclude

\[ \dot{x}_i = \frac{\partial H}{\partial p_i} \]  \hspace{1cm} (1003) 
\[ \dot{p}_i = -\frac{\partial H}{\partial x_i} \]  \hspace{1cm} (1004) 

as required. Notice that it is the full action, and not Hamilton’s principal function, that we have varied here. We cannot vary the position and momentum within the integral form of the principal function because they are fixed on the classical path.
7.6 Liouville’s theorem

Liouville’s theorem provides an excellent example of the usefulness of differential forms. The theorem states that the region of phase space occupied by a system is independent of time. Since the theorem involves the invariance of volume elements, differential forms are the right tool.

Let’s clarify the meaning of the theorem first. Consider a physical system with \( s \) degrees of freedom, with its motion determined by a particular, fixed Hamiltonian, \( H(x_i, p_i, t) \). Now consider a region of \( 2s \)-dimensional phase space, \( R \). Since the \( 2s \) Hamilton equations of motion are first order differential equations, the solutions depend on \( 2s \) initial constants. These constants are just enough to specify an initial point in phase space. Therefore, through any point in phase space, there is exactly one curve representing the subsequent evolution of the system if it were to start at that point.

Suppose we follow the evolution from each point in the region \( R \) for an amount of time, \( \tau \). For each starting point in \( R \) we find a new point. These new points fill a new region \( R' \). Liouville’s theorem states that the volume of the two regions, \( R \) and \( R' \), is the same.

The proof is as follows. First, we show there is a single canonical transformation that takes each initial point \( (x_i(t), p_i(t)) \) in \( R \), the final points \( (x_i(t + \tau), p_i(t + \tau)) \) in \( R' \). To see this, notice that the final point may be written as a function of the initial point, since the initial point provides the entire set of initial conditions. Let these relationships be given by

\[
\begin{align*}
x_i(t + \tau) &= f_i(x_j(t), p_j(t), t, \tau) \\
p_i(t + \tau) &= g_i(x_j(t), p_j(t), t, \tau)
\end{align*}
\]  

where we treat \( \tau \) as a parameter. This expresses the time evolution as a transformation of some sort. To see that the transformation is canonical, it is sufficient to show that there is a generating function that gives these functions. Hamilton’s principal function does the trick. Now since we regard \( (x_i(t), p_i(t)) \) as the original variables and \( (x_i(t + \tau), p_i(t + \tau)) \) as the new variables we need a generating function, \( F \), that satisfies

\[
\begin{align*}
p_i(t) &= \frac{\partial F}{\partial x_i(t)} \\
p_i(t + \tau) &= -\frac{\partial F}{\partial q_i(t + \tau)}
\end{align*}
\]  

186
\[ H(t + \tau) = H(t) + \frac{\partial F}{\partial t} \]  

(1009)

But consider Hamilton’s principal function, including variations of both endpoints:

\[ S(x_i(t), t; x_i(t + \tau), \tau) = \int_t^{t+\tau} L(x(t), \dot{x}(t), t) dt \]  

(1010)

As shown in our original discussion of Hamilton’s function, we have

\[ \frac{\partial S}{\partial x_i(t + \tau)} = p_i(t + \tau) \]  

(1011)

at the upper limit, while at the lower limit we have

\[ \frac{\partial S}{\partial x_i(t)} = -p_i(t) \]  

(1012)

Similarly, the partial derivatives of \( S \) with respect to time, picks up both limits

\[ \frac{\partial S}{\partial t} = H(t) - H(t + \tau) \]  

(1013)

Comparing the last three equations with the requirements for our generating function, we see that we can take \( F = -S \). Therefore, the transformation from \( R \) to \( R' \) is canonical.

Now we can turn to the main conclusion of the theorem. Since the transformation from \( R \) to \( R' \) is canonical, we need only show that canonical transformations preserve phase space volumes. Write the symplectic form as

\[ \Omega = \frac{1}{2} \Omega_{ab} d\xi_a \land d\xi_b \]  

(1014)

The factor of \( 1/2 \) is so that \( \Omega \) reduces to simply

\[ \Omega = dp_i \land dx^i \]  

(1015)

As we have shown, a transformation is canonical if and only if it preserves \( \Omega \). It follows that if we wedge together two copies of \( \Omega \), the result is also unchanged by any canonical transformation:

\[ \Omega \land \Omega = dp_i \land dx^i \land dp_j \land dx^j \]  

(1016)
Since both \(i\) and \(j\) are summed over all coordinates, this 4-form doesn’t vanish. There is always some product of terms that doesn’t repeat. If we are working with a 4-dimensional phase space spanned by \((x,y,p_x,p_y)\) then we already have the volume form. For this simple example,

\[
\Omega = dp_i \wedge dx^i \\
= dp_x \wedge dx + dp_y \wedge dy
\]

and therefore

\[
\Omega \wedge \Omega = (dp_x \wedge dx + dp_y \wedge dy) \wedge (dp_x \wedge dx + dp_y \wedge dy)
= dp_x \wedge dx \wedge dp_y \wedge dy + dp_x \wedge dx \wedge dp_y \wedge dy \\
+ dp_y \wedge dy \wedge dp_x \wedge dx + dp_y \wedge dy \wedge dp_x \wedge dx
= 0 + dp_x \wedge dx \wedge dp_y \wedge dy + dp_y \wedge dy \wedge dp_x \wedge dx + 0
= -2dx \wedge dy \wedge dp_x \wedge dp_y
\]

which is clearly proportional to the volume element of the full 4-dimensional phase space.

In higher dimensional phase spaces, we can continue this process as long as we don’t run out of independent coordinates (recall that, for example, \(dx_1 \wedge dx_1 = 0\)):

\[
\Omega^k \equiv \Omega \wedge \ldots \wedge \Omega = dp_i \wedge dx_i \wedge \ldots \wedge dp_j \wedge dx_j \quad (1017)
\]

\[
k \text{ times} \quad (1018)
\]

The process only stops when \(k = s\), for then each differential in the space occurs once:

\[
\Omega^s \equiv s! (dp_1 \wedge dx_1 \wedge \ldots \wedge dp_s \wedge dx_s) \quad (1019)
\]

But forms are the things that go under integral signs, so this is precisely \(s!\) times the volume element of phase space! Therefore, since \(\Omega' = \Omega\),

\[
V_R = \int_R \Omega^s = \int_{R'} (\Omega')^s = V_{R'} \quad (1020)
\]

completing the proof of Liouville’s theorem.

Notice that, along the way we have also proved that all of the even dimensional \((2k)\) volume forms \(\Omega^k \sim d^k pd^k x\) are also invariant under canonical transformations.
7.7 General solution in Hamiltonian dynamics

We conclude with the crowning theorem of Hamiltonian dynamics: a proof that for any Hamiltonian dynamical system there exists a canonical transformation to a set of variables on phase space such that the paths of motion reduce to single points. Clearly, this theorem shows the power of canonical transformations! The theorem relies on describing solutions to the Hamilton-Jacobi equation, which we introduce first.

7.7.1 The Hamilton-Jacobi Equation

We have the following equations governing Hamilton’s principal function.

\[
\begin{align*}
\frac{\partial S}{\partial p_i} &= 0 \\
\frac{\partial S}{\partial x_i} &= p_i \\
\frac{\partial S}{\partial t} &= -H
\end{align*}
\]  

(1021)  
(1022)  
(1023)

Since the Hamiltonian is a given function of the phase space coordinates and time, \( H = H(x_i, p_i, t) \), we combine the last two equations:

\[
\frac{\partial S}{\partial t} = -H(x_i, p_i, t) = -H(x_i, \frac{\partial S}{\partial x_i}, t)
\]

(1024)

This first order differential equation in \( s + 1 \) variables \((t, x_i; i = 1, \ldots, s)\) for the principal function \( S \) is the Hamilton-Jacobi equation. Notice that the Hamilton-Jacobi equation has the same general form as the Schrödinger equation and is equally difficult to solve for all but special potentials. Nonetheless, we are guaranteed that a complete solution exists, and we will assume below that we can find it. Before proving our central theorem, we digress to examine the exact relationship between the Hamilton-Jacobi equation and the Schrödinger equation.

7.7.2 Quantum Mechanics and the Hamilton-Jacobi equation

The Hamiltonian-Jacobi equation provides the most direct link between classical and quantum mechanics. There is considerable similarity between
the Hamilton-Jacobi equation and the Schrödinger equation:

\[ \frac{\partial S}{\partial t} = -H \left( x_i, \frac{\partial S}{\partial x_i}, t \right) \]  
\[ i\hbar \frac{\partial \psi}{\partial t} = H(\hat{x}_i, \hat{p}_i, t) \]

We make the relationship precise as follows.

Suppose the Hamiltonian in each case is that of a single particle in a potential:

\[ H = \frac{p^2}{2m} + V(x) \]

Write the quantum wave function as

\[ \psi = A e^{i\phi} \]

The Schrödinger equation becomes

\[ i\hbar \frac{\partial}{\partial t} \left( A e^{i\phi} \right) = -\frac{\hbar^2}{2m} \nabla^2 \left( A e^{i\phi} \right) + V \left( A e^{i\phi} \right) \]
\[ i\hbar \frac{\partial A}{\partial t} e^{i\phi} - A e^{i\phi} \frac{\partial \phi}{\partial t} = -\frac{\hbar^2}{2m} \nabla \cdot \left( e^{i\phi} \nabla A + \frac{i}{\hbar} A e^{i\phi} \nabla \phi \right) + VA e^{i\phi} \]
\[ = -\frac{\hbar^2}{2m} e^{i\phi} \left( \frac{i}{\hbar} \nabla \phi \nabla A + \nabla^2 A \right) \]
\[ -\frac{\hbar^2}{2m} e^{i\phi} \left( \frac{i}{\hbar} \nabla A \cdot \nabla \phi + \frac{i}{\hbar} A \nabla^2 \phi \right) \]
\[ -\frac{\hbar^2}{2m} \left( \frac{i}{\hbar} \right)^2 e^{i\phi} \left( A \nabla \phi \cdot \nabla \phi \right) \]
\[ +VA e^{i\phi} \]

Then cancelling the exponential,

\[ i\hbar \frac{\partial A}{\partial t} - A \frac{\partial \phi}{\partial t} = -\frac{i\hbar}{2m} \nabla \phi \nabla A - \frac{\hbar^2}{2m} \nabla^2 A \]
\[ -\frac{i\hbar}{2m} A \cdot \nabla \phi - \frac{i\hbar}{2m} \nabla A \nabla^2 \phi \]
\[ +\frac{1}{2m} \left( A \nabla \phi \cdot \nabla \phi \right) + VA \]
Collecting by powers of $h$,

\[
O(h^0) : \quad \frac{\partial \varphi}{\partial t} = \frac{1}{2m} \nabla \varphi \cdot \nabla \varphi + V
\]

\[
O(h^1) : \quad \frac{1}{A} \frac{\partial A}{\partial t} = -\frac{1}{2m} \left( \frac{2}{A} \nabla A \cdot \nabla \varphi + \nabla^2 \varphi \right)
\]

\[
O(h^2) : \quad 0 = -\frac{\hbar^2}{2m} \nabla^2 A
\]  

(1029)

The zeroth order terms is the Hamilton-Jacobi equation, with $\varphi = S$:

\[
-\frac{\partial S}{\partial t} = \frac{1}{2m} \nabla S \cdot \nabla S + V
\]

\[
= \frac{1}{2m} p^2 + V(x)
\]

where $p = \nabla S$. Therefore, the Hamilton-Jacobi equation is the $h \to 0$ limit of the Schrödinger equation.

### 7.7.3 Trivialization of the motion

We now seek a solution, in principle, to the complete mechanical problem. The solution is to find a canonical transformation that makes the motion trivial. Hamilton’s principal function, the solution to the Hamilton-Jacobi equation, is the generating function of this canonical transformation.

To begin, suppose we have a solution to the Hamilton-Jacobi equation of the form

\[
S = g(t, x_1, \ldots, x_s, \alpha_1, \ldots, \alpha_s) + A
\]  

(1030)

where the $\alpha_i$ and $A$ provide $s + 1$ constants describing the solution. Such a solution is called a complete integral of the equation, as opposed to a general integral which depends on arbitrary functions. We will show below that a complete solution leads to a general solution. We use $S$ as a generating function.

Our canonical transformation will take the variables $(x_i, p_i)$ to a new set of variables $(\beta^i, \alpha_i)$. Since $S$ depends on the old coordinates $x_i$ and the new momenta $\alpha_i$, we have the relations

\[
p_i = \frac{\partial S}{\partial x_i}
\]  

(1031)
\begin{align}
\beta_i &= \frac{\partial S}{\partial \alpha_i} \quad (1032) \\
H' &= H + \frac{\partial S}{\partial t} \quad (1033)
\end{align}

Notice that the new Hamiltonian, \( H' \), vanishes because \( S \) satisfies the Hamiltonian-Jacobi equation!. With \( H' = 0 \), Hamilton’s equations in the new canonical coordinates are simply

\begin{align}
\frac{d\alpha_i}{dt} &= \frac{\partial H'}{\partial \beta_i} = 0 \quad (1034) \\
\frac{d\beta_i}{dt} &= -\frac{\partial H'}{\partial \alpha_i} = 0 \quad (1035)
\end{align}

with solutions

\begin{align}
\alpha_i &= \text{const.} \quad (1036) \\
\beta_i &= \text{const.} \quad (1037)
\end{align}

The system remains at the phase space point \((\alpha_i, \beta_i)\). To find the motion in the original coordinates as functions of time and the \(2s\) constants of motion,

\[ x_i = x_i(t; \alpha_i, \beta_i) \]

we can algebraically invert the \(s\) equations

\[ \beta_i = \frac{\partial g(x_i, t, \alpha_i)}{\partial \alpha_i} \]

The momenta may be found by differentiating the principal function,

\[ p_i = \frac{\partial S(x_i, t, \alpha_i)}{\partial x_i} \]

Therefore, solving the Hamilton-Jacobi equation is the key to solving the full mechanical problem. Furthermore, we know that a solution exists because Hamilton’s equations satisfy the integrability equation for \( S \).

We note one further result. While we have made use of a complete integral to solve the mechanical problem, we may want a general integral of the Hamilton-Jacobi equation. The difference is that a complete integral of an equation in \(s + 1\) variables depends on \(s + 1\) constants, while a general
integral depends on \( s \) functions. Fortunately, a complete integral of the equation can be used to construct a general integral, and there is no loss of generality in considering a complete integral. We see this as follows. A complete solution takes the form

\[
S = g(t, x_1, \ldots, x_s, \alpha_1, \ldots, \alpha_s) + A
\]  

(1038)

To find a general solution, think of the constant \( A \) as a function of the other \( s \) constants, \( A(\alpha_1, \ldots, \alpha_s) \). Now replace each of the \( \alpha_i \) by a function of the coordinates and time, \( \alpha_i \to h_i(t, x_i) \). This makes \( S \) depend on arbitrary functions, but we need to make sure it still solves the Hamilton-Jacobi equation. It will provided the partials of \( S \) with respect to the coordinates remain unchanged. In general, these partials are given by

\[
\frac{\partial S}{\partial x_i} = \left( \frac{\partial S}{\partial x_i} \right)_{h_i=\text{const.}} + \left( \frac{\partial S}{\partial h_k} \right)_{x=\text{const.}} \frac{\partial h_k}{\partial x_i}
\]  

(1039)

We therefore still have solutions provided

\[
\left( \frac{\partial S}{\partial h_k} \right)_{x=\text{const.}} \frac{\partial h_k}{\partial x_i} = 0
\]

and since we want \( h_k \) to be an arbitrary function of the coordinates, we demand

\[
\left( \frac{\partial S}{\partial h_k} \right)_{x=\text{const.}} = 0
\]

Then

\[
\frac{\partial S}{\partial h_k} = \frac{\partial}{\partial h_k} (g(t, x_i, \alpha_i) + A(\alpha_i)) = 0
\]  

(1040)

and we have

\[
A(\alpha_1, \ldots, \alpha_s) = \text{const.} - g
\]

This just makes \( A \) into some specific function of \( x^i \) and \( t \).

Since the partials with respect to the coordinates are the same, and we haven’t changed the time dependence,

\[
S = g(t, x_1, \ldots, x_s, h_1, \ldots, h_s) + A(h_i)
\]  

(1041)

is a general solution to the Hamilton-Jacobi equation.
**Example 1: Free particle**  The simplest example is the case of a free particle, for which the Hamiltonian is

\[ H = \frac{p^2}{2m} \]

and the Hamilton-Jacobi equation is

\[ \frac{\partial S}{\partial t} = -\frac{1}{2m} (S')^2 \]

Let

\[ S = f(x) - Et \]

Then \( f(x) \) must satisfy

\[ \frac{df}{dx} = \sqrt{2mE} \]

and therefore

\[ f(x) = \sqrt{2mE}x - c \]

\[ = \pi x - c \]

where \( c \) is constant and we write the integration constant \( E \) in terms of the new (constant) momentum. Hamilton’s principal function is therefore

\[ S(x, \pi, t) = \pi x - \frac{\pi^2}{2m}t - c \]

Then, from eqs.(965), we have

\[ p = \frac{\partial S}{\partial x} = \pi \]

\[ q = \frac{\partial S}{\partial \pi} = x - \frac{\pi}{m}t \]

\[ H' = H + \frac{\partial S}{\partial t} = H - E \]

(1042)

Because \( E = H \), the new Hamiltonian, \( H' \), is zero. This means that both \( q \) and \( \pi \) are constant. The solution for \( x \) and \( p \) follows immediately:

\[ x = q + \frac{\pi}{m}t \]

\[ p = \pi \]

194
We see that the new canonical variables \((q, \pi)\) are just the initial position and momentum of the motion, and therefore do determine the motion. The fact that knowing \(q\) and \(\pi\) is equivalent to knowing the full motion rests here on the fact that \(S\) generates motion along the classical path. In fact, given initial conditions \((q, \pi)\), we can use Hamilton’s principal function as a generating function but treat \(\pi\) as the old momentum and \(x\) as the new coordinate to reverse the process above and generate \(x(t)\) and \(p\).

**Example 2: Simple harmonic oscillator** For the simple harmonic oscillator, the Hamiltonian becomes

\[
H = \frac{p^2}{2m} + \frac{1}{2}kx^2
\]

and the Hamilton-Jacobi equation is

\[
\frac{\partial S}{\partial t} = -\frac{1}{2m}(S')^2 + \frac{1}{2}kx^2
\]

Letting

\[
S = f(x) - Et
\]

as before, \(f(x)\) must satisfy

\[
\frac{df}{dx} = \sqrt{2m \left( E - \frac{1}{2}kx^2 \right)}
\]

and therefore

\[
\begin{align*}
f(x) &= \int \sqrt{2m \left( E - \frac{1}{2}kx^2 \right)} \, dx \\
&= \int \sqrt{\pi^2 - mkx^2} \, dx
\end{align*}
\]

where we have set \(E = \frac{\pi^2}{2m}\). Now let \(\sqrt{mk}x = \pi y\). The integral is immediate:

\[
\begin{align*}
f(x) &= \int \sqrt{\pi^2 - mkx^2} \, dx \\
&= \frac{\pi^2}{\sqrt{mk}} \int \cos^2 y \, dy \\
&= \frac{\pi^2}{2\sqrt{mk}} (y + \sin y \cos y)
\end{align*}
\]
Hamilton’s principal function is therefore

\[
S(x, \pi, t) = \frac{\pi^2}{2\sqrt{mk}} \left( \sin^{-1} \left( \sqrt{\frac{mkx}{\pi}} \right) + \sqrt{mk} \frac{x}{\pi} \sqrt{1 - mkx^2} \right) - \frac{\pi^2}{2m} t - c
\]

\[
= \frac{\pi^2}{2\sqrt{mk}} \sin^{-1} \left( \sqrt{\frac{mkx}{\pi}} \right) + \frac{x}{2} \sqrt{\pi^2 - mkx^2} - \frac{\pi^2}{2m} t - c
\]

and we may use it to generate the canonical change of variable.

This time we have

\[
p = \frac{\partial S}{\partial x}
\]

\[
= \frac{\pi}{2} \frac{1}{\sqrt{1 - mkx^2}} + \frac{1}{2} \sqrt{\pi^2 - mkx^2} + \frac{x}{2} \sqrt{\pi^2 - mkx^2} - \frac{mkx}{\sqrt{\pi^2 - mkx^2}}
\]

\[
= \sqrt{\pi^2 - mkx^2} \left( \frac{\pi^2}{2} + \frac{1}{2} \left( \pi^2 - mkx^2 \right) - \frac{mkx^2}{2} \right)
\]

\[
q = \frac{\partial S}{\partial \pi}
\]

\[
= \frac{\pi}{\sqrt{mk}} \sin^{-1} \left( \sqrt{\frac{mkx}{\pi}} \right) - \frac{\pi^2}{2\sqrt{mk}} \frac{1}{\sqrt{1 - mkx^2}} \left( -\sqrt{\frac{mkx}{\pi^2}} \right)
\]

\[
+ \frac{x}{2} \sqrt{\pi^2 - mkx^2} - \frac{\pi}{m} t
\]

\[
= \frac{\pi}{\sqrt{mk}} \sin^{-1} \left( \sqrt{\frac{mkx}{\pi}} \right) - \frac{\pi}{m} t
\]

\[
H' = H + \frac{\partial S}{\partial t} = H - E = 0
\]

The first equation relates \( p \) to the energy and position, the second gives the new position coordinate \( q \), and third equation shows that the new Hamiltonian is zero. Hamilton’s equations are trivial, so that \( \pi \) and \( q \) are constant, and we can invert the expression for \( q \) to give the solution. Setting \( \omega = \sqrt{\frac{k}{m}} \), the solution is

\[
x(t) = \frac{\pi}{m\omega} \sin \left( \frac{m\omega}{\pi} q + \omega t \right)
\]

196
Example 3: One dimensional particle motion

Now suppose a particle with one degree of freedom moves in a potential $U(x)$. Little is changed. The Hamiltonian becomes

$$H = \frac{p^2}{2m} + U$$

and the Hamilton-Jacobi equation is

$$\frac{\partial S}{\partial t} = -\frac{1}{2m} (S')^2 + U(x)$$

Letting

$$S = f(x) - Et$$

as before, $f(x)$ must satisfy

$$\frac{df}{dx} = \sqrt{2m (E - U(x))}$$

and therefore

$$f(x) = \int \sqrt{2m (E - U(x))} dx$$

$$= \int \sqrt{\pi^2 - 2mU(x)} dx$$

where we have set $E = \frac{\pi^2}{2m}$. Hamilton’s principal function is therefore

$$S(x, \pi, t) = \int \sqrt{\pi^2 - 2mU(x)} dx - \frac{\pi^2}{2m} t - c$$

and we may use it to generate the canonical change of variable.
This time we have
\[\begin{align*}
p &= \frac{\partial S}{\partial x} = \sqrt{\pi^2 - 2mU(x)} \\
q &= \frac{\partial S}{\partial \pi} = \frac{\partial}{\partial \pi} \left( \int_{x_0}^x \sqrt{\pi^2 - 2mU(x)} \, dx \right) - \frac{\pi}{m} t \\
H' &= H + \frac{\partial S}{\partial t} = H - E = 0 \quad (1043)
\end{align*}\]

The first and third equations are as expected, while for \(q\) we may interchange the order of differentiation and integration:
\[\begin{align*}
q &= \frac{\partial}{\partial \pi} \left( \int \sqrt{\pi^2 - 2mU(x)} \, dx \right) - \frac{\pi}{m} t \\
&= \int \frac{\partial}{\partial \pi} \left( \sqrt{\pi^2 - 2mU(x)} \right) \, dx - \frac{\pi}{m} t \\
&= \int \frac{\pi \, dx}{\sqrt{\pi^2 - 2mU(x)}} - \frac{\pi}{m} t \quad (1044)
\end{align*}\]

To complete the problem, we need to know the potential. However, even without knowing \(U(x)\) we can make sense of this result by combining the expression for \(q\) above to our previous solution to the same problem. There, conservation of energy gives a first integral to Newton’s second law,
\[\begin{align*}
E &= \frac{p^2}{2m} + U \\
&= \frac{1}{2} m \left( \frac{dx}{dt} \right)^2 + U
\end{align*}\]

so we arrive at the familiar quadrature
\[t - t_0 = \int dt = \int_{x_0}^x \frac{mdx}{\sqrt{2m(E-U)}}\]

Substituting into the expression for \(q\),
\[\begin{align*}
q &= \int \frac{\pi \, dx}{\sqrt{\pi^2 - 2mU(x)}} - \frac{\pi}{m} \int_{x_0}^x \frac{mdx}{\sqrt{2m(E-U)}} - \frac{\pi}{m} t_0 \quad (1045) \\
&= \int \frac{\pi \, dx}{\sqrt{\pi^2 - 2mU(x)}} - \int_{x_0}^x \frac{\pi \, dx}{\sqrt{\pi^2 - 2mU(x)}} - \frac{\pi}{m} t_0 \quad (1046) \\
&= \int_{x_0}^x \frac{\pi \, dx}{\sqrt{\pi^2 - 2mU(x)}} - \frac{\pi}{m} t_0 \quad (1047)
\end{align*}\]
We once again find that $q$ is a constant characterizing the initial configuration. Since $t_0$ is the time at which the position is $x_0$ and the momentum is $p_0$, we have the following relations:

\[
\frac{p^2}{2m} + U(x) = \frac{p_0^2}{2m} + U(x_0) = E = \text{const}.\]

and

\[
t - t_0 = \int_{x_0}^{x} \frac{dx}{\sqrt{\frac{2}{m}(E - U)}}
\]

which we may rewrite as

\[
t - \int_{x_0}^{x} \frac{dx}{\sqrt{\frac{2}{m}(E - U)}} = t_0 - \int_{x_0}^{x_0} \frac{dx}{\sqrt{\frac{2}{m}(E - U)}} = \frac{m}{\pi} q = \text{const}.\]
8 Appendices

8.1 Some rules for functional derivatives

We would like to define higher order functional derivatives. However, notice that the functional derivative of the functional \( S [x (t)] \) defined above is a function. This is a problem if we wish to differentiate a second time. However, before we take the limit on \( n \), we still have a functional. Therefore, we may take the functional derivative of

\[
\left( \frac{\delta S}{\delta x} \right)_{t',n} [x] = \frac{dS [x (t, t', \alpha, n)]}{d\alpha} \bigg|_{\alpha=0}
\]

\[
= \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial x(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}(t)} \right) h_n (t, t') \, dt
\]

regarding \( h_n (t, t') \) as given. Then the functional derivative of \( S [x] \) is

\[
\frac{\delta S}{\delta x} (t') = \lim_{n \to \infty} \left( \frac{\delta S}{\delta x} \right)_{t',n} [x]
\]

\[
= \lim_{n \to \infty} \frac{dS [x (t, t', \alpha, n)]}{d\alpha} \bigg|_{\alpha=0}
\]

Since \( \left( \frac{\delta S}{\delta x} \right)_{t',n} [x] \) is a functional, we may take its functional derivative:

\[
\frac{\delta}{\delta x} \left( \frac{\delta S}{\delta x} \right)_{t',n} [x]
\]

This procedure may be iterated as many times as desired. We now formalize the entire series.

Let \( S [x] = \int_{t_1}^{t_2} L (x, \dot{x}, \ldots) \, dt \) be a functional of a function \( x(t) \). We define a sequence of functionals by induction. Let

\[
S_{t_1,n_1} [x (t)] = \frac{dS [x (t, t_1, \alpha, n_1)]}{d\alpha} \bigg|_{\alpha=0}
\]

Now suppose \( S_{t_1, \ldots, t_k, n_1, \ldots, n_k} [x] \) is the \( k \)th functional of the sequence. Then the next functional in the sequence is defined by

\[
S_{t_1, \ldots, t_k, t_{k+1}, n_1, \ldots, n_k, n_{k+1}} [x (t)] = \frac{dS_{t_1, \ldots, t_k, n_1, \ldots, n_k} [x (t, t_{k+1}, \alpha, n_{k+1})]}{d\alpha} \bigg|_{\alpha=0}
\]
We define the $k^{th}$ functional derivative to be the function of $k$ variables given by the taking the limits,

$$\frac{\delta^k S}{\delta x^k} (t_1, \ldots, t_k) = \lim_{n_k \to \infty} \lim_{n_{k-1} \to \infty} \ldots \lim_{n_1 \to \infty} S_{t_1, \ldots, t_k, n_k, n_{k-1}, \ldots, n_1} [x(t)] \ldots$$

As an example, we compute the second and third functional derivatives of the classical, 1-particle action,

$$S [x] = \int_{t_A}^{t_B} L (x (t), \dot{x} (t)) \, dt$$

First, we construct the sequence of functionals. For the first,

$$S_{t_1, n_1} [x] = \left. \frac{d}{d \alpha} \int_{t_A}^{t_B} L (x (t, t_1, \alpha, n_1), \dot{x} (t, t_1, \alpha, n_1)) \, dt \right|_{\alpha = 0}$$

$$= \int_{t_A}^{t_B} \left( \frac{\partial L}{\partial x (t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x} (t)} \right) h_{n_1} (t, t_1) \, dt$$

Now, in more detail, the second: Again:

Suppose we take the limits on $n_1$ and $n_2$. Then we get Dirac delta functions. If we integrate this function over $t_1$, we get:

$$\int_{t_A}^{t_B} S (t_1, t) \, dt_1 = \frac{\partial^2 L (x (t))}{\partial x \partial x} - \frac{d}{dt} \left( \frac{\partial^2 L (x (t))}{\partial x \partial \dot{x}} \right)$$

$$+ \frac{d^2}{dt^2} \left( \frac{\partial^2 L (x (t))}{\partial \dot{x}^2} \right) - \frac{d^2}{dt^2} \frac{\partial^2 L (x (t))}{\partial \dot{x}^2}$$

$$= \frac{\partial^2 L (x (t))}{\partial x \partial x} - \frac{d}{dt} \left( \frac{\partial^2 L (x (t))}{\partial x \partial \dot{x}} \right)$$

Expand $S$ in a power series in $\alpha$:

$$S [x + \alpha h] = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{d^n S [x + \alpha h]}{d \alpha^n} \bigg|_{\alpha = 0} \alpha^n h^n$$

$$\frac{d S [x (t) + \alpha h (t)]}{d \alpha} = \left. \frac{d}{d \alpha} \int_{t_A}^{t_B} L (x (t) + \alpha h, \dot{x} (t) + \alpha \dot{h}) \, dt \right|_{\alpha = 0}$$

$$= \int_{t_A}^{t_B} \left( \frac{\partial L}{\partial x} h + \frac{\partial L}{\partial \dot{x}} \dot{h} \right) \, dt$$

$$= \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right) h \, dt$$

Then
8.2 Some intuitions

Some general ideas about variations.

\[
\frac{\delta S}{\delta x}(t') = \lim_{n \to \infty} \left. \frac{dS[x(t, t', \alpha, n)]}{d \alpha} \right|_{\alpha = 0}
\]

\[
= \lim_{n \to \infty} \int_{t_1}^{t_2} \left( \frac{\partial L}{\partial q(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}(t)} \right) h_n(t, t') dt
\]

\[
= \partial L \frac{\partial L}{\partial q(t')} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}(t')}
\]

Setting \( \delta x = h(t) \) we may write this as

\[
\delta S = \int \left( \frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \right) \delta x(t') dt'
\]  

(1048)

Comparing this to the differential of a multivariable function,

\[
df(x_1, \ldots, x_n) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i
\]

we see that \( \delta S \) is like the differential of a function with an uncountable infinity of independent variables, \( x(t) \) – that is, a functional! The summation is replaced by an integral. Formally rewriting,

\[
\delta S = \int \frac{\delta S}{\delta x(t)} \delta x(t) dt = \int \left( \frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \right) \delta x(t) dt
\]  

(1049)

we conjecture a transfinite chain rulerelation of the form

\[
\frac{\delta S}{\delta x(t)} = \int \left( \frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \right) \frac{\delta x(t')}{\delta x(t)} dt'
\]  

(1050)

This is somewhat approximate because we cannot really multiply and divide by \( \delta x(t) \), with may vanish for some values of \( t \). However, it is suggestive

We might write this much by just using the chain rule. What we need is to evaluate the basic functional derivative,

\[
\frac{\delta x(t')}{\delta x(t)}
\]  

(1051)
To see what this might be, consider the analogous derivative for a countable number of degrees of freedom. Beginning with

\[ \frac{\partial q^j}{\partial q^i} = \delta^j_i \tag{1052} \]

we notice that when we sum over the \( i \) index holding \( j \) fixed, we have

\[ \sum_i \frac{\partial q^j}{\partial q^i} = \sum_j \delta^j_i = 1 \tag{1053} \]

since \( j = i \) for only one value of \( j \). We demand the continuous version of this relationship. The sum over independent coordinates becomes an integral, \( \sum_i \to \int dt' \), so we demand

\[ \int \frac{\delta x(t')}{\delta x(t)} \, dt' = 1 \tag{1054} \]

This will be true provided we use a Dirac delta function for the derivative:

\[ \frac{\delta x(t')}{\delta x(t)} = \delta(t' - t) \tag{1055} \]

Substituting this expression into eq. (1050) gives the desired result for \( \frac{\delta S}{\delta x(t)} : \)

\[ \frac{\delta S}{\delta x(t)} = \int \left( \frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \right) \delta(t' - t) \, dt' \tag{1056} \]

\[ = \frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \tag{1057} \]

Notice how the Dirac delta function enters this calculation. When finding the extrema of \( S \) as before, we reach a point where we demand

\[ 0 = \delta S = \int \left( \frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} \right) h(t) \, dt \tag{1058} \]

for every function \( h(t) \). To complete the argument, we imagine \( h(t) \) of smaller and smaller compact support near a point at time \( t_0 \). The result of this limiting process is to conclude that the integrand must vanish at \( t_0 \). Since this limiting argument holds for any choice of \( t_0 \), we must have

\[ \frac{\partial L(x, \dot{x})}{\partial x} - \frac{d}{dt} \frac{\partial L(x, \dot{x})}{\partial \dot{x}} = 0 \tag{1059} \]
The Dirac delta function simply streamlines this limiting process; indeed, the Dirac delta is defined by just such a limiting procedure.

Let’s summarize the procedure. Given a functional

\[ f[x(t)] = \int g(x(t'), \dot{x}(t'), \ldots) \, dt' \]

the functional derivative is given by

\[
\frac{\delta f[x(t)]}{\delta x(t)} = \frac{\delta}{\delta x(t)} \int g(x(t'), \dot{x}(t'), \ldots) \, dt' \\
= \int \lim_{n \to \infty} \frac{dg(\varepsilon, x(t'), \dot{x}(t'), \ldots)}{d\varepsilon} \bigg|_{\varepsilon=0} \\
= \int \lim_{n \to \infty} \frac{dg(x(t') + \varepsilon h_n(t, t'), \dot{x}(t') + \varepsilon \dot{h}_n(t, t'), \ldots)}{d\varepsilon} \bigg|_{\varepsilon=0} \\
= \int \lim_{n \to \infty} \left( \frac{\partial g}{\partial x} h_n(t, t') + \frac{\partial g}{\partial x} \dot{h}_n(t, t') + \ldots \right) dt' \\
= \int \lim_{n \to \infty} h_n(t, t') \left( \frac{\partial g}{\partial x} - \frac{d}{dt'} \frac{\partial g}{\partial \dot{x}} + \ldots \right) dt' \\
= \int \left( \frac{\partial g}{\partial x} - \frac{d}{dt} \frac{\partial g}{\partial \dot{x}} + \ldots \right) \delta(t - t') dt' \\
= \frac{\partial g}{\partial x} - \frac{d}{dt} \frac{\partial g}{\partial \dot{x}} + \ldots
\]

A convenient shorthand notation for this procedure is

\[
\frac{\delta f[x(t)]}{\delta x(t)} = \frac{\delta}{\delta x(t)} \int g(x(t'), \dot{x}(t'), \ldots) \, dt' \\
= \int \frac{\delta g}{\delta x(t)} \frac{\delta x(t')}{dt'} dt' \\
= \int \frac{\delta g}{\delta x(t')} \delta(t - t') dt' \\
= \frac{\partial g}{\partial x} - \frac{d}{dt} \frac{\partial g}{\partial \dot{x}} + \ldots
\]

One advantage of treating variations in this more formal way is that we can equally well apply the technique to classical field theory.
8.3 Field equations as functional derivatives

We can vary field actions in the same way, and the results make sense directly. Consider varying the scalar field action:

\[ S = \frac{1}{2} \int (\partial^a \varphi \partial_\alpha \varphi - m^2 \varphi^2) \, d^4x \]  

(1071)

with respect to the field \( \varphi \). Setting the functional derivative of \( S \) to zero, we have

\[ 0 = \frac{\delta S[\varphi]}{\delta \varphi(x)} \]

(1072)

\[ = \frac{1}{2} \frac{\delta}{\delta \varphi(x)} \int (\partial^a \varphi \partial_\alpha \varphi - m^2 \varphi^2) \, d^4x' \]

(1073)

\[ = \int \left( \partial^a \varphi \frac{\partial}{\partial x^\alpha} \delta \varphi(x') - m^2 \varphi \frac{\partial}{\partial \varphi(x)} \delta \varphi(x) \right) d^4x' \]

(1074)

\[ = \int \left( -\partial_\alpha \partial^\alpha \varphi - m^2 \varphi \right) \delta \varphi(x') d^4x' \]

(1075)

\[ = \int \left( -\partial_\alpha \partial^\alpha \varphi - m^2 \varphi \right) \delta^3(x' - x) d^3x' \]

(1076)

\[ = -\Box \varphi - m^2 \varphi \]

(1077)

and we have the field equation.

With this new tool at our disposal, we turn to quantization. To see the sort of question we face, suppose we have a function \( f(x(t), x(t + a)) \), which depends nonlocally on \( x(t) \). If we want to find \( \frac{df}{dt} \) at any point \( t_0 \), ordinary differentiation will suffice. We can just let \( y = x(t + a) \), and treat \( f \) as a function of two variables, \( f(x(t), y(t)) \), and use the multivariable chain rule:

\[ \left. \frac{df}{dt} \right|_{t=t_0} = \frac{\partial f}{\partial x} \frac{dx}{dt} \bigg|_{t=t_0} + \frac{\partial f}{\partial y} \frac{dy}{dt} \bigg|_{t=t_0} \]

(1078)

\[ = \frac{\partial f}{\partial x} \bigg|_{t=t_0} + \frac{\partial f}{\partial y} \bigg|_{t=t_0+a} \]

(1079)

This method generalizes to any finite nonlocal dependence. Where we run into the real issue is when the nonlocality is uncountable, as in

\[ S[x(t)] = \int f(x(t)) \, dt \]  

(1080)
To define a derivative of a functional, we need to displace each of an infinite number of independent variables, and sum over their collective changes. Let’s write it as a continuum limit of the discrete case. Let \( f = f(x(t_0 + t_i)) \) for \( i = 1, 2, \ldots, N \). Then defining \( x_i = x(t_i) \),

\[
\frac{df}{dt} \bigg|_{t=t_0} = \sum \frac{\partial f}{\partial x_i} \frac{dx_i}{dt} \bigg|_{t=t_0} \rightarrow \int \frac{\partial f}{\partial x(t)} \frac{dx(t)}{dt} \bigg|_{t=t_0} dt
\]

(1081)

The first problem to address is what the evaluation at \( t_0 \) means. Our functional \( f[x(t)] \) depends on \( x \) at all \( t \), so in setting \( t = t_0 \), what we’re really doing is evaluating at \( x(t) \). It may make more sense to work from the definition of a derivative. Let’s displace \( x(t) \) by an amount \( h \), possibly different at each \( t \), so that denoting the functional derivative by \( \frac{\delta f}{\delta x(t)} \), we want

\[
\frac{\delta f}{\delta x(t)} = \lim_{h(t) \to 0} \frac{f[x(t) + h(t)] - f[x(t)]}{h(t)}
\]

(1082)

The only problem is achieving a uniform convergence. We need to insure that \( h(t) \) tends to zero at all values of \( t \) at once. We can do this by writing \( \varepsilon h(t) \) and letting \( \varepsilon \to 0 \), simultaneously demanding that the resulting expression hold for all choices of \( h(t) \). Then we have

\[
\frac{\delta f}{\delta x(t)} = \lim_{\varepsilon \to 0} \frac{f[x(t) + \varepsilon h(t)] - f[x(t)]}{\varepsilon h(t)}
\]

(1083)

But this can be reformulated as an ordinary derivative. All we have to do is let \( f[x(t)] \) be replaced by

\[
f(\varepsilon, [x]) \equiv f[x(t) + \varepsilon h(t)]
\]

(1084)

The notation on the left is intended to convey that \( f \) is simultaneously a function of \( \varepsilon \) and a functional of \( x(t) \). Now we have

\[
\frac{\delta f}{\delta x(t)} = \lim_{\varepsilon \to 0} \frac{f[x(t) + \varepsilon h(t)] - f[x(t)]}{\varepsilon h(t)}
\]

(1085)

\[
= \lim_{\varepsilon \to 0} \frac{f(\varepsilon, [x(t)]) - f(0, [x(t)])}{\varepsilon h(t)}
\]

(1086)

\[
= \frac{1}{h(t)} \frac{df(\varepsilon, [x])}{d\varepsilon} \bigg|_{\varepsilon=0}
\]

(1087)
Let’s try this and see what happens. Suppose \( f[x] \) is path length in the \( xt \)-plane,

\[
f[x] = \int \sqrt{dx^2 + dt^2} = \int dt \sqrt{1 + \left(\frac{dx}{dt}\right)^2}
\]

(1088)  
(1089)

Then let

\[
f(\varepsilon, [x]) = \int dt \sqrt{1 + \left(\frac{d(x + \varepsilon h(t))}{dt}\right)^2}
\]

(1090)

and compute

\[
\frac{\delta f}{\delta x(t)} = \left. \frac{1}{h(t)} \frac{df(\varepsilon, [x])}{d\varepsilon} \right|_{\varepsilon=0}
\]

(1091)

\[
= \frac{1}{h(t)} \frac{d}{d\varepsilon} \int dt \sqrt{1 + \left(\frac{d(x + \varepsilon h(t))}{dt}\right)^2} \bigg|_{\varepsilon=0}
\]

(1092)

\[
= \frac{1}{2h(t)} \int dt \left(1 + \left(\frac{d(x + \varepsilon h(t))}{dt}\right)^2\right)^{-1/2} \left(2(x + \varepsilon h(t)) \frac{dh(t)}{dt}\right) \bigg|_{\varepsilon=0}
\]

(1093)

\[
= \frac{1}{h(t)} \int dt \frac{dh(t)}{dt} \frac{x(t)}{\sqrt{1 + \left(\frac{dx}{dt}\right)^2}}
\]

(1094)

\[
= -\frac{1}{h(t)} \int dt \ h(t) \frac{d}{dt} \frac{x(t)}{\sqrt{1 + \left(\frac{dx}{dt}\right)^2}}
\]

(1095)

Now, this has to be true for any choice of \( h(t) \), which is only the case if

\[
\frac{d}{dt} \frac{x(t)}{\sqrt{1 + \left(\frac{dx}{dt}\right)^2}} = 0
\]

(1096)

This is fine, but we need a systematic way to get rid of the arbitrary function, \( h(t) \). Suppose we can find a function \( h(t) \) such that

\[
\frac{1}{h(t)} \int dt \ h(t) \ f(t) = f(t)
\]

(1097)
for any choice of \( f(t) \). This would allow us to extract the consequence we want without making the argument each time. Try to solve:

\[
\int dt \ h(t) \ f(t) = f(t)h(t) \tag{1098}
\]

\[
h f = f' h + f \ h' \tag{1099}
\]

for all \( f \). Therefore,

\[
\int dt \ h(t) \ f(t) = f(t)h(t) \tag{1100}
\]

\[
0 = \frac{f'}{f} + \frac{h' - h}{h} \tag{1101}
\]

and this is impossible. We need a new thing.

Functional differentiation is required whenever the function we wish to differentiate is uncountably nonlocal.

may be developed formally by defining it as ordinary differentiation of a one-parameter family of functionals. However, we’ll take a bit of a shortcut to the main results. Suppose we want to take the limit of a derivative as the number of degrees of freedom becomes infinite. For example, consider the replacement,

\[
\frac{\partial H}{\partial q^i} \to \frac{\delta H}{\delta \varphi(x)} \tag{1102}
\]

We can handle a derivative like this using a few properties of differentiation.

### 8.4 The minimal condition for integrability

Suppose we look at the integrability of the action functional directly. That is, we consider

\[
S [x] = \int L dt
\]

\[
= \int (p_i dx^i - H dt)
\]

for a fixed initial point, and ask for path independence of the integral. Then, since the integral to a general point \( x \) must be independent of path,
we require

\[ 0 = \oint (p_i \, dx^i - H \, dt) \]

\[ = \int \int d (p_i \, dx^i - H \, dt) \]

or

\[ 0 = d (p_i \, dx^i - H \, dt) \]

As above, this implies

\[ 0 = \left( dp_i + \frac{\partial H}{\partial x_i} \, dt \right) \wedge \left( dx_i - \frac{\partial H}{\partial p_i} \, dt \right) \]

The most general solution to this equation is for the two 1-forms to be proportional:

\[ dp_i + \frac{\partial H}{\partial x_i} \, dt = \lambda \left( dx_i - \frac{\partial H}{\partial p_i} \, dt \right) \]

\[ dp_i = -\frac{\partial H}{\partial x_i} \, dt + \lambda \left( dx_i - \frac{\partial H}{\partial p_i} \, dt \right) \] (1103)

Taking the exterior derivative again, we have

\[ A = \frac{\partial^2 H}{\partial x \partial x^i} dx^j \wedge dt + \frac{\partial^2 H}{\partial p_j \partial x^i} dp_j \wedge dt \]

\[ = d \lambda \wedge \left( dx_i - \frac{\partial H}{\partial p_i} \, dt \right) - \lambda \frac{\partial^2 H}{\partial x^j \partial p_i} dx^j \wedge dt \] (1104)

\[ + \frac{\partial^2 H}{\partial p_j \partial p_i} dp_j \wedge dt \] (1105)

Expanding \( d \lambda \),

\[ d \lambda \wedge \left( dx_i - \frac{\partial H}{\partial p_i} \, dt \right) = \left( \frac{\partial \lambda}{\partial x_i} \, dx_i + \frac{\partial \lambda}{\partial p_i} \, dp_i + \frac{\partial \lambda}{\partial t} \, dt \right) \wedge \left( dx_i - \frac{\partial H}{\partial p_i} \, dt \right) \]

\[ = \frac{\partial \lambda}{\partial x_j} \, dx_j \wedge dx_i - \frac{\partial H}{\partial p_i} \frac{\partial \lambda}{\partial x_j} \, dx_j \wedge dt + \frac{\partial \lambda}{\partial p_j} \, dp_j \wedge dx_i \]

\[ - \frac{\partial H}{\partial p_i} \frac{\partial \lambda}{\partial p_j} \, dp_j \wedge dt + \frac{\partial \lambda}{\partial t} \, dt \wedge dx_i \] (1106)
Equating corresponding terms, we must have
\[
\frac{\partial \lambda}{\partial x_j} dx_j \wedge dx_i = 0
\]
\[
- \frac{\partial \lambda}{\partial t} dx_i \wedge dt - \frac{\partial H}{\partial p_i} \frac{\partial \lambda}{\partial x_j} dx_j \wedge dt = \frac{\partial^2 H}{\partial x^i \partial x^j} dx^j \wedge dt + \lambda \frac{\partial^2 H}{\partial p_j \partial p_i} dx^j \wedge dt
\]
\[
\frac{\partial \lambda}{\partial p_j} dp_j \wedge dx_i = 0
\]
\[
- \frac{\partial H}{\partial p_i} \frac{\partial \lambda}{\partial x_j} dp_j \wedge dt = \frac{\partial^2 H}{\partial p_j \partial x^i} dp_j \wedge dt + \lambda \frac{\partial^2 H}{\partial p_j \partial p_i} dp_j \wedge dt
\]
The first and third imply \( \lambda = \lambda(t) \), leaving
\[
- \frac{\partial \lambda}{\partial t} = \frac{\partial^2 H}{\partial x^j \partial x^i} + \lambda \frac{\partial^2 H}{\partial x^i \partial p_i}
\]
\[
0 = \frac{\partial^2 H}{\partial p_j \partial x^i} + \lambda \frac{\partial^2 H}{\partial p_j \partial p_i}
\]
Let’s work on these:
\[
x_0 \frac{d\lambda(0)}{dt} - x^i \frac{d\lambda(t)}{dt} = \frac{\partial H}{\partial x^i} + \lambda \frac{\partial H}{\partial p_i}
\]
\[
= \lambda \dot{x}_i - \dot{p}_i
\]
\[
0 = \frac{\partial}{\partial p_j} \left( \frac{\partial H}{\partial x^i} + \lambda \frac{\partial H}{\partial p_i} \right)
\]
\[
= \frac{\partial}{\partial p_j} (\lambda \dot{x}_i - \dot{p}_i)
\]
\[
\dot{p}_i + \frac{\partial H}{\partial x_i} = \lambda \left( \dot{x}_i - \frac{\partial H}{\partial p_i} \right) = \lambda \dot{x}_i - \lambda \frac{\partial H}{\partial p_i} \tag{1107}
\]
\[
\frac{\partial H}{\partial x_i} + \lambda \frac{\partial H}{\partial p_i} = \lambda \dot{x}_i - \dot{p}_i \tag{1108}
\]
or,
\[
- \frac{\partial \lambda}{\partial t} = \frac{\partial}{\partial x^j} (\lambda \dot{x}_i - \dot{p}_i)
\]
\[
0 = \frac{\partial}{\partial p_j} (\lambda \dot{x}_i - \dot{p}_i)
\]
210
or,

\[
\begin{align*}
\frac{\partial \lambda}{\partial t} &= \frac{\partial}{\partial x^i} \left( \frac{\partial H}{\partial x^i} + \lambda \frac{\partial H}{\partial p_i} \right) \\
0 &= \frac{\partial}{\partial p_j} \left( \frac{\partial H}{\partial x^i} + \lambda \frac{\partial H}{\partial p_i} \right)
\end{align*}
\]

Therefore, the combination

\[
\left( \frac{\partial}{\partial x^i} + \lambda \frac{\partial}{\partial p_i} \right) H
\]

is a function of \(x^i\) and \(t\) only:

\[
f(x^i, t) = \frac{\partial H}{\partial x^i} + \lambda(t) \frac{\partial H}{\partial p_i}
\]

Perhaps we can find a canonical transformation for which

\[
\frac{\partial}{\partial x^i} + \lambda \frac{\partial}{\partial p_i} = \frac{\partial}{\partial \pi_i}
\]

### 8.5 Dualing Brackets

We can define the dual of a Poisson bracket. In this form, we get the same result but derive the bracket from a mapping on 2-forms. Let \(\xi^A\) be canonical coordinates and let \(f\) be a dynamical variable. Then consider the differential form

\[
df = \frac{\partial f}{\partial \xi^A} d\xi^A
\]

The wedge product of two such objects is

\[
df \wedge dg = \frac{\partial f}{\partial \xi^A} d\xi^A \wedge \frac{\partial g}{\partial \xi^B} d\xi^B
\]

\[
= \frac{1}{2} \left( \frac{\partial f}{\partial \xi^A} \frac{\partial g}{\partial \xi^B} - \frac{\partial g}{\partial \xi^B} \frac{\partial f}{\partial \xi^A} \right) d\xi^A \wedge d\xi^B
\]

This suggest that we can think of the symplectic form as a mapping from 2-forms to the reals:

\[
\Omega : \Lambda_2 \to R
\]

211
with components defined by

\[
\Omega (\omega) = \Omega (\omega_{AB} d\xi^A \wedge d\xi^B) = \omega_{AB} \Omega (d\xi^A \wedge d\xi^B) = \omega_{AB} \Omega^{AB}
\]

where we have set

\[
\Omega (d\xi^A \wedge d\xi^B) = \Omega^{AB}
\]

If \(\xi^A\) is canonical, so that

\[
\{\xi^A, \xi^B\} = \Omega^{CD} \frac{\partial \xi^A}{\partial \xi^C} \frac{\partial \xi^B}{\partial \xi^D}
\]

\[
= \Omega^{AB}
\]

\[
= \begin{pmatrix}
-1 & 1
\end{pmatrix}
\]

we have

\[
\Omega (\omega) = \omega_{AB} \Omega^{AB} = \frac{1}{2} \omega_{i} d x^{i} - \frac{1}{2} \omega^{i}_{i} = \omega_{i}
\]

where we expand \(\omega\) in \((x^i, p_i)\) coordinates as

\[
\omega = \frac{1}{2} \omega_{ij} dx^i \wedge dx^j + \frac{1}{2} \omega^i_{\ j} dp_i \wedge dx^j + \frac{1}{2} \omega^i_{\ j} dx^i \wedge dp_j + \frac{1}{2} \omega^{ij} dp_i \wedge dp_j
\]

Antisymmetry requires

\[
\omega_{ij} = -\omega_{ji}
\]
\[
\omega^i_{\ j} = -\omega^i_{\ j}
\]
\[
\omega^{ij} = -\omega^{ji}
\]

Finally, we define the bracket as

\[
\{f, g\} \equiv \Omega (df \wedge dg)
\]
\[
\begin{align*}
\Omega \left( \frac{\partial f}{\partial \xi^A} d\xi^A \wedge \frac{\partial g}{\partial \xi^B} d\xi^B \right) \\
= \frac{\partial f}{\partial \xi^A} \frac{\partial g}{\partial \xi^B} \Omega (d\xi^A \wedge d\xi^B) \\
= \frac{\partial f}{\partial \xi^A} \frac{\partial g}{\partial \xi^B} \Omega^{AB}
\end{align*}
\]

Let’s look at \( d\Omega (d g \wedge d h) \):

\[
\begin{align*}
d\Omega (d g \wedge d h) &= \Omega^{AB} \left( \frac{\partial^2 f}{\partial \xi^C \partial \xi^A} \frac{\partial g}{\partial \xi^B} + \frac{\partial f}{\partial \xi^A} \frac{\partial^2 g}{\partial \xi^C \partial \xi^B} \right) d\xi^C \\
&= \Omega^{AB} \left( d \left( \frac{\partial f}{\partial \xi^A} \right) \frac{\partial g}{\partial \xi^B} + \frac{\partial f}{\partial \xi^A} d \left( \frac{\partial g}{\partial \xi^B} \right) \right)
\end{align*}
\]

Now consider the Jacobi identity. We have

\[
\{ f, \{ g, h \} \} = \Omega (df \wedge d\Omega (dg, dh))
\]

\[
= \Omega \left( df \wedge d \left( \frac{\partial f}{\partial \xi^A} \frac{\partial g}{\partial \xi^B} \Omega^{AB} \right) \right)
\]

\[
= \Omega \left( df \wedge d \left( \frac{\partial f}{\partial \xi^A} \frac{\partial g}{\partial \xi^B} \Omega^{AB} \right) \right)
\]

\[
= \Omega \left( \frac{\partial f}{\partial \xi^A} d\xi^A \wedge \frac{\partial g}{\partial \xi^B} d\xi^B \right)
\]

\[
= \frac{\partial f}{\partial \xi^A} \frac{\partial g}{\partial \xi^B} \Omega (d\xi^A \wedge d\xi^B)
\]

\[
= \frac{\partial f}{\partial \xi^A} \frac{\partial g}{\partial \xi^B} \Omega^{AB}
\]

\[
213
\]
8.6 Canonical transformations

Now we return to our earlier claim that transformations certain fundamental Poisson brackets preserve Hamilton’s equations and preserve all Poisson brackets. Specifically, we show that a transformation from one set of phase space coordinates \((x_i, \pi_i)\) to another \((q_i, p_i)\) as canonical if and only if it preserves the fundamental Poisson brackets

\[
[x_i, x_j]_{q\pi} = [p_i, p_j]_{q\pi} = 0 \quad (1109)
\]

\[
[p_i, x_j]_{q\pi} = -[x_i, p_j]_{q\pi} = \delta_{ij} \quad (1110)
\]

Here the subscript on the bracket, \([\cdot]_{q\pi}\), means that the partial derivatives defining the bracket are taken with respect to \(q_i\) and \(\pi_i\). We claim that brackets \([f, g]_{qp}\) taken with respect to the new variables \((q_i, p_i)\) are identical to those \([f, g]_{x\pi}\) with respect to \((x_i, \pi_i)\) if and only if the transformation is canonical. Next, we seek necessary and sufficient conditions for a transformation to be canonical.

Landau and Lifshitz make the following argument. Since Hamilton’s equations are preserved under canonical transformations and the time evolution of any dynamical variable is independent of the phase space coordinates,

\[
\frac{df}{dt} = \frac{\partial f}{\partial t} + [H, f]_{p,q} = \frac{\partial f}{\partial t} + [H', f]_{x,x} \quad (1111)
\]

and therefore

\[
[H, f]_{p,q} = [H', f]_{x,x} \quad (1112)
\]

where the transformation from \((p_i, q_i)\) to \((x_i, \pi_i)\) is canonical. However, since time enters as a parameter only, it is sufficient to consider time-independent transformations only, in which case, \(H = H'\)

\[
[H, f]_{p,q} = [H, f]_{x,x} \quad (1113)
\]

Furthermore, since we have placed no restrictions on what functions of phase space may be used for the Hamiltonian, \(H\), we may write \(H = g\), with \(g\) arbitrary. Therefore

\[
[g, f]_{p,q} = [g, f]_{x,x} \quad (1114)
\]

for all functions \(f\) and \(g\). Thus, canonical transformations preserve Poisson brackets. Finally, choosing \(f\) and \(g\) to be any of the variables \((\pi_i, x_i)\), we
The fundamental Poisson brackets are therefore a necessary property of any canonical transformation.

However, there is a problem with this reasoning. We arrived at the equality of the brackets $[H,f]_{p,q}$ and $[H,f]_{\pi,x}$ by our claim that the time evolution of $f$ is independent of the phase space coordinates. But that time evolution is not independent of the choice of Hamiltonian so that when we replace $f$ and $g$ in the final step with $\pi_i$ and $x_i$, we are in fact specializing the Hamiltonian to be a component of $\pi_i$ or $x_i$. The correct conclusion from the argument is only that if the Hamiltonian is $\pi$ then the time evolution of $x_i$ is preserved.

Let’s try the following argument instead. We’ll compute the brackets we desire directly, using a generating function $f = f(x_i, q_i, t)$. Using eq.(957) (with $\lambda = 1$) we have

$$[\pi_i, x_j]_{p,q} = \frac{\partial \pi_i}{\partial q_k} \frac{\partial x_j}{\partial p_k} - \frac{\partial x_i}{\partial q_k} \frac{\partial \pi_j}{\partial p_k}$$

(1118)

The second term vanishes because $\frac{\partial f}{\partial q_i}$ is a function of $(x_i, q_i)$ only. Thus,

$$[\pi_i, x_j]_{p,q} = - \frac{\partial^2 f}{\partial q_i \partial q_j} \frac{\partial x_j}{\partial p_k}$$

(1119)

Finally, we use the independence of the $q_i$ and $p_j$, written in terms of $x_i$ and $q_j$, to write:

$$\delta_{ij} = \frac{\partial \pi_i}{\partial \pi_j} = \left( \frac{\partial x_i}{\partial x_k} \frac{\partial \pi_j}{\partial x_k} + \frac{\partial q_i}{\partial q_k} \frac{\partial \pi_j}{\partial q_k} \right)$$

(1121)

$$\delta_{ij} = \frac{\partial p_i}{\partial p_j} = \left( \frac{\partial x_i}{\partial p_k} \frac{\partial p_j}{\partial x_k} + \frac{\partial q_i}{\partial q_k} \frac{\partial p_j}{\partial q_k} \right)$$

(1122)
Substituting for \( \pi_i \) in terms of \( f \) we have

\[
\delta_{ij} = \frac{\partial \pi_i}{\partial \pi_j} = -\frac{\partial x_k}{\partial \pi_j} \frac{\partial^2 f}{\partial x_k \partial q_i} - \frac{\partial q_k}{\partial \pi_j} \frac{\partial^2 f}{\partial q_k \partial q_i} - \frac{\partial q_k}{\partial \pi_j} \frac{\partial^2 f}{\partial q_k \partial q_i}, \tag{1123}\]

\[
= -\frac{\partial q_k}{\partial \pi_j} \frac{\partial^2 f}{\partial q_k \partial q_i}, \tag{1124}\]

\[
\delta_{ij} = \frac{\partial p_i}{\partial p_j} = \frac{\partial x_k}{\partial p_j} \frac{\partial p_i}{\partial x_k} + \frac{\partial q_k}{\partial p_j} \frac{\partial p_i}{\partial q_k} + \frac{\partial q_k}{\partial p_j} \frac{\partial^2 f}{\partial q_k \partial x_i} - \frac{\partial q_k}{\partial p_j} \frac{\partial^2 f}{\partial q_k \partial x_i}, \tag{1125}\]

\[
= -\frac{\partial q_k}{\partial \pi_j} \frac{\partial^2 f}{\partial q_k \partial q_i}, \tag{1126}\]

\[
\delta_{ij} = \frac{\partial p_i}{\partial p_j} = \left( \frac{\partial x_k}{\partial p_j} \frac{\partial p_i}{\partial x_k} + \frac{\partial q_k}{\partial p_j} \frac{\partial p_i}{\partial q_k} \right), \tag{1127}\]

where the last step follows by the independence of \( x_k \) and \( \pi_j \).

\[
p_i = \frac{\partial f}{\partial x_i}, \tag{1128}\]

\[
\pi_i = -\frac{1}{\lambda} \frac{\partial f}{\partial q_i}, \tag{1129}\]

\[
H' = \frac{1}{\lambda} \left( H + \frac{\partial f}{\partial t} \right), \tag{1130}\]

The authors assert without proof that these conditions are also sufficient, though the proof is not difficult. Indeed, suppose we have the Poisson bracket, \([f, g]_{qp} \) with respect to \((q_i, p_i)\) of any two functions \( f \) and \( g \). Then compute

\[
[f, g]_{xp} = \frac{\partial g}{\partial x_i} \frac{\partial f}{\partial \pi_i} - \frac{\partial g}{\partial \pi_i} \frac{\partial f}{\partial x_i}, \tag{1132}\]

\[
= \left( \frac{\partial g}{\partial q_j} \frac{\partial q_j}{\partial x_i} + \frac{\partial g}{\partial p_j} \frac{\partial p_j}{\partial x_i} \right) \left( \frac{\partial f}{\partial q_k} \frac{\partial q_k}{\partial \pi_i} + \frac{\partial f}{\partial p_k} \frac{\partial p_k}{\partial \pi_i} \right) \left( \frac{\partial f}{\partial q_k} \frac{\partial q_k}{\partial x_i} + \frac{\partial f}{\partial p_k} \frac{\partial p_k}{\partial x_i} \right), \tag{1133}\]

\[
- \left( \frac{\partial g}{\partial q_j} \frac{\partial q_j}{\partial \pi_i} + \frac{\partial g}{\partial p_j} \frac{\partial p_j}{\partial \pi_i} \right) \left( \frac{\partial f}{\partial q_k} \frac{\partial q_k}{\partial x_i} + \frac{\partial f}{\partial p_k} \frac{\partial p_k}{\partial x_i} \right), \tag{1134}\]
\[
\begin{align*}
&= \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial q_k} \frac{\partial q_j}{\partial x_i} \frac{\partial q_k}{\partial \pi_i} + \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial p_j} \frac{\partial q_k}{\partial \pi_i} \\
&\quad + \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial q_k} \frac{\partial p_j}{\partial \pi_i} + \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial p_j} \frac{\partial p_k}{\partial \pi_i} \\
&\quad - \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial q_k} \frac{\partial q_k}{\partial \pi_i} \frac{\partial x_i}{\partial \pi_i} - \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial p_j} \frac{\partial p_k}{\partial \pi_i} \\
&\quad - \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial p_j} \frac{\partial p_k}{\partial \pi_i} \frac{\partial q_k}{\partial \pi_i} \frac{\partial x_i}{\partial \pi_i} \\
&\quad - \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial p_j} \frac{\partial p_k}{\partial \pi_i} \frac{\partial x_i}{\partial \pi_i} \\
&\quad - \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial p_j} \frac{\partial p_k}{\partial \pi_i} \frac{\partial x_i}{\partial \pi_i} \\
&\quad = \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial q_k} \left( \frac{\partial q_j}{\partial x_i} \frac{\partial q_k}{\partial \pi_i} - \frac{\partial q_j}{\partial \pi_i} \frac{\partial x_i}{\partial \pi_i} \right) \\
&\quad + \frac{\partial g}{\partial p_j} \frac{\partial f}{\partial q_k} \left( \frac{\partial p_j}{\partial x_i} \frac{\partial q_k}{\partial \pi_i} - \frac{\partial p_j}{\partial \pi_i} \frac{\partial x_i}{\partial \pi_i} \right) \\
&\quad + \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial p_k} \left( \frac{\partial q_j}{\partial x_i} \frac{\partial p_k}{\partial \pi_i} - \frac{\partial q_j}{\partial \pi_i} \frac{\partial x_i}{\partial \pi_i} \right) \\
&\quad + \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial p_k} \left( \frac{\partial p_j}{\partial x_i} \frac{\partial p_k}{\partial \pi_i} - \frac{\partial p_j}{\partial \pi_i} \frac{\partial x_i}{\partial \pi_i} \right) \\
&\quad = \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial q_k} \left[ q_j, q_k \right]_{x\pi} + \frac{\partial g}{\partial p_j} \frac{\partial f}{\partial q_k} \left[ p_j, q_k \right]_{x\pi} \\
&\quad + \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial q_k} \left[ q_j, p_k \right]_{x\pi} + \frac{\partial g}{\partial p_j} \frac{\partial f}{\partial q_k} \left[ p_j, p_k \right]_{x\pi}
\end{align*}
\]

Therefore,
\[
\left[ f, g \right]_{x\pi} = \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial q_k} \left[ q_j, q_k \right]_{x\pi} + \frac{\partial g}{\partial p_j} \frac{\partial f}{\partial q_k} \left[ p_j, q_k \right]_{x\pi}
\]
\[
+ \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial q_k} \left[ q_j, p_k \right]_{x\pi} + \frac{\partial g}{\partial p_j} \frac{\partial f}{\partial q_k} \left[ p_j, p_k \right]_{x\pi}
\]
\]

Comparing this expression term by term with the Poisson bracket of \( f \) and \( g \) in the new variables,

\[
\left[ f, g \right]_{qp} = \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial q_k} - \frac{\partial g}{\partial q_k} \frac{\partial f}{\partial q_j}
\]

we see that the two brackets will be equal if and only if

\[
0 = \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial q_k} \left[ q_j, q_k \right]_{x\pi}
\]
\[ \frac{\partial g}{\partial p_k} \frac{\partial f}{\partial q_k} = \frac{\partial g}{\partial p_j} \frac{\partial f}{\partial q_k} [p_j, q_k]_{x\pi} \]  
\[ -\frac{\partial g}{\partial q_k} \frac{\partial f}{\partial p_k} = \frac{\partial g}{\partial q_j} \frac{\partial f}{\partial p_k} [q_j, p_k]_{x\pi} \]  
\[ 0 = \frac{\partial g}{\partial p_j} \frac{\partial f}{\partial p_k} [p_j, p_k]_{x\pi} \]

and therefore if and only if

\[ [q_j, q_k]_{x\pi} = 0 = [p_j, p_k]_{x\pi} \]  
\[ [p_j, q_k]_{x\pi} = \delta_{jk} = -[q_j, p_k]_{x\pi} \]

Since \( f \) and \( g \) were arbitrary, we have shown that any canonical transformation will preserve all Poisson brackets. In particular, replacing \( f \) by \( H \) and \( g \) by any of the coordinate functions \( (x_i, \pi_i) \), we see that Hamilton’s equations are preserved.