Essentials of Classical Mechanics

Ryan D. Martin

December 1, 2015

Contents

1 Formalism and generalized coordinates

1.1 Analytic and vector mechanics

In introductory physics, one learns to describe the time evolution of a system using Newton's second law of motion. This approach to mechanics is called "vector mechanics". Newton's second law:

$$\vec{F} = m\vec{a} \tag{1.1}$$

is a vector equation that is true for each component of the vectors for force and acceleration.

By contrast, in "analytic mechanics", the formalism is developed to describe the time-evolution of a system using a single equation involving only scalar quantities. The approach will be seen to be equivalent to the vector mechanics approach. The vector mechanics formalism generally suffers from limitations that make it difficult to apply in complicated situations. For example, the vector equations are not straightforward in non-Cartesian coordinate systems, which may be more suited for certain problems (motion on the surface of a sphere, orbital motion, etc.).

In general, a "theory" of classical mechanics should be able to describe the future and past of a system, given knowledge of its current state and the forces involved. We will see that it is sufficient to describe the current state of a system by specifying the positions of all of the particles that are involved. Furthermore, given the particular forces (or potential energies involved), knowing the positions and velocities of all particles is sufficient in order to describe the entire history (past and future) of the system. The analytic treatment of mechanics is formulated using the positions and velocities of the particles involved (as opposed to the vectorial approach that deals with accelerations).

EXAMPLE 1-1: The Compound Pendulum is a typical example of a simple system that is awkward to describe using vectorial analysis, but straightforward using analytic mechanics.

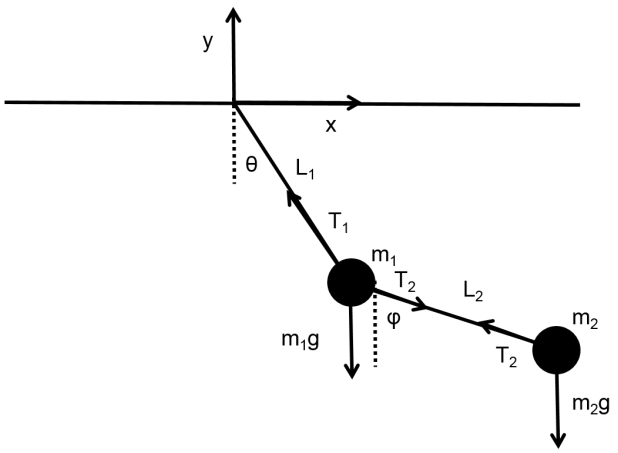

Figure 1.1: Vectorial analysis of the compound pendulum

1.2 Definitions and notation

1.2.1 Particle, position, velocity, acceleration

We will define a "particle" to be a point-like object with a mass. The location in space of such a particle is described by a position vector, $\vec{r}(t)$, that can change as a function of time. The components of the position vector are the "coordinates" of the particle. The rate of change of the position vector is called "velocity" (and that of velocity, "acceleration"). If one knows the values of the position and velocity vectors at some instant in time, the whole "path" of the particle (backwards and forwards in time and space) can be deduced. The equations that relate position, velocity, and acceleration are called "equations of motion". Newton's second law gives the equation of motion for $\vec{r}(t)$ in differential form:

$$\vec{F} = m\vec{a} = m\frac{d^2}{dt^2}\vec{r} = m\ddot{\vec{r}}$$
$$\ddot{\vec{r}} = \frac{1}{m}\vec{F} \tag{1.2}$$

where we have introduced the use of "dots" to signify derivatives with respect to time:

$$\vec{F} = m\vec{a} = m\frac{d^2}{dt^2}\vec{r} = m\ddot{\vec{r}}$$
$$\dot{x} \equiv \frac{d}{dt}x$$
$$\ddot{x} \equiv \frac{d^2}{dt^2}x \tag{1.3}$$

1.2.2 Rigid body

A "rigid body" is an object made from many particles that are fixed in position relative to each other. The rigid body can thus be treated without knowing the specifics of the particles that make it up. Often, it is sufficient to describe a rigid body by the position of its center of mass, its angles of rotation about three axes, its total mass, and its moment of inertia tensor.

1.2.3 Degrees of freedom

The "number of degrees of freedom" is the number of scalar quantities that are required to specify the state of a system. For a particle in three dimensional space, one requires 3 coordinates to specify the location of a particle. For a rigid object, one requires 6 degrees of freedom (3 coordinates of the center of mass and 3 rotation angles).

1.2.4 System

A "system" is an ensemble of particles and rigid bodies that one wants to describe. A system will generally have 3N+6M degrees of freedom, if there are N particles and M rigid bodies.

1.2.5 Constraints

When there are "constraints" applied to the coordinates of the particles, the number of degrees of freedom is reduced. For example, if 2 particles are connected by a rigid mass-less bar of length, l, the number of degrees of freedom is 5. The constraint of a rigid bar can be written as 1 equation of constraint, requiring that the distance between the two particles remain constant.

$$|\vec{r_1} - \vec{r_2}|^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 = l^2$$
$$\therefore l^2 - (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 = 0 \tag{1.4}$$

we call this a "constraint" equation. In general, the number of degrees of freedom is reduced by the number of constraint equations that exist. For two masses constrained by a bar, we have 5 degrees of freedom. This makes sense; we can completely specify the position of the two masses by the 3 coordinates of the first mass and 2 angles that specify the rotation of the bar.

EXAMPLE 1-2: How many degrees of freedom are required to describe a system of 4 particles connected by 4 rigid massless rods? Write out the constraint equations. What if the particles are constrained to the

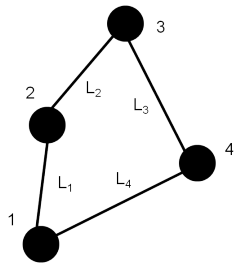

Figure 1.2: Four particles connected by rods

xy-plane?

1.2.6 Types of constraints

Different types of constraints have different properties (and names).

Holonomic Constraint equations that can be written using the coordinates of the particles in the system are called "holonomic". The constraint equation with two masses held by a rigid bar is an example of a holonomic constraint (see eqn. 1.4). A holonomic constraint can depend on time.

EXAMPLE 1-3: Rolling without slipping in 1D. Write out the constraint for a disk of radius r to roll without slipping along the x-axis.

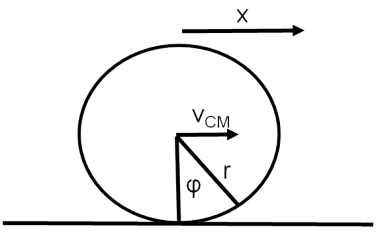

Figure 1.3: A disk rolling without slipping

Non-holonomic Non-holonomic constraint equations cannot be written only in terms of the coordinates, or contain inequalities. For example, the constraint of "rolling without slipping" is a constraint on the instantaneous velocity of the point of contact. The equation is expressed in terms of the velocity rather than the position coordinates (it can be integrated in 1 dimension to get a holonomic constraint, but this does not generalize to 2-dimensions). Another type of non-holonomic constraints are those that contain inequalities such as particles that are constrained to move within a volume (the constraint is on the position vectors, but it is expressed as an inequality). In general, it is difficult to deal with non-holonomic constraints.

EXAMPLE 1-4: The constraint equation for particles constrained to be inside of a sphere of radius R:

EXAMPLE 1-5: Rolling without slipping in 2D. Write out the constraint for a sphere of radius r to roll without slipping in the xy-plane

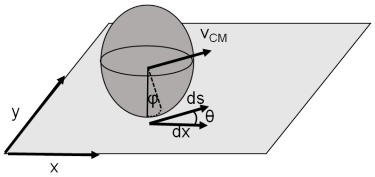

Figure 1.4: A sphere rolling without slipping

Sceleronomous A sceleronomous constraint does not depend explicitly on time (the partial derivative with respect to time is zero).

Rheonomous A rheonomous contraint depends (explicitly) on time (partial derivative with respect to time is non-zero).

1.2.7 Generalized coordinates

In general, one has the freedom to choose the coordinate system (Cartesian, polar, something more exotic) to describe a system. If we have a system with N particles and k holonomic constraints, there will be $3N$ coordinates, but only $3N - k$ of them will be independent (the number of degrees of freedom).

It is often useful to restrict the number of coordinates to describe a system so that there are only as many coordinates as there are degrees of freedom. These coordinates are called "generalized coordinates", and often denoted as the set $\{q_1, q_2, q_3, \ldots, q_{3N-k}\}$. Their time derivatives are called "generalized velocities" $\{\dot{q_1}, \dot{q_2}, \dot{q_3}, \ldots, \dot{q}_{3N-k}\}$. The generalized coordinates should not necessarily be thought of as a set of coordinates in space (like Cartesian or Polar); rather, a set of numbers that describes the state of the system.

Using generalized coordinates, one does not need to think in terms of the coordinates of each particle in the system. That is, the generalized coordinates may not be the position of specific particles, but rather numbers that can describe the ensemble of particles. The system is entirely described by the generalized coordinates. There will exist "transformation equations" to convert the generalized coordinates back to the coordinates of the individual particles in the system:

$$\begin{aligned} x_1 &= x_1(q_1, q_2, q_3, \ldots, q_{3N-k}, t) \\ y_1 &= y_1(q_1, q_2, q_3, \ldots, q_{3N-k}, t) \\ z_1 &= z_1(q_1, q_2, q_3, \ldots, q_{3N-k}, t) \\ &\cdots \\ x_N &= x_N(q_1, q_2, q_3, \ldots, q_{3N-k}, t) \\ y_N &= y_N(q_1, q_2, q_3, \ldots, q_{3N-k}, t) \\ z_N &= z_N(q_1, q_2, q_3, \ldots, q_{3N-k}, t) \end{aligned} \tag{1.5}$$

The generalized coordinates could have different units than length and could, in general, be quite abstract. The transformation equations to the generalized coordinates can depend on time. Similarly, the generalized velocities (and accelerations), can also be obtained from the transformation equations:

$$\begin{aligned} \dot{x}_1 &= \frac{\partial x_1}{\partial q_1}\dot{q}_1 + \cdots + \frac{\partial x_1}{\partial q_{3N-k}}\dot{q}_{3N-k} + \frac{\partial x_1}{\partial t} \\ &\cdots \\ \dot{z}_N &= \frac{\partial z_N}{\partial q_1}\dot{q}_1 + \cdots + \frac{\partial z_N}{\partial q_{3N-k}}\dot{q}_{3N-k} + \frac{\partial z_N}{\partial t} \end{aligned}$$

The partial derivative with respect to time comes from writing $dx = \frac{\partial x_1}{\partial q_1}dq_1 + \cdots + \frac{\partial x_1}{\partial t}dt$ and then dividing out by dt to get $\dot{x}$ on the left hand side.

EXAMPLE 1-6: Write out generalized coordinates to describe the motion of a bead of mass m that is constrained to slide along a mass-less rod of length, L, that is rotating at a constant angular frequency ω in the vertical plane about its pivot point (Figure 1.5). Write out the constraint equations in the form $f_i(x_1, \ldots, x_n) = 0$. Write out the generalized velocities.

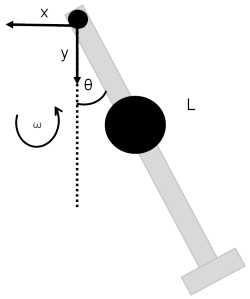

Figure 1.5: Bead that can slide on a rotating rod.

1.2.8 Configuration space

The n generalized coordinates corresponding to the n degrees of freedom of a system can be thought of in terms of an n-dimensional Euclidean space called "configuration space". As the system evolves through time, one can relate this to a path in configuration space. The equations of motion dictate that path. Note that the initial conditions of a system will dictate which path it takes in coordinate space based on the equations of motion. For example, a cannon ball has the same equations of motion regardless of the inclination of the cannon its initial velocity as it exits the cannon. However, depending on the initial conditions of its position and velocity, the cannon ball will trace different trajectories (paths in configuration space).

1.3 Problems

Problem 1-1: Transformations to generalized coordinates
Given the following transformations between the cartesian coordinates ($x_i(t)$,$y_i(t)$,$z_i(t)$) of a particle i and a set of generalized coordinates ($q_i(t)$), write the velocities and accelerations of the cartesian coordinates in terms of the velocities and accelerations of the generalized coordinates. Note that unless specified, other variables should be taken as constants with respect to time, t.
a)

$$\begin{aligned} x_1 &= L\cos(q_1) \\ y_1 &= L\sin(q_1) \\ z_1 &= q_2 \end{aligned}$$

b)

$$\begin{aligned} x_1 &= L\cos(q_1)q_2 \\ y_1 &= L\sin(q_1)q_2 \\ z_1 &= R\sin(\omega t) \end{aligned}$$

c)

$$\begin{aligned} x_1 &= \sqrt{q_1^2 + q_2^2} \\ y_1 &= \tan^{-1}\left(\frac{q_1}{q_2}\right) \end{aligned}$$

d)

$$\begin{aligned} x_1 &= -\frac{1}{2}gt^2 + L\sin q_1 \\ y_1 &= vt + L\cos q_1 \end{aligned}$$

e)

$$\begin{aligned} x_1 &= L\cos(q_1) \\ y_1 &= L\sin(q_1) \\ z_1 &= q_2\cos(\omega t) \\ x_2 &= x_1 + L\cos(q_3) \\ y_2 &= y_1 - L\sin(q_3) \\ z_2 &= z_1 \end{aligned}$$

Problem 1-2: Degrees of freedom and kinetic energy
For the following situations in two dimensional space, give n, the number of degrees of freedom, then choose n generalized coordinates and write out the kinetic energy of the system in terms of the corresponding generalized velocities (start by writing the kinetic energy in Cartesian coordinates $\sum \frac{1}{2}m_i(\dot{x}_i^2 + \dot{y}_i^2)$.
a)Two masses, m, are connected by a massless rigid rod of length L.
b) Three masses, m, that are connected by 2 massless rigid rods of length l and a massless spring and are constrained to move in the plane (see figure)

c) A pendulum consisting of a mass, m, connected to a rigid massless bar of length, L, whose other end is constrained to move downwards with a known velocity, v (see figure)

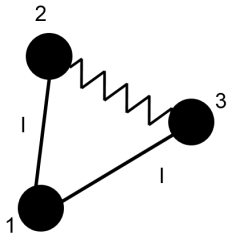

Figure 1.6: Three masses connected by 2 rods and a spring, problem 1-2

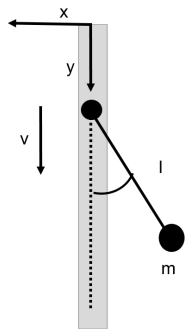

Figure 1.7: Moving pendulum, problem 1-2

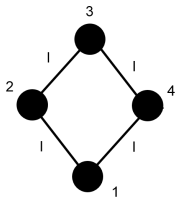

Figure 1.8: Four masses connected by four rods, problem 1-2

d) Four masses, m connected by 4 massless rigid rods of length l, and are constrained to the move in the plane (see figure)

Problem 1-3: Block sliding down a ramp

The figure shows a block of mass, m, sliding down a ramp of length L_1 and a slope given by an angle θ which is connected to a second "launching" ramp of length L_2 with angle ϕ. Assume that the block starts at the top of the first ramp and that the origin is as shown. Furthermore, assume that coefficient of kinetic friction between the block and the ramp is given by μ

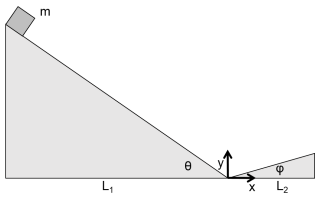

Figure 1.9: Block sliding down ramp„ problem 1-3

a) Draw a free body diagram of the forces on the block, and write the differential equations of motion for

the x and y components of the velocity of the block. Do this for each ramp.
b) Solve the differential equations of motion from part a) (using an initial velocity of zero) to determine where the components of the velocity vector as the block leaves the second ramp.
c) Use the result from part b) to determine the distance from the origin at which the block will land
d) Repeat the problem using conservation of energy to find the point at which the block will land.

Problem 1-4: Disk rolling down a ramp
The block from problem 1-4 is replaced by a disk of radius r, and mass m, that rolls without slipping, and has moment of inertia $I = \frac{1}{2}mr^2$.
a) Draw a free body diagram of the forces and torques on the disk, and write the differential equations of motion for the x and y components of the velocity of the disk, as well as for its angular speed, ω. Do this for each ramp.
b) Solve the differential equations of motion from part a) (using an initial velocity of zero) to determine where the components of the velocity vector and the magnitude of the angular velocity as the disk leaves the second ramp.
c) Use the result from part b) to determine the distance from the origin at which the disk will land
d) Repeat the problem using conservation of energy to find the point at which the disk will land and its angular velocity just before landing.

Problem 1-5: Person on a ladder
The figure shows a person of mass m standing in the middle of a ladder of mass M and length L inclined against a friction-less vertical wall. What is the minimum value for the coefficient of static friction, μ, between the ladder and the ground for the ladder not to slide when inclined at an angle θ?

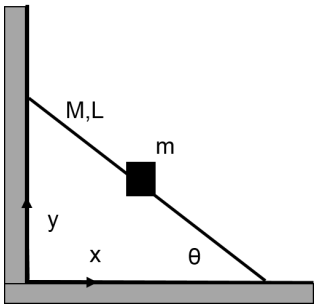

Figure 1.10: Person on a ladder„ problem 1-5

Problem 1-6: Compound pendulum
Consider the compound pendulum in Example 1-1. Use Newton's Laws to do the following:
a) Write out the differential equations of motion for $\ddot{x}_1$, $\ddot{y}_1$, $\ddot{x}_2$, $\ddot{y}_2$
b) Show that the system can be described by the generalized coordinates θ, and ϕ, and write out the differential equations of motion for $\ddot{\theta}$ and $\ddot{\phi}$.
c) Use a computer to solve the differential equations of motion and make plots of $\theta(t)$, and $\phi(t)$ for $t = 0 \ldots 10\,s$. Use L_1=1 m, L_2=0.75 m, m_1=1 kg, m_2=2 kg, and initial conditions at $t = 0$ of $\theta = \frac{\pi}{2}$ and $\phi = 0$.

Problem 1-7: Two masses and two springs
The figure shows two masses, m_1 and m_2, each connected to two springs with spring constants k_1 and k_2. Mass m_1 is constrained to slide without friction along the x-axis, whereas mass m_2 is constrained to move in the vertical direction, constrained by a massless frictionless vertical rod that is attached to m_1. Both springs have a resting length of L. **a)** Write out the differential equations of motion for x_1, x_2, y_1, and y_2.
b) How many degree of freedom, n, are there? Choose n generalized coordinates and write out their differential equations of motion.
c) Write out the total energy of the sytem (kinetic + potential) in terms of the generalized coordinates and velocities.

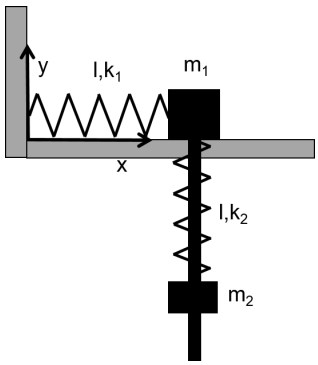

Figure 1.11: Two masses and two springs,, problem 1-7

d) Use a computer to plot the path in the xy plane for mass m_2 for $t = 0 \ldots 10$s given the following values and initial conditions at $t = 0$: $m_1 = 1\,\mathrm{kg}$, $m_2 = 5\,\mathrm{kg}$, $L = 1\,\mathrm{m}$, $k_1 = 10\,\mathrm{N/m}$, $k_2 = 2\,\mathrm{N/M}$, $x_1 = 1.2\,\mathrm{m}$, $v_{1x} = 0\,\mathrm{m/s}$, $y_2 = -0.5\,\mathrm{m}$, $v_{2y} = 0\,\mathrm{m/s}$.

2 Calculus of variations

The calculus of variations is fundamental in understanding the principles of analytic mechanics. The main problems to be considered are those of finding stationary values of functions and of integrals. Technically, the analysis of stationary values of functions is called "calculus", while that of integrals is called "calculus of variations"

2.1 Virtual displacement

The calculus of variations deals with "virtual" displacements, rather than "differential" displacements, and "variations of functions". For example, when considering the time-evolution of some coordinate, $q(t)$, one can write the differential, dq, which is a measure of the infinitesimal change in q for a given change, dt, in time.

A virtual displacement, denoted δq, is a change in $q(t)$, without the corresponding change in time. That is, the value of $q(t)$ is infinitesimally displaced from the value that it should have at a particular time. In some respects, this is an un-physical change in q.

As will be seen later, it is very useful to look at these un-physical changes in quantities to set restriction on how they can change physically, thus determining their equations of motion. Another example is the potential energy of a marble at the bottom of the bowl; it may be interesting for us to consider how the potential energy would change if the marble were to be (un-physically) displaced horizontally away from the bottom of the bowl. Those considerations will lead us to understanding how the marble can behave physically.

2.2 Variation of a function

Consider a function, $y = f(x)$, and a new function, $\bar{f}(x)$:

$$\bar{f}(x) = f(x) + \epsilon\phi(x) \tag{2.1}$$

where ϵ is a number that we can make arbitrarily small, and $\phi(x)$ is a function that is continuous and differentiable in some interval $x = a \ldots b$ over which $f(x)$ is defined, continuous and differentiable. We define the variation of a function as:

$$\delta y \equiv \bar{f}(x) - f(x) = \epsilon\phi(x) \tag{2.2}$$

The change in $f(x)$ is infinitesimal and "virtual", meaning that we can choose any arbitrary well behaved $\phi(x)$. Thus δy does not represent a "real" change in the function from a change in its dependent variables. Again, note the difference between δy and dy; dy is the (familiar) change in $f(x)$ from a corresponding change in x by dx, while δy is a change in $f(x)$ **without** a change in x, and is a **new** function. This is illustrated in Figure 2.1. For the calculus of variations, we only consider changes in the dependent variables, thus:

$$\delta x = 0 \tag{2.3}$$

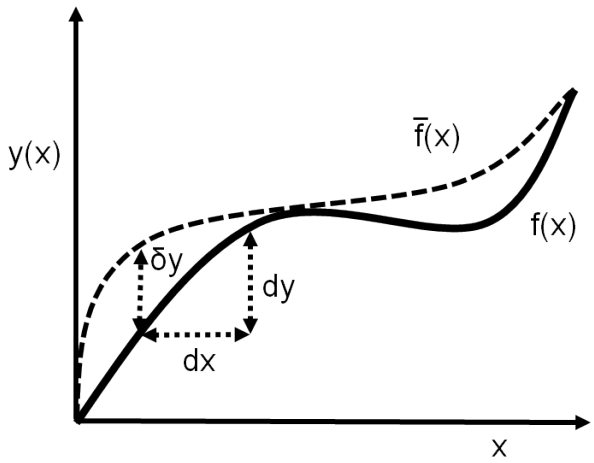

Figure 2.1: A function $y = f(x)$ showing a differential change dx which corresponds to a change in x of dx. The variation of the function at a point, δy, is also shown, which is independent of a change in x and leads to a different function, $\bar{f}(x)$.

2.2.1 Properties of the δ operator

Commutation with differentiation

Consider the "derivative of the variation":

$$\frac{d}{dx}\delta y = \frac{d}{dx}\left[\bar{f}(x) - f(x)\right] = \frac{d}{dx}\epsilon\phi(x) = \epsilon\phi'(x) \tag{2.4}$$

where we have used an apostrophe (') to designate derivatives with respect to x. Now consider the "variation of the derivative":

$$\delta\left(\frac{d}{dx}f(x)\right) \equiv \frac{d\bar{f}}{dx} - \frac{df}{dx} = \frac{d}{dx}(f(x) + \epsilon\phi(x)) - \frac{df}{dx} = \epsilon\phi'(x) \tag{2.5}$$

We thus find that δ is commutative with $\frac{d}{dx}$:

$$\therefore \delta\left(\frac{dy}{dx}\right) = \frac{d}{dx}(\delta y) \tag{2.6}$$

Commutation with integration

Next, consider the variation of a definite integral:

$$\begin{aligned} \delta\int_a^b f(x)dx &\equiv \int_a^b \bar{f}(x)dx - \int_a^b f(x)dx = \int_a^b (\bar{f}(x) - f(x))dx \\ \therefore \delta\int_a^b f(x)dx &= \int_a^b \delta f(x)dx \end{aligned} \tag{2.7}$$

The delta operator also commutes with respect to integration.

Chain Rule

If y is a function of x, the variation of a function, $F(y)$, is given by:

$$\delta F(y) = F(y + \delta y) - F(y) \tag{2.8}$$

We can expand $F(y+\delta y)$ using a Taylor series, where δy is small:

$$F(y+\delta y) = F(y) + \frac{dF}{dy}\delta y + \frac{1}{2!}\frac{d^2F}{dy^2}(\delta y)^2 + \dots \tag{2.9}$$

Given that δy are small, we can ignore the terms in powers higher than 2:

$$\begin{aligned} F(y+\delta y) &= F(y) + \frac{dF}{dy}\delta y \\ \therefore \delta F(y) &= \frac{dF}{dy}\delta y \end{aligned} \tag{2.10}$$

which is (like) the Chain Rule.

EXAMPLE 2-1: Determine $\delta(\cos\theta)$

If the function F is a function of multiple functions, $y(x), z(x), \dots$, we proceed the same way, but using partial derivatives:

$$\begin{aligned} \delta F(y,z,\dots) &= F(y+\delta y, z+\delta z, \dots) - F(y,z,\dots) \\ &= \frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial z}\delta z + \dots + \frac{1}{2!}\left(\frac{\partial^2 F}{\partial y}\delta^2 y + \frac{\partial^2 F}{\partial z}\delta^2 z + 2\frac{\partial^2 F}{\partial y \partial z}\delta y \delta z + \dots\right) + \dots \end{aligned} \tag{2.11}$$

where again, we will neglect the terms in δ^2 and higher.

EXAMPLE 2-2: Compare the differential displacement and virtual displacement of a position vector $\vec{r} = \vec{r}(q_1, q_2, \dots, q_n, t)$ that is a function of generalized coordinates, q.

2.3 Stationary value of a function

In analytic mechanics, we will be interested in the stationary value of an integral. We start by considering the stationary value of a function. For example, an n-dimensional function $F(q_1, q_2, \dots, q_n)$, can be pictured as a surface in a space with n+1 dimension (for example, imagine a 2-dimensional surface in 3-dimensional space). The stationary points of the functions are locations where the surface is "flat", and can correspond to local "extrema" (minima or maxima) or "saddle points". We can use the formalism of variations to find the conditions for such points.

If the point, P, is a stationary value of the function, $F(q_1, q_2, \ldots, q_n)$, then, the variation of the function near P in the direction of an arbitrary virtual displacement, $\delta\vec{q}$, is 0 (the function is flat in the infinitesimal region near the point). The variation of the function is (technically called the "first variation", as we drop the higher order terms in the Taylor series):

$$\delta F = \frac{\partial F}{\partial q_1}\delta q_1 + \frac{\partial F}{\partial q_2}\delta q_2 + \cdots + \frac{\partial F}{\partial q_N}\delta q_N \tag{2.12}$$

We can write this in terms of finite numbers if we write the virtual displacements as:

$$\delta q_i = \epsilon \alpha_i \tag{2.13}$$

where $\vec{\alpha}$ is a vector in the direction of the virtual displacement and ϵ is a number that tends to zero. The rate of change of the function, F, in the direction of $\vec{\alpha}$ is:

$$\frac{\delta F}{\epsilon} = \frac{\partial F}{\partial q_1}\alpha_1 + \frac{\partial F}{\partial q_2}\alpha_2 + \cdots + \frac{\partial F}{\partial q_N}\alpha_N \tag{2.14}$$

which must vanish for a stationary point (writing the above equation as a sum):

$$\sum_i \frac{\partial F}{\partial q_i}\alpha_i = 0 \tag{2.15}$$

In order to have a stationary point, the rate of the change of the function must vanish for any direction, $\vec{\alpha}$, so that each term must be equal to zero, independent of the α_i:

$$\begin{aligned} &\frac{\partial F}{\partial q_i}\alpha_i = 0 \\ &\therefore \frac{\partial F}{\partial q_i} = 0 \end{aligned} \tag{2.16}$$

and we recover the familiar result from calculus that the partial derivatives must vanish at P, for that point to be a stationary point of the function. The second order derivatives ("second variations") are needed in order to know if this is a local minimum, maximum or saddle point. For analytic mechanics, we will generally only need to know if the point is stationary.

2.4 Constraints and Lagrange multipliers

2.4.1 Single constraint

In some cases, the problem of finding a stationary point does not generalize to any virtual displacement, as constraints between coordinates limit the choice of direction for $\vec{\alpha}$. For example, in the case finding the minimum of the potential energy of a marble in a bowl, we would have a constraint equation in the form:

$$x^2 + y^2 + z^2 = (R - r)^2 \tag{2.17}$$

where (x, y, z) are the position of the center of mass of the ball, R is the radius of the bowl, and r is the radius of the ball. The ball is constrained to be in the bowl, so we cannot just consider any point in space as the minimum of the potential energy.

In general then, we seek to minimize the function $F(q_1, q_2, \ldots, q_n)$, subject to some constraint equation(s) between the coordinates:

$$f(q_1, q_2, \ldots, q_n) = 0 \tag{2.18}$$

The obvious (and valid) way to handle this is to use the constraint equation to eliminate one of the variables from F and work with n-1 independent (generalized) coordinates. However, it may not be convenient

to eliminate a coordinate, and there might not be an obvious choice of which coordinate should be the "dependent" one. Lagrange's method of multipliers allows one to preserve all the coordinates while including the constraints (or "auxiliary conditions"). We start by finding the stationary point subject to a single auxiliary condition. We have:

$$\delta f = \frac{\partial f}{\partial q_1}\delta q_1 + \frac{\partial f}{\partial q_2}\delta q_2 + \cdots + \frac{\partial f}{\partial q_n}\delta q_n = 0 \tag{2.19}$$

In order to have a stationary point of F, we also have:

$$\delta F = \frac{\partial F}{\partial q_1}\delta q_1 + \frac{\partial F}{\partial q_2}\delta q_2 + \cdots + \frac{\partial F}{\partial q_n}\delta q_n = 0 \tag{2.20}$$

however, unlike in the un-constrained case, the δq_i, are not independent, thus this no longer implies that each term (and hence each partial derivative) is zero. Since the two above equations are zero, we can create a linear combination of them, which will still equal zero:

$$\begin{aligned}\delta F + \lambda\delta f &= \frac{\partial F}{\partial q_1}\delta q_1 + \frac{\partial F}{\partial q_2}\delta q_2 + \cdots + \frac{\partial F}{\partial q_n}\delta q_n + \lambda(\frac{\partial f}{\partial q_1}\delta q_1 + \frac{\partial f}{\partial q_2}\delta q_2 + \cdots + \frac{\partial f}{\partial q_n}\delta q_n) = 0 \\ \sum_{i=1}^{n}\left(\frac{\partial F}{\partial q_i} + \lambda\frac{\partial f}{\partial q_i}\right)\delta q_i &= 0\end{aligned} \tag{2.21}$$

where λ is called a "Lagrange multiplier", and is not known a priori. Note again that only the sum is zero, not the individual terms. We now **choose** λ so that the n-th term is zero:

$$\frac{\partial F}{\partial q_n} + \lambda\frac{\partial f}{\partial q_n} = 0 \tag{2.22}$$

Given the chosen value of λ, the sum now contains one less term:

$$\sum_{i=1}^{n-1}\left(\frac{\partial F}{\partial q_i} + \lambda\frac{\partial f}{\partial q_i}\right)\delta q_i = 0 \tag{2.23}$$

At this point however, we effectively imposed the constraint to eliminate δq_n; the remaining coordinates can now be varied independently ("freely"), so that the remaining terms (given λ) are individually zero:

$$\frac{\partial F}{\partial q_i} + \lambda\frac{\partial f}{\partial q_i} = 0 \text{ (i=1, 2, \ldots n-1)} \tag{2.24}$$

We thus have $n-1$ equations as above, and 1 equation to determine λ, resulting in n equations to find the point where F is stationary given the auxiliary condition $f = 0$. By inspection, those equations are the same and the distinction between dependent and independent variables vanishes!

This method can thus be written a little bit more generally by considering all the δq_i as independent instead of (artificially) eliminating δq_n, and considering the variation:

$$\delta F + \lambda\delta f \tag{2.25}$$

We have:

$$\begin{aligned}\delta(F + \lambda f) &= \delta F + (\delta\lambda)f + \lambda(\delta f) \\ &= \delta F + \lambda\delta f\end{aligned} \tag{2.26}$$

since $f = 0$. We can then construct a new function, $\bar{F}$:

$$\bar{F} = F + \lambda f \tag{2.27}$$

and consider the stationary points of $\bar{F}$ without having to worry about auxiliary conditions, thus treating all variations of the coordinates as unconstrained ("free variations"). In this situation, we have n equations by setting the partial derivatives of $\bar{F}$ to zero and 1 equation from the constraint equation $f = 0$. This allows us to solve for the n values of q at the minimum and the unknown λ.

EXAMPLE 2-3: Calculate the dimensions, x, y, z of a box that maximizes the volume for a fixed surface area of 2A.

2.4.2 Multiple constraints

The treatment is similar when we have multiple auxiliary conditions (say k such constraints):

$$\begin{aligned} f_1(q_1, q_2, \ldots, q_n) &= 0 \\ f_2(q_1, q_2, \ldots, q_n) &= 0 \\ &\cdots \\ f_k(q_1, q_2, \ldots, q_n) &= 0 \end{aligned} \tag{2.28}$$

where we now have $n - k$ degrees of freedom. The variation of the auxiliary conditions are:

$$\delta f_i = \frac{\partial f_i}{\partial q_1}\delta q_1 + \frac{\partial f_i}{\partial q_2}\delta q_2 + \cdots + \frac{\partial f_i}{\partial q_n}\delta q_n = 0 \tag{2.29}$$

Again, we wish to find the stationary point of F, leading to:

$$\delta F = \sum_{i=1}^{n} \frac{\partial F}{\partial q_i}\delta q_i = 0 \tag{2.30}$$

We now introduce a Lagrange multiplier λ_i for each of the k auxiliary conditions f_i, so that the following linear combination is still zero:

$$\begin{gathered} \delta F + \lambda_1 \delta f_1 + \lambda_2 \delta f_2 + \cdots + \lambda_k \delta f_k = 0 \\ \sum_{i=1}^{n} \left(\frac{\partial F}{\partial q_i} + \lambda_1 \frac{\partial f_1}{\partial q_i} + \cdots + \lambda_k \frac{\partial f_k}{\partial q_i} \right) \delta q_i = 0 \end{gathered} \tag{2.31}$$

We now choose to solve for the k Lagrange multipliers by eliminating the last k terms in the sum, leading to k equations for the k unknown λ:

$$\begin{aligned} &\frac{\partial F}{\partial q_n} + \lambda_1 \frac{\partial f_1}{\partial q_n} + \cdots + \lambda_k \frac{\partial f_k}{\partial q_n} = 0 \\ &\frac{\partial F}{\partial q_{n-1}} + \lambda_1 \frac{\partial f_1}{\partial q_{n-1}} + \cdots + \lambda_k \frac{\partial f_k}{\partial q_{n-1}} = 0 \\ &\frac{\partial F}{\partial q_{n-2}} + \lambda_1 \frac{\partial f_1}{\partial q_{n-2}} + \cdots + \lambda_k \frac{\partial f_k}{\partial q_{n-2}} = 0 \\ &\cdots \\ &\frac{\partial F}{\partial q_{n-k+1}} + \lambda_1 \frac{\partial f_1}{\partial q_{n-k+1}} + \cdots + \lambda_k \frac{\partial f_k}{\partial q_{n-k+1}} = 0 \end{aligned} \tag{2.32}$$

The rest of the $n - k$ coordinates in equation 2.31 can now be varied freely, so that each corresponding terms in the sum must vanish:

$$\sum_{i=1}^{n-k} \left(\frac{\partial F}{\partial q_i} + \lambda_1 \frac{\partial f_1}{\partial q_i} + \cdots + \lambda_k \frac{\partial f_k}{\partial q_i} \right) \delta q_i = 0 \tag{2.33}$$

thus giving us $n - k$ equations:

$$\begin{aligned} &\frac{\partial F}{\partial q_1} + \lambda_1 \frac{\partial f_1}{\partial q_1} + \cdots + \lambda_k \frac{\partial f_k}{\partial q_1} = 0 \\ &\frac{\partial F}{\partial q_2} + \lambda_1 \frac{\partial f_1}{\partial q_2} + \cdots + \lambda_k \frac{\partial f_k}{\partial q_2} = 0 \\ &\cdots \\ &\frac{\partial F}{\partial q_{n-k}} + \lambda_1 \frac{\partial f_1}{\partial q_{n-k}} + \cdots + \lambda_k \frac{\partial f_k}{\partial q_{n-k}} = 0 \end{aligned} \tag{2.34}$$

Again, by inspection, these equations are the same as those in equation 2.32, removing any distinction between dependent and independent variables. We can thus construct the function $\bar{F}$:

$$\bar{F} = F + \lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_k f_k \tag{2.35}$$

and treat all of the coordinates in the same way. That is, instead of performing the variation of F subject to the auxiliary conditions, we can perform the variation on $\bar{F}$ and "ignore" the auxiliary conditions. This gives us the n equations:

$$\frac{\partial F}{\partial q_i} + \lambda_1 \frac{\partial f_i}{\partial q_i} + \cdots + \lambda_k \frac{\partial f_k}{\partial q_i} = 0 \text{ (i=0...n)} \tag{2.36}$$

together with the k constraint equations to determine the n coordinates q_i and the k Lagrange multipliers.

2.5 Stationary value of an integral

In analytic mechanics, we will find that we need to find the stationary value of an integral. For example, we may need to find the curve $y(x)$ that results in the definite integral of some function $L(y, y', x)$ being maximized (or stationary). We called the definite integral, I, a "functional":

$$I = \int_a^b L(y, y', x) dx \tag{2.37}$$

One common example is the "Brachistochrone" problem, in which we wish to find the shape of a wire, $y(x)$, that minimizes the time for a frictionless bead to slide down under gravity from some point a to another point b. This problem in fact led to the development of the field of variational calculus.

EXAMPLE 2-4: Derive the functional corresponding to a bead sliding under gravity down a frictionless wire given by the function $y(x)$

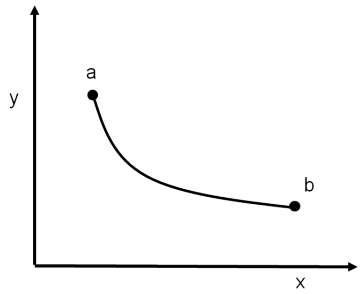

Figure 2.2: Bead sliding under gravity on a wire from a to b.

We thus want to find the condition on L such that the functional, I, is stationary, given the boundary condition that the function, $y(x)$, passes through a and b.

In analogy with finding the stationary point of a function, we examine the rate of change of the functional as we vary $y(x)$. We start with the variation of the integrand:

$$\delta L(y, y', x) = L(y + \delta y, y' + \delta y', x) - L(y, y', x) \tag{2.38}$$

where, as you recall, we do not consider any variations in x. We can use the first order terms in δ from a Taylor series to expand $L(y + \delta y, y' + \delta y', x)$:

$$L(y + \delta y, y' + \delta y', x) = L(y, y', x) + \frac{\partial L}{\partial y}\delta y + \frac{\partial L}{\partial y'}\delta y' \tag{2.39}$$

Thus:

$$\begin{aligned} \delta I &= \delta \int_a^b L(y, y', x)dx \\ &= \int_a^b \delta L(y, y', x)dx \\ &= \int_a^b \left(\frac{\partial L}{\partial y}\delta y + \frac{\partial L}{\partial y'}\delta y' \right) dx \end{aligned} \tag{2.40}$$

Recalling from equations 2.2 and 2.6 that we can write:

$$\begin{aligned} \delta y &= \epsilon\phi(x) \\ \delta y' &= \epsilon\phi'(x) \end{aligned} \tag{2.41}$$

where $\phi(x)$ is a small variation of the function $y(x)$. The functional becomes:

$$\delta I = \epsilon \int_a^b \left(\frac{\partial L}{\partial y}\phi(x) + \frac{\partial L}{\partial y'}\phi'(x) \right) dx \tag{2.42}$$

Now, since we want to find a stationary value of I, we require the rate of change of δI with respect to ϵ to be zero:

$$\frac{d\delta I}{d\epsilon} = \frac{\delta I}{\epsilon} = \int_a^b \left(\frac{\partial L}{\partial y}\phi(x) + \frac{\partial L}{\partial y'}\phi'(x) \right) dx = 0 \tag{2.43}$$

The second part can be integrated by parts:

$$\int_a^b \frac{\partial L}{\partial y'}\phi'(x)dx = \phi(x)\frac{\partial L}{\partial y'}\Big|_a^b - \int_a^b \frac{d}{dx}\left(\frac{\partial L}{\partial y'} \right)\phi(x)dx \tag{2.44}$$

The first term of integration by parts is zero, because the function $y(x)$ is fixed at the end points of the integration range, so $\phi(x=a)=\phi(x=b)=0$. We thus have:

$$\begin{aligned}\frac{\delta I}{\epsilon} &= \int_a^b \left(\frac{\partial L}{\partial y}\phi(x) - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)\phi(x)\right)dx \\ &= \int_a^b \phi(x)\left(\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)\right)dx = 0\end{aligned} \tag{2.45}$$

The only way for this integral to vanish, regardless of the choice of $\phi(x)$ is for the remaining part of the integrand to always be zero. This leads to the "Euler-Lagrange" equation, which is the condition for the functional $\int L(y,y',x)dx$ to be stationary:

$$\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) - \frac{\partial L}{\partial y} = 0 \tag{2.46}$$

Noting that $\phi(x)=\delta y$, we could also have written equation 2.45 as:

$$\frac{\delta I}{\epsilon} = \int_a^b \delta y\left(\frac{\partial L}{\partial y} - \frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right)\right)dx = 0 \tag{2.47}$$

a notation which we will find convenient in the next section.

In analytic mechanics, we will generally have a function, L, that depends on multiple coordinates, q_i, their generalized velocities, $\dot{q}$, and time will play the role of the independent variable:

$$L = L(q_1, q_2, \ldots, \dot{q_1}, \dot{q_2}, \ldots, t) \tag{2.48}$$

We will want to find the values of $q_i(t)$ that lead to a stationary value of:

$$S = \int_a^b L(q_1, q_2, \ldots, \dot{q_1}, \dot{q_2}, \ldots, t)dt \tag{2.49}$$

Following the same procedure as above, and collecting the terms for each coordinate, we can show that we get an Euler-Lagrange equation for each coordinate:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0 \text{ (i=1...n)} \tag{2.50}$$

2.6 Stationary value of an integral that does not explicitly depend on x

We consider the special case when the function L does not depend explicitly on x (that is, it still depends explicitly on the functions $y(x)$ and $y'(x)$):

$$\begin{aligned}L &= L(y,y') \\ \frac{\partial L}{\partial x} &= 0\end{aligned} \tag{2.51}$$

Consider the following identity:

$$\begin{aligned}\frac{d}{dx}\left(y'\frac{\partial L}{\partial y'} - L\right) &= y''\frac{\partial L}{\partial y'} + y'\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) - \frac{dL}{dx} \\ &= y''\frac{\partial L}{\partial y'} + y'\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) - \left(\frac{\partial L}{\partial y}\frac{dy}{dx} + \frac{\partial L}{\partial y'}\frac{dy'}{dx} + \frac{\partial L}{\partial x}\right) \\ &= y''\frac{\partial L}{\partial y'} + y'\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) - \left(\frac{\partial L}{\partial y}y' + \frac{\partial L}{\partial y'}y''\right) \\ &= y'\left[\frac{d}{dx}\left(\frac{\partial L}{\partial y'}\right) - \frac{\partial L}{\partial y}\right]\end{aligned} \tag{2.52}$$

where we have used the fact that:

$$dL = \frac{\partial L}{\partial y}dy + \frac{\partial L}{\partial y'}dy' + \frac{\partial L}{\partial x}dx \tag{2.53}$$

where the last term is zero.

If the function L satisfies the Euler-Lagrange equations, this implies that:

$$\begin{aligned} \frac{d}{dx}\left(y'\frac{\partial L}{\partial y'} - L\right) &= 0 \\ \therefore y'\frac{\partial L}{\partial y'} - L &= k \end{aligned} \tag{2.54}$$

where k is a constant. This is equivalent to the Euler-Lagrange equations for the case where L does not explicitly depend on x. It is an easier equation to solve for $y(x)$, since it only contains first order derivatives.

2.6.1 The case when there are multiple functions

Suppose that the function L depends on multiple functions $y_1(x), \ldots, y_n(x)$ and their derivatives, $y_1'(x), \ldots, y_n'(x)$, $L = L(y_1(x), \ldots, y_n(x), y_1'(x), \ldots, y_n'(x))$. If L does not explicitly depend on x, we have:

$$\begin{aligned} \frac{dL}{dx} &= \frac{\partial L}{\partial y_1}y_1' + \cdots + \frac{\partial L}{\partial y_n}y_n' + \frac{\partial L}{\partial y_1'}y_1'' + \cdots + \frac{\partial L}{\partial y_n'}y_n'' \\ &= \sum_{i=1}^{n}\left(\frac{\partial L}{\partial y_i}y_i' + \frac{\partial L}{\partial y_i'}y_i''\right) \end{aligned} \tag{2.55}$$

We thus consider the identity:

$$\begin{aligned} \frac{d}{dx}\left(\sum_{i=1}^{n}\left(y_i'\frac{\partial L}{\partial y_i'}\right) - L\right) &= \sum_{i=1}^{n}\left(y_i''\frac{\partial L}{\partial y_i'} + y_i'\frac{d}{dx}\frac{\partial L}{\partial y_i'}\right) - \frac{dL}{dx} \\ &= \sum_{i=1}^{n}\left(y_i''\frac{\partial L}{\partial y_i'} + y_i'\frac{d}{dx}\frac{\partial L}{\partial y_i'}\right) - \sum_{i=1}^{n}\left(\frac{\partial L}{\partial y_i}y_i' + \frac{\partial L}{\partial y_n'}y_n''\right) \\ &= \sum_{i=1}^{n}\left(y_i'\left[\frac{d}{dx}\frac{\partial L}{\partial y_i'} - \frac{\partial L}{\partial y_i}\right]\right) \end{aligned} \tag{2.56}$$

Again, the term on the right is zero since the Euler-Lagrange equations are satisfied by each $y_i(x)$. We thus obtain the equivalent result as we did with a single function (noting that we have a sum on the left):

$$\begin{aligned} \frac{\partial L}{\partial x} &= 0 \\ \therefore \sum_{i=1}^{n}\left(y_i'\frac{\partial L}{\partial y_i'}\right) - L &= k \end{aligned} \tag{2.57}$$

where k is a constant.

2.7 Stationary value of an integral with auxiliary conditions

We conclude this chapter by consider the stationary value of a functional with integrand, $L(q_1, q_2, \ldots, \dot{q_1}, \dot{q_2}, \ldots, t)$, subject to k constraints on the, $q_i(t)$, functions. Here, in analogy with mechanics, we chose $q(t)$ and $\dot{q}(t)$ as functions that depend on the independent variable, t (instead of $y(x)$ as we had in the previous section).

$$\begin{aligned} f_1(q_1, q_2, \ldots, q_n, t) &= 0 \\ f_2(q_1, q_2, \ldots, q_n, t) &= 0 \\ &\cdots \\ f_k(q_1, q_2, \ldots, q_n, t) &= 0 \end{aligned} \tag{2.58}$$

where the constraints depend on time. As before, it would be possible to eliminate k of the $q_i(t)$ functions, but this may not be the most mathematically convenient and may arbitrarily make some functions "dependent" and others "independent". Instead, we can use the method of the Lagrangian multipliers. The variation of the constraints are zero for all times:

$$\begin{aligned}
\delta f_1 &= \frac{\partial f_1}{\partial q_1}\delta q_1 + \frac{\partial f_1}{\partial q_2}\delta q_2 + \cdots + \frac{\partial f_1}{\partial q_n}\delta q_n = 0 \\
\delta f_2 &= \frac{\partial f_2}{\partial q_1}\delta q_1 + \frac{\partial f_2}{\partial q_2}\delta q_2 + \cdots + \frac{\partial f_2}{\partial q_n}\delta q_n = 0 \\
&\cdots \\
\delta f_k &= \frac{\partial f_k}{\partial q_1}\delta q_1 + \frac{\partial f_k}{\partial q_2}\delta q_2 + \cdots + \frac{\partial f_k}{\partial q_n}\delta q_n = 0
\end{aligned} \tag{2.59}$$

We can then construct a new functional of which we want to find a stationary value:

$$\delta I = \int_{t_a}^{t_b} \delta L(q_1, q_2, \ldots, \dot{q_1}, \dot{q_2}, \ldots, t)dt + \int_a^b (\lambda_1 \delta f_1 + \lambda_2 \delta f_2 + \cdots + \lambda_k \delta f_k)dt = 0 \tag{2.60}$$

where we have effectively added zero to our original functional. Note that the λ_i are in principle functions of t, since the f_i are functions of t. The first term of the variation, after integration by parts will have a term containing δq_i of the form:

$$\int_{t_a}^{t_b} \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} \right) \delta q_i dt \tag{2.61}$$

and there will be a matching term in the part with the Lagrange multipliers:

$$\int_{t_a}^{t_b} \left(\lambda_1 \frac{\partial f_1}{\partial q_i} + \lambda_2 \frac{\partial f_2}{\partial q_i} + \cdots + \lambda_k \frac{\partial f_k}{\partial q_i} \right) \delta q_i dt \tag{2.62}$$

This results in the functional being

$$\delta I = \sum_{i=1}^{n} \int_a^b \left[\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} \right) + \left(\lambda_1 \frac{\partial f_1}{\partial q_i} + \lambda_2 \frac{\partial f_2}{\partial q_i} + \cdots + \lambda_k \frac{\partial f_k}{\partial q_i} \right) \right] \delta q_i dt \tag{2.63}$$

In principle, the term in square brackets is **not** zero, because the δq_i cannot all be varied independently because of the constraint equations 2.59. We can however **choose** the λ_i such that the last k terms are zero:

$$\left[\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} \right) + \left(\lambda_1 \frac{\partial f_1}{\partial q_i} + \lambda_2 \frac{\partial f_2}{\partial q_i} + \cdots + \lambda_k \frac{\partial f_k}{\partial q_i} \right) \right] = 0 \text{ (i=n-k+1...n)} \tag{2.64}$$

Now, this leaves us with the remaining $n - k$ terms for which the δq_i, can be varied independently. Of course, the whole point of the Lagrange multiplier method is that this gives equations that are identical for all q_i, removing the distinction between dependent and independent variables. The above equation is thus true for all values of $i = 1 \ldots n$!

We can thus re-express the problem of varying the integrand L subject to auxiliary conditions as the equivalent problem of varying a modified integrand, L', without worrying about auxiliary conditions:

$$L' = L - (\lambda_1 f_1 + \lambda_2 f_2 + \cdots + \lambda_k f_k) \tag{2.65}$$

where the auxiliary conditions can be used in addition to the Euler-Lagrange equations to determine all of

the $q_i(t)$ and the $\lambda_i(t)$. This leads to the Euler-Lagrange equations which we can re-write as:

$$\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i}\right) = -\lambda_1 \frac{\partial f_1}{\partial q_i} - \lambda_2 \frac{\partial f_2}{\partial q_i} - \cdots - \lambda_k \frac{\partial f_k}{\partial q_i}$$
$$\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i}\right) = -\sum_{j=1}^{k} \lambda_j \frac{\partial f_j}{\partial q_i}$$
$$\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i}\right) = Q_i \tag{2.66}$$

where we have introduced:

$$Q_i \equiv -\sum_{j=1}^{k} \lambda_j \frac{\partial f_j}{\partial q_i} \tag{2.67}$$

and the overall negative sign is not important, since the Lagrange multipliers can be chosen.

EXAMPLE 2-5: Show that the Euler-Lagrange equation for $L' = L - \lambda_1 f_1 - \lambda_2 f_2 - \cdots - \lambda_k f_k$ is equivalent to equation 2.64. That is, show that the sign in front of the Lagrange multipliers is arbitrary.

2.64

2.7.1 Case when auxiliary condition is in the form of an integral

In certain cases, the auxiliary condition may be in the form of an integral:

$$\int_{t_a}^{t_b} f(q_1, q_2, \ldots, t)dt = C \tag{2.68}$$

where C is a constant, and there may be any number of such constraint equations. Again, we can take the variation of this integral:

$$\delta \int_{t_a}^{t_b} f(q_1, q_2, \ldots, t)dt = \int_{t_a}^{t_b} \sum \frac{\partial f}{\partial q_i} \delta q_i dt = 0 \tag{2.69}$$

which must be zero. We can thus use a Lagrange multiplier and add the variation of the constraint to our original integral:

$$\delta \int_{t_a}^{t_b} L(q_1, q_2, \ldots, \dot{q}_1, \dot{q}_2, \ldots, t)dt + \lambda\delta \int_{t_a}^{t_b} f(q_1, q_2, \ldots, t)dt$$
$$= \delta \int_{t_a}^{t_b} \left[L(q_1, q_2, \ldots, \dot{q}_1, \dot{q}_2, \ldots, t) + \lambda f(q_1, q_2, \ldots, t)\right] dt \tag{2.70}$$

We can solve this problem by considering the stationary value of a new function, $\bar{L}$:

$$\bar{L}(q_1, q_2, \ldots, \dot{q_1}, \dot{q_2}, \ldots, t) \equiv L(q_1, q_2, \ldots, \dot{q_1}, \dot{q_2}, \ldots, t) + \lambda f(q_1, q_2, \ldots, t) \tag{2.71}$$

and apply the Euler-Lagrange equation to $\bar{L}$.

2.8 Problems

Problem 2-1: Highest point on a surface
Find the coordinates of the highest point on the surface given by $z(x, y) = x^2y + xy$ that when projected onto the xy-plane lies on a circle of radius $r = 5$ centered at $x = 2$ and $y = 3$. Use a computer to illustrate the result. You will also likely need to solve the equations numerically.

Problem 2-2: Box with no lid
Find the dimensions that maximize the volume of a box with no lid, if the total surface area of the box is $1\,\mathrm{m}^2$.

Problem 2-3: Euler-Lagrange equations for integrand that depends on more than one function
Show that the stationary value of the functional $S = \int_a^b L(q_1(t), q_2(t), \ldots, \dot{q_1}, \dot{q_2}, \ldots, t)dt$ gives an Euler-Lagrange equation for each $q_i(t)$:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q_i}}\right) - \frac{\partial L}{\partial q_i} = 0$$

Problem 2-4: Euler-Lagrange equation when there is a second order derivative
Show that the equivalent of the Euler-Lagrange equation when the function in the integrand depends on the second derivative of y:

$$I = \int_a^b L(y, y', y'', x)dx$$

is given by:

$$\frac{d^2}{dx^2}\frac{\partial L}{\partial y''} - \frac{d}{dx}\frac{\partial L}{\partial y'} + \frac{\partial L}{\partial y} = 0$$

Note that you will have to integrate by parts twice and that the variation of $y'(x)$ is zero at the end points of the integral.

Problem 2-5: Brachistrochrone
Refer to the problem of example 2-4.
a) Find the differential equation for the function $y(x)$ that solves the brachistrochrone problem, by minimizing T:

$$\sqrt{2g}T = \int_0^{x_b} \frac{\sqrt{1 + y'^2}}{\sqrt{y}}dx$$

Note that we have set the problem up such that one end of the wire is at the origin, gravity is in the positive y direction, and the constant $2g$ will not change the shape of the wire. Note that a closed form solution in the form $y = f(x)$ is not possible, and only parametric solutions can be obtained (so called "cycloids").

b) Repeat part a), with the wire being constrained to have a length of L.

Problem 2-6: Geodesic
Determine the functional for the shortest distance between two points on a sphere of radius a, and write

the differential equation that gives the function $\phi(\theta)$, where ϕ and θ are the azimuthal and polar angles in spherical coordinates, respectively.

Problem 2-7: Straight line
Show that the shortest distance between two points in a plane is a straight line.

Problem 2-8: Catenary
A rope of uniform linear mass density, μ, and total length, L, is hung across a precipice of length, H (both sides of the precipice are at the same height). Minimize the potential energy of the rope:

$$V = \int_0^H \mu g y ds$$

to obtain the differential equation for the shape of the rope. Use a computer to solve and plot the resulting curve (choose reasonable values).

Problem 2-9: Surface and volume of revolution
a) Find the shape of the curve, $y(x)$, passing through two points, a and b, that gives the minimal surface of revolution when rotated about the x-axis (see figure). Note that this would be the shape of a soap film formed between two hoops.
b) Find the shape of the curve, $y(x)$, passing through two points, a and b, that creates a solid of revolution with the minimal moment of inertia about the x-axis (see figure). Assume that the volume is made of a material with uniform density. It will not be possible to get a closed form, write the solution as a differential equation for $y(x)$ or give x in terms of an integral over y.

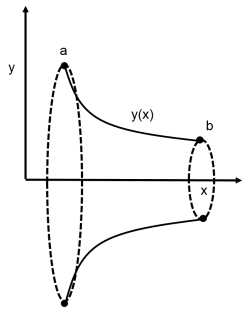

Figure 2.3: Minimal surface of revolution between two points, problem 2-9

2.9 Solutions

Problem 2-1:

3 Virtual work and D'Alembert's principle

One of the advantages of analytic mechanics is that it removes the need to deal with "internal forces". Consider the problem of two masses attached by a rigid mass-less rod; in vector mechanics, one needs to know the tension in the rod to be able to write the forces on each mass and obtain the equation of motion. In principle, the rod itself is made of an almost infinite number of particles each exerting forces on each other, and a complete vectorial approach is not tractable. In analytic mechanics, we can distinguish "internal forces" and usually ignore them. Typically, these internal forces can be handled easily by using a constraint (e.g. the rod is a rigid object).

3.1 Virtual work

Virtual work, δW_i is the work done by a force, $\vec{F}_i$ given a "reversible virtual displacement" that is in harmony with the given constraints, $\delta\vec{r}_i$:

$$\delta W_i = \vec{F}_i \cdot \delta\vec{r}_i \tag{3.1}$$

A reversible displacement is one where $\delta\vec{r}_i$ can be replaced with $-\delta\vec{r}_i$ without violating any of the constraints $\delta\vec{r}_i$ is the variation of the position vector where the force is applied. By definition, a virtual displacement that is parallel to a normal force is not reversible.

3.1.1 Principal of virtual work and static equilibrium

The principle of virtual work states that for a static equilibrium, the sum of virtual work done by all forces is equal to zero.

$$\sum_{i=1}^{N} \delta W_i = 0$$

$$\sum_{i=1}^{N} \vec{F}_i \cdot \delta\vec{r}_i = 0 \tag{3.2}$$

We can divide this up into "external forces" (such as gravity, electric fields), $\vec{F}_i^E$, and "internal forces" (such as tension in a rod, normal reaction forces), $\vec{F}_j^I$.

$$\sum_{i=1}^{N} \delta W_i = \sum_{i=1}^{N} \vec{F}_i^E \cdot \delta\vec{r}_i + \sum_{i=1}^{N} \vec{F}_j^I \cdot \delta\vec{r}_j \tag{3.3}$$

Typically, the internal forces are related to constraints. "Workless constraints" are those that lead to forces that do no virtual work.

EXAMPLE 3-1: The tension force in a rigid rod that holds two masses together is workless

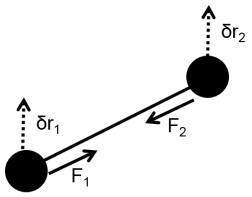

Figure 3.1: Two masses constrained by a mass-less rigid rod.

EXAMPLE 3-2: The normal force on a frictionless surface is workless

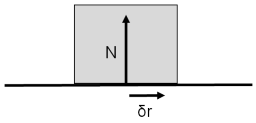

Figure 3.2: Block on a surface

In the case of workless constraints, the total virtual work is given by the work of the external forces. It is generally true (although difficult to prove) that most constraint forces are workless and can be ignored.

For static equilibrium, the "Principal of Virtual Work" states that the total virtual work (which is done by external forces) must be zero:

$$\therefore \sum_{i=1}^{N} \delta W_i = \sum_{i=1}^{N} \vec{F}_i^E \cdot \delta \vec{r}_i = 0 \text{ (workless constraints)} \tag{3.4}$$

This is in contrast to the vectorial approach with requires the sum of all forces (external and internal) to be zero.

3.1.2 Generalized forces

We continue by ignoring the internal forces and consider the virtual work done by N external forces:

$$\sum_{i=1}^{N} \delta W_i = \sum_{i=1}^{N} \vec{F}_i^E \cdot \delta \vec{r}_i = 0 \tag{3.5}$$

If we have n degrees of freedom, we can re-write this in terms of the generalized coordinates (where holonomic constraints are used to reduce the number of coordinates, and the coordinates are thus all independent):

$$\begin{aligned}
\vec{r}_i &= \vec{r}_i(q_1, \ldots, q_n) \\
\delta\vec{r}_i &= \sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j \\
\therefore \sum_{i=1}^{N} \delta W_i &= \sum_{i=1}^{N} \vec{F}_i^E \cdot \left(\sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j \right) \\
\sum_{i=1}^{N} \delta W_i &= \sum_{j=1}^{n} \left(\sum_{i=1}^{N} \vec{F}_i^E \cdot \frac{\partial \vec{r}_i}{\partial q_j} \right) \delta q_j \\
\sum_{i=1}^{N} \delta W_i &= \sum_{j=1}^{n} Q_j \delta q_j
\end{aligned} \tag{3.6}$$

where we have introduced the components, Q_j, of the "generalized force" which do not necessarily have the dimensions of force:

$$Q_j \equiv \sum_{i=1}^{N} \vec{F}_i^E \cdot \frac{\partial \vec{r}_i}{\partial q_j} \tag{3.7}$$

The total work is thus the scalar product of the generalized force, $\vec{Q}$, and a virtual displacement vector, $\delta\vec{q}$, in configuration space:

$$\sum_{i=1}^{N} \delta W_i = \vec{Q} \cdot \delta\vec{q} \tag{3.8}$$

If any displacement is allowed in configuration space, $\delta\vec{q}$, then all components of the generalized force must be zero for static equilibrium. In a holonomic system, where all coordinates are independent of each other, δq_i all correspond to allowed displacements and the generalized force components are therefore zero. Another way to picture this is that, given generalized coordinates q_j, the generalized force Q_j is the force that does work $Q_j \delta q_j$ when the system is displaced in the direction δq_j. Since q_j is not necessarily a cartesian coordinate (it could be an angle), Q_j does not necessarily have the units of force.

3.1.3 Using the Principal of Virtual Work to solve statics problems

We proceed with a few examples for solving statics problems using the principal of virtual work. The general procedure will be as follows:

1. Identify the number of degrees of freedom and choose generalized coordinates

2. Identify the forces that can perform virtual work (forces acting at points where a virtual displacement is possible given the constraints, and forces that are not perpendicular to the allowable virtual displacements)

3. Write the position vectors of the points where the forces are applied in terms of the generalized coordinates. Then, take the variations in those vectors to obtain the $\delta\vec{r}_i$

4. Write out the virtual work and set it to zero

5. (Alternatively) Evaluate the components of the generalized force and set them equal to zero

EXAMPLE 3-3: What is the magnitude, F, of the horizontal force required to maintain a pivoting bar (restricted to pivot in a plane) of mass m, length $2L$, and negligible diameter (see Figure 3.3) at an angle θ?

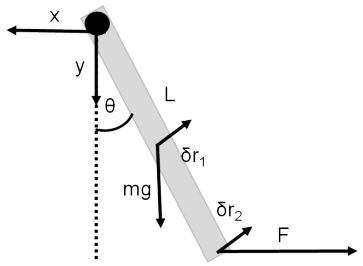

Figure 3.3: A pivoting bar, held in static equilibrium by a force, F.

EXAMPLE 3-4: Find the magnitude of the force, F, required to keep the system in equilibrium from Figure 3.4 at an angle θ. The blocks of mass m slide with no friction and are held by a mass-less rigid rod of length, L. Their dimensions are negligible.

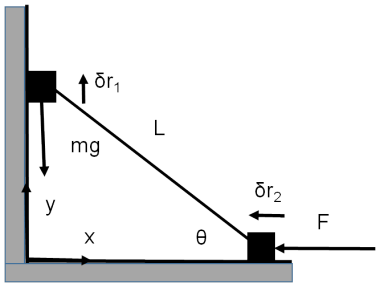

Figure 3.4: Two blocks constrained by a massless rigid rod, held in equilibrium by a force F.

EXAMPLE 3-5: Find the value of θ for the two blocks in Figure 3.5 to be in equilibrium. The blocks of mass m and $2m$ are sitting on a frictionless sphere of radius R and are connected by a mass-less inextensible string of length L.

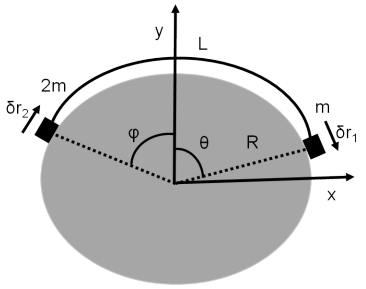

Figure 3.5: Two masses connected by a string sitting on a frictionless sphere.

3.2 D'Alembert's principle

D'Alembert's principle can be used to extend the Principle of Virtual Work to dynamics problems. Starting with Newton's Second Law:

$$\vec{F} = m\vec{a} \tag{3.9}$$

we introduce a new vector, for the "negative force of inertia", I:

$$\begin{aligned} \vec{I} &\equiv -m\vec{a} \\ \therefore \vec{F} + \vec{I} &= 0 \end{aligned} \tag{3.10}$$

With the introduction of the this "force of inertia", we have effectively changed the mathematics of a problem of dynamics to the formalism of statics, where the sum of the forces and torques must be zero. One can also think of finding a frame of reference where the body is instantaneously at rest (imagine the force of inertia on a body inside a car going around a turn). D'Alembert's Principle thus consists of applying the principle of virtual work to a system when the forces of inertia are included. This means that the total virtual work done by all the forces must be zero. Again, considering virtual displacements $\delta\vec{r}$ that are in harmony with the constraints of motion, we can write:

$$\sum_{i=1}^{N} \delta W_i = \sum_{i=1}^{N} (\vec{F_i} + \vec{I_i}) \cdot \delta\vec{r_i} = 0 \tag{3.11}$$

where the sum is over the N particles in the system, and we have introduced the effective forces $\vec{\mathcal{F}}_i$. In principle, the F_i contain both "applied forces" and "forces of constraint" (internal forces). However, most forces of constraint cannot perform virtual work, and it is safe to generally ignore them. We can thus state that F_i are only the applied forces without any substantial loss in generality (although D'Alembert's priniciple does apply in general). Note that for a system of particles, F_i can be identified with the net force applied on particle i.

D'Alembert's principle thus extends the principle of virtual work to the realm of dynamics (and removes the need to worry about internal forces). This leads to the (differential) equations of motion for the particles, since the inertial force will contain the derivatives associated with acceleration.

The effective individual forces in equation 3.11 are not necessarily all equal to zero, as the virtual displacements $\delta\vec{r}_i$ are not necessarily independent. As we did for the principle of virtual work, we can change coordinates to express D'Alembert's principle using the n independent generalized coordinates (if this is a holonomic system).

$$\begin{aligned}\sum_{i=1}^{N}\delta W_i &= \sum_{i=1}^{N}(\vec{F}_i + \vec{I}_i)\cdot\delta\vec{r}_i = 0\\ &= \sum_{i=1}^{N}\vec{F}_i\cdot\delta\vec{r}_i + \sum_{i=1}^{N}\vec{I}_i\cdot\delta\vec{r}_i = 0\\ &= \sum_{j=1}^{n}Q_j\delta q_j + \sum_{i=1}^{N}\vec{I}_i\cdot\delta\vec{r}_i = 0\end{aligned} \tag{3.12}$$

where we have introduced the generalized forces from from equation 3.7:

$$Q_j \equiv \sum_{i=1}^{N}\vec{F}_i\cdot\frac{\partial\vec{r}_i}{\partial q_j} \tag{3.13}$$

The second term in the virtual work must also be transformed to generalized coordinates, it is however a little more tricky because it contains the acceleration vectors $\vec{a} = \ddot{\vec{r}}$:

$$\begin{aligned}\sum_{i=1}^{N}\vec{I}_i\cdot\delta\vec{r}_i &= -\sum_{i=1}^{N}m_i\ddot{\vec{r}}_i\cdot\delta\vec{r}_i\\ &= -\sum_{i=1}^{N}m_i\ddot{\vec{r}}_i\cdot\sum_{j=1}^{n}\frac{\partial\vec{r}_i}{\partial q_j}\delta q_j\end{aligned} \tag{3.14}$$

Consider the following way to re-write this term using the product rule:

$$\begin{aligned}\frac{d}{dt}\left(m_i\dot{\vec{r}}_i\cdot\sum_{j=1}^{n}\frac{\partial\vec{r}_i}{\partial q_j}\delta q_j\right) &= m_i\ddot{\vec{r}}_i\cdot\sum_{j=1}^{n}\frac{\partial\vec{r}_i}{\partial q_j}\delta q_j + m_i\dot{\vec{r}}_i\cdot\frac{d}{dt}\left(\sum_{j=1}^{n}\frac{\partial\vec{r}_i}{\partial q_j}\delta q_j\right)\\ &= m_i\ddot{\vec{r}}_i\cdot\sum_{j=1}^{n}\frac{\partial\vec{r}_i}{\partial q_j}\delta q_j + m_i\dot{\vec{r}}_i\cdot\sum_{j=1}^{n}\frac{\partial}{\partial q_j}\frac{d\vec{r}_i}{dt}\delta q_j\end{aligned} \tag{3.15}$$

where we have use the fact that we can interchange $\frac{d}{dt}$ and $\frac{\partial}{\partial q_j}$. We can now re-arrange and remove the summation sign and the (δq_j) (since the term in front of the sum can be brought into the sum, the equality must be true for each term:

$$\begin{aligned}\therefore m_i\ddot{\vec{r}}_i\cdot\frac{\partial\vec{r}_i}{\partial q_j} &= \frac{d}{dt}\left(m_i\dot{\vec{r}}_i\cdot\frac{\partial\vec{r}_i}{\partial q_j}\right) - m_i\dot{\vec{r}}_i\cdot\frac{\partial}{\partial q_j}\frac{d\vec{r}_i}{dt}\\ &= \frac{d}{dt}\left(m_i\dot{\vec{r}}_i\cdot\frac{\partial\vec{r}_i}{\partial q_j}\right) - m_i\dot{\vec{r}}_i\cdot\frac{\partial\dot{\vec{r}}_i}{\partial q_j}\end{aligned} \tag{3.16}$$

Also note the following relation:

$$\begin{aligned}\frac{\partial \dot{\vec{r}}_i}{\partial \dot{q}_j} &= \frac{\partial}{\partial \dot{q}_j}\frac{d\vec{r}_i}{dt} \\ &= \frac{\partial}{\partial \dot{q}_j}\left(\sum_{k=1}^{n}\frac{\partial \vec{r}_i}{\partial q_k}\dot{q}_k + \frac{\partial \vec{r}_i}{\partial t}\right) \\ &= \frac{\partial \vec{r}_i}{\partial q_k}\delta_{jk} \\ \therefore \frac{\partial \dot{\vec{r}}_i}{\partial \dot{q}_j} &= \frac{\partial \vec{r}_i}{\partial q_j}\end{aligned} \tag{3.17}$$

where we have used the Kronecker Delta (δ_{jk}). We can put this back into equation 3.16:

$$m_i\ddot{\vec{r}}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} = \frac{d}{dt}\left(m_i\dot{\vec{r}}_i \cdot \frac{\partial \dot{\vec{r}}_i}{\partial \dot{q}_j}\right) - m_i\dot{\vec{r}}_i \cdot \frac{\partial \dot{\vec{r}}_i}{\partial q_j} \tag{3.18}$$

Note the following relation that can be used to modify the first term:

$$\begin{aligned}\frac{\partial}{\partial \dot{q}_j}\dot{\vec{r}}_i \cdot \dot{\vec{r}}_i = \dot{\vec{r}}_i \cdot \frac{\partial \dot{\vec{r}}_i}{\partial \dot{q}_j} + \frac{\partial \dot{\vec{r}}_i}{\partial \dot{q}_j} \cdot \dot{\vec{r}}_i = 2\dot{\vec{r}}_i \cdot \frac{\partial \dot{\vec{r}}_i}{\partial \dot{q}_j} \\ \therefore \dot{\vec{r}}_i \cdot \frac{\partial \dot{\vec{r}}_i}{\partial \dot{q}_j} = \frac{1}{2}\frac{\partial}{\partial \dot{q}_j}\dot{\vec{r}}_i \cdot \dot{\vec{r}}_i \\ = \frac{1}{2}\frac{\partial}{\partial \dot{q}_j}\dot{r}_i^2\end{aligned} \tag{3.19}$$

Similarly:

$$\dot{\vec{r}}_i \cdot \frac{\partial \dot{\vec{r}}_i}{\partial q_j} = \frac{1}{2}\frac{\partial}{\partial q_j}\dot{r}_i^2 \tag{3.20}$$

Thus:

$$\begin{aligned}m_i\ddot{\vec{r}}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} &= \frac{d}{dt}\left(m_i\frac{1}{2}\frac{\partial}{\partial \dot{q}_j}\dot{r}_i^2\right) - m_i\frac{1}{2}\frac{\partial}{\partial q_j}\dot{r}_i^2 \\ &= \frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_j}(\frac{1}{2}m_i\dot{r}_i^2)\right) - \frac{\partial}{\partial q_j}(\frac{1}{2}m_i\dot{r}_i^2) \\ &= \frac{d}{dt}\left(\frac{\partial T_i}{\partial \dot{q}_j}\right) - \frac{\partial T_i}{\partial q_j}\end{aligned} \tag{3.21}$$

where we have introduced:

$$T_i \equiv \frac{1}{2}m_i\dot{r}_i^2 \tag{3.22}$$

which we can recognize as the kinetic energy of particle i. We are now ready to put this back into the

equation for the virtual work of the inertial force (remember, we got a little side-tracked at equation 3.14!):

$$\begin{aligned}
\sum_{i=1}^{N} \vec{I}_i \cdot \delta\vec{r}_i &= -\sum_{i=1}^{N} m_i \ddot{\vec{r}}_i \cdot \delta\vec{r}_i \\
&= -\sum_{i=1}^{N} m_i \ddot{\vec{r}}_i \cdot \sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j} \delta q_j \\
&= -\sum_{i=1}^{N} \sum_{j=1}^{n} \left(\frac{d}{dt}\left(\frac{\partial T_i}{\partial \dot{q}_j} \right) - \frac{\partial T_i}{\partial q_j} \right) \delta q_j \\
&= -\sum_{j=1}^{n} \left(\frac{d}{dt}\left(\frac{\partial \sum_{i=1}^{N} T_i}{\partial \dot{q}_j} \right) - \frac{\partial \sum_{i=1}^{N} T_i}{\partial q_j} \right) \delta q_j \\
&= -\sum_{j=1}^{n} \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \right) \delta q_j
\end{aligned} \tag{3.23}$$

where we have introduced the "total kinetic energy" of the particles in the system (and taken the liberty of swapping the order of summation, and bringing the summation into the derivatives):

$$T = \sum_{i=1}^{N} T_i = \sum_{i=1}^{N} \frac{1}{2} m_i \dot{r}_i^2 \tag{3.24}$$

Finally, the total virtual work from the external forces and the inertial forces is given by:

$$\begin{aligned}
\sum_{i=1}^{N} \delta W_i &= \sum_{i=1}^{N} (\vec{F}_i + \vec{I}_i) \cdot \delta\vec{r}_i = 0 \\
&= \sum_{j=1}^{n} Q_j \delta q_j + \sum_{i=1}^{N} \vec{I}_i \cdot \delta\vec{r}_i \\
&= \sum_{j=1}^{n} Q_j \delta q_j - \sum_{j=1}^{n} \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \right) \delta q_j \\
&= \sum_{j=1}^{n} \left[Q_j - \left(\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} \right) \right] \delta q_j
\end{aligned} \tag{3.25}$$

Since the virtual displacements of the generalized coordinates are independent (they can be varied independently from each other), each term in square brackets must be zero. We thus have one equation per degree of freedom q_j:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j} \right) - \frac{\partial T}{\partial q_j} = Q_j \tag{3.26}$$

which holds for a holonomic system (since we were able to transform to n independent generalized coordinates). This equation is equivalent to Newton's second law and thus gives the differential equations of motion for the system. Note the similarity of equation 3.26 and the equation that we obtained when calculating the variation of an integral subject to a constraint (equation 2.66 from Chapter 2). In Chapter 2, we had found that the integral of a function $L(q_1, q_2, \ldots, \dot{q}_1, \dot{q}_2, \ldots, t)$ subject to constraints $f_k(q_1, q_2, \ldots, t) = 0$

was stationary if:

$$\delta S = \int_a^b L(q_1, q_2, \ldots, \dot{q_1}, \dot{q_2}, \ldots, t)dt = 0$$

$$\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j}\right) = Q_j$$

$$Q_j \equiv \sum_{i=1}^{k} \lambda_i \frac{\partial f_i}{\partial q_j} \tag{3.27}$$

Thus, D'Alembert's principle is equivalent to requiring that the integral of of the kinetic energy, T, is stationary subject to a constraints related to how the position of the particles can vary with respect to the generalized coordinates:

$$\delta S = \int_a^b T(q_1, q_2, \ldots, \dot{q_1}, \dot{q_2}, \ldots, t)dt = 0$$

$$\left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j}\right) = Q_j$$

$$Q_j \equiv \sum_{i=1}^{N} \vec{F}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \tag{3.28}$$

where the applied forces appear to be related to Lagrange multipliers and the equations for the constraints are related to the coordinate transformations.

For completeness, we can re-write the total kinetic energy in terms of the generalized coordinates:

$$\begin{aligned}
T &= \sum_{i=1}^{N} \frac{1}{2} m_i \dot{r}_i^2 \\
&= \sum_{i=1}^{N} \frac{1}{2} m_i \frac{d\vec{r}_i}{dt} \frac{d\vec{r}_i}{dt} \\
&= \sum_{i=1}^{N} \frac{1}{2} m_i \left(\sum_{j=1}^{n} \frac{\partial \vec{r}_i}{\partial q_j}\dot{q}_j + \frac{\partial \vec{r}_i}{\partial t}\right)\left(\sum_{k=1}^{n} \frac{\partial \vec{r}_i}{\partial q_k}\dot{q}_k + \frac{\partial \vec{r}_i}{\partial t}\right) \\
&= \sum_{j=1}^{n}\sum_{k=1}^{n} A_{jk}\dot{q}_j\dot{q}_k + \sum_{k=1}^{n} B_k\dot{q}_k + C \\
&= T(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)
\end{aligned} \tag{3.29}$$

where the derivation of the terms A_{jk}, B_k, C, are left as an exercise. Note that if the coordinate transformations do not explicitly depend on time ($\frac{\partial \vec{r}_i}{\partial t} = 0$), then the kinetic energy reduces to:

$$T = \sum_{j=1}^{n}\sum_{k=1}^{n} A_{jk}\dot{q}_j\dot{q}_k \tag{3.30}$$

and is said to be quadratic in the velocities.

EXAMPLE 3-6: Find the terms A_{jk}, B_k, C for expressing the kinetic energy in polar coordinates for a system composed of a single particle in free space.

EXAMPLE 3-7: We can use D'Alembert's principle to obtain the equations of motion for a particle in a conservative field (such as gravity).

3.26

3.3 Conservation of energy from D'Alembert's principle

If the applied forces are conservative, they can be written as the gradient of a potential energy:

$$\vec{F_i} = -\nabla V_i = -\left(\frac{\partial V_i}{\partial x}\hat{x} + \frac{\partial V_i}{\partial y}\hat{y} + \frac{\partial V_i}{\partial z}\hat{z}\right) \tag{3.31}$$

The virtual work done by such a force is thus:

$$\begin{aligned}
\delta W &= \sum_{i=1}^{N} \vec{F_i} \cdot \delta\vec{r_i} \\
&= -\sum_{i=1}^{N} (\nabla V_i) \cdot \delta\vec{r_i} \\
&= -\sum_{i=1}^{N} \left(\frac{\partial V_i}{\partial x}\hat{x} + \frac{\partial V_i}{\partial y}\hat{y} + \frac{\partial V_i}{\partial z}\hat{z}\right)(\delta x\hat{x} + \delta y\hat{y} + \delta z\hat{z}) \\
&= -\sum_{i=1}^{N} \left(\frac{\partial V_i}{\partial x}\delta x + \frac{\partial V_i}{\partial y}\delta y + \frac{\partial V_i}{\partial z}\delta z\right) \\
&= -\sum_{i=1}^{N} \delta V_i
\end{aligned} \tag{3.32}$$

the negative of the change in potential energy, which makes sense. We can then write D'Alembert's principle as:

$$\begin{aligned}
\sum_{i=1}^{N} \vec{F_i} \cdot \delta\vec{r_i} - \sum_{i=1}^{N} m_i \ddot{\vec{r}}_i \cdot \delta\vec{r_i} = 0 \\
\sum_{i=1}^{N} \delta V_i + \sum_{i=1}^{N} m_i \ddot{\vec{r}}_i \cdot \delta\vec{r_i} = 0
\end{aligned} \tag{3.33}$$

We now consider a special case of the virtual displacement, namely, the case when $\delta\vec{r_i} = d\vec{r_i}$. Since we are free to choose any virtual displacement, we can choose the one that coincides with the true displacement in time. In the generalized coordinates, we can always choose $\delta q = dq$ without loss of generality. However, this does not always imply that $\delta\vec{r_i} = d\vec{r_i}$. We can only do this in the "scleronomic" case, when the generalized coordinates to not depend explicitly on time:

$$\begin{aligned}
\vec{r_i} &= \vec{r_i}(q_1, \dots q_n) \\
\frac{\partial \vec{r_i}}{\partial t} &= 0 \\
\therefore \delta q = dq \rightarrow \delta\vec{r_i} &= d\vec{r_i}
\end{aligned} \tag{3.34}$$

In the "rheonomic" case, when the transformations depend on time:

$$\begin{aligned}
\vec{r_i} &= \vec{r_i}(q_1, \dots q_n, t) \\
\frac{\partial \vec{r_i}}{\partial t} &\neq 0 \\
\therefore \delta q = dq \not\rightarrow \delta\vec{r_i} &= d\vec{r_i}
\end{aligned} \tag{3.35}$$

Similarly, the variation of the potential, δV, will equal the true change in the potential, if the potential energy does not depend explicitly on time:

$$\begin{aligned} V &= V(q_1, \ldots, q_n) \\ \frac{\partial V}{\partial t} &= 0 \\ \therefore \delta q = dq \rightarrow \delta V &= dV \end{aligned} \tag{3.36}$$

Keeping in mind that the following only holds for a scleronomic system when the potential energy does not depend on time, we can write the variations as true differentials with respect to time:

$$\begin{aligned} \sum_{i=1}^{N} \delta V_i + \sum_{i=1}^{N} m_i \ddot{\vec{r}}_i \cdot \delta \vec{r}_i \rightarrow & \sum_{i=1}^{N} dV_i + \sum_{i=1}^{N} m_i \ddot{\vec{r}}_i \cdot d\vec{r}_i \\ &= \sum_{i=1}^{N} dV_i + \sum_{i=1}^{N} m_i \ddot{\vec{r}}_i \cdot d\vec{r}_i \frac{dt}{dt} \\ &= \sum_{i=1}^{N} dV_i + \sum_{i=1}^{N} m_i \ddot{\vec{r}}_i \cdot \dot{\vec{r}}_i dt \\ &= \sum_{i=1}^{N} dV_i + \sum_{i=1}^{N} \frac{1}{2} m_i \frac{d}{dt}(\dot{r}_i^2) dt \\ &= \sum_{i=1}^{N} dV_i + \sum_{i=1}^{N} \frac{1}{2} m_i d(\dot{r}_i^2) \\ &= \sum_{i=1}^{N} dV_i + \sum_{i=1}^{N} dT_i \\ &= d \sum_{i=1}^{N} V_i + d \sum_{i=1}^{N} T_i \\ &= d(V + T) = 0 \\ \therefore \int d(V + T) = V + T \equiv E &= \text{constant} \end{aligned} \tag{3.37}$$

where we have introduced the kinetic energy, T_i for each particle (and its potential energy, V_i), as well as the total kinetic and potential energies of the system, T, and, V. We have implicitly assumed that the masses m_i are constant in time. We can see that in the specific case of a system that is scleronomic in potential energy and in transformation equations, the quantity $E = T + V$ is a constant of motion. Of course, we can identify this with the conservation of energy of the system, and the usual conditions for energy to be conserved.

3.4 Inertial forces in accelerated coordinate systems

In principle, if one is constrained into an elevator, it is impossible to distinguish whether the elevator is stationary in a gravitational field $-\vec{g}$ or whether it is in free space and accelerating "upwards" with an acceleration $\vec{g}$. This is illustrated in Figure 3.6, where a frame of reference (x',y') is shown relative to an "absolute fixed inertial frame of reference" (x,y). Measurements performed in the moving frame of reference are denoted with primes ('). The origin of the moving frame of reference is at a position $\vec{C}$ as measured in the absolute frame of reference.

If the position of a particle is described by a vector $\vec{r'}$ in the moving frame of reference, then in the absolute

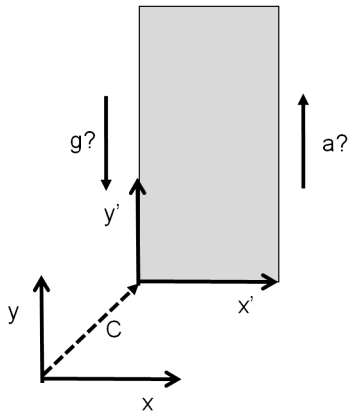

Figure 3.6: A moving frame of reference (x',y') relative to a fixed frame of reference (x,y).

frame of reference it is given by:

$$\vec{r} = \vec{C} + \vec{r'} \tag{3.38}$$

If the velocities and acceleration in the moving reference are given by $\dot{\vec{r'}}$, $\ddot{\vec{r'}}$, respectively, then, in the absolute frame of reference, they are given by:

$$\begin{aligned} \vec{v} &= \frac{d}{dt}(\vec{C} + \vec{r'}) \\ &= \dot{\vec{C}} + \dot{\vec{r'}} \\ \vec{a} &= \frac{d^2}{dt^2}(\vec{C} + \vec{r'}) \\ &= \ddot{\vec{C}} + \ddot{\vec{r'}} \end{aligned} \tag{3.39}$$

If we now consider D'Alembert's principle, we have:

$$\begin{aligned} \vec{F} + \vec{I} &= \vec{F} - m\vec{a} \\ &= \vec{F} - m\ddot{\vec{C}} - m\ddot{\vec{r'}} = 0 \end{aligned} \tag{3.40}$$

where we have an "apparent inertial force", $-m\ddot{\vec{C}}$, that is applied in addition to the inertial force, $-m\ddot{\vec{r'}}$. It is impossible to tell if one is in a system with a real force $-m\ddot{\vec{C}}$ or whether one is in an accelerated system with an apparent inertial force $-m\ddot{\vec{C}}$. The fact that this inertial force is proportional to the same mass as that which appears in the gravitational force forms the basis of the "equivalence principle" that leads to the General Theory of Relativity.

We also have the result that it is possible to choose a reference frame where a particle is at rest (possibly introducing an apparent force). This blurs the line of what we mean when referring to "inertial frames of reference". We do however recover Gallileo's relativity principle that there is no "absolute" reference frame, and that the description of a system between inertial frames of references is unchanged. By inertial frame of reference, we mean frames of reference that move with respect to each other in a straight line at a constant velocity. The forces that are apparent in a rotating frame of reference can be derived in a similar fashion, resulting in the apparent Coriolis force.

3.5 Problems

Problem 3-1: Suspension system
The mass m is sitting on top of two massless rods of length L. The left rod is fixed at point A and free to rotate about that point. The two rods are joined by a hinge at point B, and the rod on the right is free to slide without friction along the ground at point C. A spring with spring constant k is pushing horizontally against the rods at point C. When no mass is present, the system is in equilibrium with $\theta = \theta_0$ (that is, θ_0 corresponds to the case when the spring is not compressed, since the rods are massless).

Find the angle θ when the mass m is present. If you cannot solve for θ directly, give an equation that can be solved for the angle.

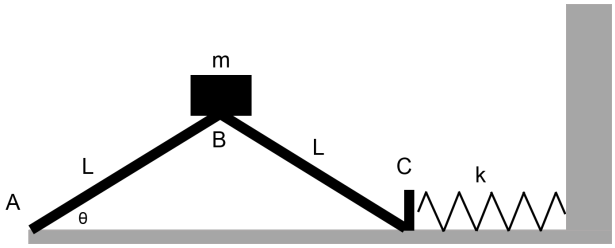

Figure 3.7: Mass on hinge and spring (Problem 3-1)

Problem 3-2: Masses on a scale
A simple scale is constructed with a massless plank and two masses m_1 and m_2 (see figure). If m_1, L_1, and L_2 are known, use the principle of virtual work to express m_2 in terms of the known quantities.

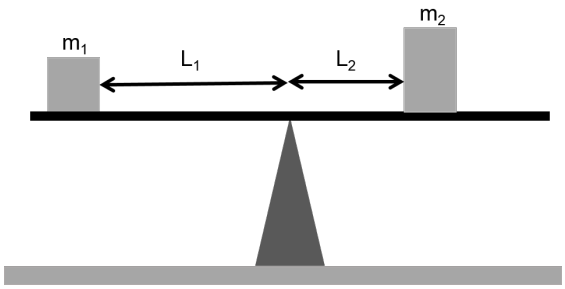

Figure 3.8: Two masses on a scale (Problem 3-2)

Problem 3-3: Kinetic energy in generalized coordinates
Evaluate the terms A_{jk}, B_k, and C from equation 3.29. For example, the term A_{jk} is given by:

$$A_{jk} = \frac{1}{2}\sum_{i=1}^{N} m_i \frac{\partial \vec{r}_i}{\partial q_j}\frac{\partial \vec{r}_i}{\partial q_k}$$

Problem 3-4: Generalized kinetic energy
A free particle has a kinetic energy $T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$ when expressed in Cartesian coordinates in a fixed reference system. Write the kinetic energy of the particle using the following systems of coordinates:
a) Polar(r, ϕ, z)
b) Spherical (r, ϕ, θ)
c) A Cartesian coordinate system (x', y', z') that is rotating about the z-axis with angular speed ω (assume that at $t = 0$ the xyz axes of the moving system coincided with that of the fixed coordinate system)
d) A polar coordinate system (r', ϕ', z') that is rotating about the z-axis with angular speed ω (assume that at $t = 0$ the xyz axes of the moving system coincided with that of the fixed coordinate system)

Problem 3-5: Moving pendulum
The pendulum in Figure 3.9 is composed of a mass m attached to a mass-less rigid rod of length L. The pendulum can swing in the xy-plane. The pivot point of the bar moves downwards at a fixed, known, speed v.

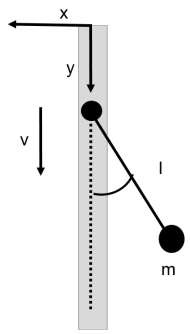

Figure 3.9: The mass m is attached by mass-less rigid rod of length L and free to swing in the xy-plane under the action of gravity. The pivot point moves with a fixed, known velocity, v, and was at the origin at time t = 0. (Problem 3-5)

a)Choose suitable generalized coordinates and write the kinetic energy in terms of the generalized coordinates
b)Use D'Alembert's principle to write the equations of motion in terms of the generalized coordinates.

Problem 3-6: Two masses and two springs
The figure shows two masses, m_1 and m_2, each connected to two springs with spring constants k_1 and k_2. Mass m_1 is constrained to slide without friction along the x-axis, whereas mass m_2 is constrained to move in the vertical direction, constrained by a massless frictionless vertical rod that is attached to m_1. Both springs have a resting length of L.
a) Choose suitable generalized coordinates, and write out the kinetic energy of the system in terms of those coordinates
b) Use D'Alembert's principle to write out the equations of motion for the generalized coordinates

Problem 3-7: Pendulum with a spring
The figure shows a bead of mass, m, that can slide freely along a long massless rail which has one end fixed at the origin, forming a pendulum. The mass is connected to the pivot point at the origin by a massless spring of resting length, L, and spring constant k. The motion is constrained to be in the vertical plane

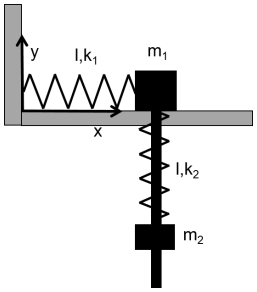

Figure 3.10: Two masses and two springs, problem 3-6

(gravity pointing downwards in the figure).

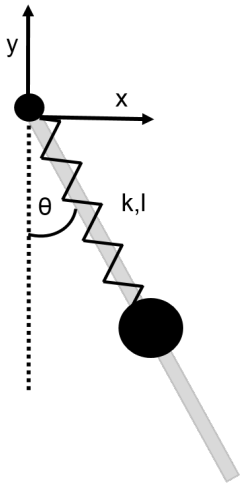

Figure 3.11: Pendulum with a spring, problem 3-7

a) Choose suitable generalized coordinates, and write out the kinetic energy of the system in terms of those coordinates
b) Use D'Alembert's principle to write out the equations of motion for the generalized coordinates

4 The Lagrangian

In this chapter, we introduce the "Lagrangian", a function of the generalized coordinates and velocities that is a very powerful method for describing physical systems. We will see that this scalar function, $L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)$, contains all of the information that is required to describe a system. We will see that all of classical mechanics can be described using the Lagrangian and the methods of the calculus of variations. We will find that properties of the Lagrangian result in a much deeper understanding of conservation laws.

4.1 The Lagrangian from D'Alembert's Principle

4.1.1 Single particle in conservative fields

We saw in the previous chapter, Example 3-7, that for a particle that is subject to a conservative force (a force that can be obtained from the gradient of the potential energy), D'Alembert's principle leads to a particularly elegant way of solving for the equations of motion. One needs only write the Lagrangian, $L = T - V$, given by the kinetic minus the potential energy of the particle, and find the condition for the action, $S = \int L dt$ to be stationary. Let's revisit that example. First, let's motivate the statement that V is the potential energy. If we have a force, F, that is given by the negative of the gradient of V:

$$\begin{aligned} \vec{F} &= -\nabla V \\ &= -\left(\frac{\partial V}{\partial x}\hat{x} + \frac{\partial V}{\partial y}\hat{y} + \frac{\partial V}{\partial z}\hat{z}\right) \end{aligned} \tag{4.1}$$

then the work done by the force along a path is given by:

$$\begin{aligned} W &= \int_{\vec{r}_1}^{\vec{r}_2} \vec{F} \cdot d\vec{r} \\ &= \int_{\vec{r}_1}^{\vec{r}_2} -\left(\frac{\partial V}{\partial x}\hat{x} + \frac{\partial V}{\partial y}\hat{y} + \frac{\partial V}{\partial z}\hat{z}\right) \cdot (dx\hat{x} + dy\hat{y} + dz\hat{z}) \\ &= -\int_{\vec{r}_1}^{\vec{r}_2} \left(\frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz\right) \end{aligned} \tag{4.2}$$

Note that this is just dV:

$$dV = \frac{\partial V}{\partial x}dx + \frac{\partial V}{\partial y}dy + \frac{\partial V}{\partial z}dz \tag{4.3}$$

Thus, the work is:

$$\begin{aligned} W &= -\int_{\vec{r}_1}^{\vec{r}_2} dV \\ &= -\left[V(\vec{r}_2) - V(\vec{r}_1)\right] \end{aligned} \tag{4.4}$$

which is precisely the definition of the potential energy (the negative of the work done by the force).

We are thus convinced that for a conservative force acting particle, $T - V$, really is the difference between kinetic and potential energy (sometimes called the "excess kinetic energy").

Recall that D'Alembert's principle gave us:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = Q_j$$
$$Q_j \equiv \sum_{i=1}^{N} \vec{F}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \tag{4.5}$$

where the sum over i is over all the forces on the particle. For a single particle, this can be the broken up into the sum of N_C conservative forces, $\vec{F}_i^C$, plus the sum of N_{NC} non-conservative forces, $\vec{F}_i^{NC}$:

$$\begin{aligned}
Q_j &= \sum_{i=1}^{N} \vec{F}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \\
&= \sum_{i=1}^{N_C} \vec{F}_i^C \cdot \frac{\partial \vec{r}_i}{\partial q_j} + \sum_{i=1}^{N_{NC}} \vec{F}_i^{NC} \cdot \frac{\partial \vec{r}_i}{\partial q_j} \\
&= -\sum_{i=1}^{N_C} \left(\frac{\partial V_i}{\partial x}\hat{x} + \frac{\partial V_i}{\partial y}\hat{y} + \frac{\partial V_i}{\partial z}\hat{z}\right) \cdot \frac{\partial}{\partial q_j}(x_i\hat{x} + y_i\hat{y} + z_i\hat{z}) + \sum_{i=1}^{N_{NC}} \vec{F}_i^{NC} \cdot \frac{\partial \vec{r}_i}{\partial q_j} \\
&= -\sum_{i=1}^{N_C} \sum_{k=1}^{3} \frac{\partial V_i}{\partial x_k} \frac{\partial x_{ik}}{\partial q_j} + \sum_{i=1}^{N_{NC}} \vec{F}_i^{NC} \cdot \frac{\partial \vec{r}_i}{\partial q_j} \\
&= -\sum_{i=1}^{N_C} \frac{\partial V_i}{\partial q_j} + \sum_{i=1}^{N_{NC}} \vec{F}_i^{NC} \cdot \frac{\partial \vec{r}_i}{\partial q_j} \\
&= -\frac{\partial V}{\partial q_j} + \sum_{i=1}^{N_{NC}} \vec{F}_i^{NC} \cdot \frac{\partial \vec{r}_i}{\partial q_j}
\end{aligned} \tag{4.6}$$

where we have introduced a potential energy, V_i, for each conservative force, and the total potential energy, V, as the sum of the potential energies. Note that the potential energy can now be expressed in terms of the generalized coordinates. Writing D'Alembert's principle, we get:

$$\begin{aligned}
\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} &= Q_j^C + Q_j^{NC} \\
&= -\frac{\partial V}{\partial q_j} + \sum_{i=1}^{N_{NC}} \vec{F}_i^{NC} \cdot \frac{\partial \vec{r}_i}{\partial q_j}
\end{aligned} \tag{4.7}$$

where we have split up the generalized forces into a conservative part and a non-conservative part. The conservative forces are written as:

$$Q_j^C = -\frac{\partial V}{\partial q_j} \tag{4.8}$$

If V does not depend on velocity, then:

$$\begin{aligned}
\frac{\partial V}{\partial \dot{q}_j} = 0 \rightarrow \frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_j} T - V\right) &= \frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_j} T\right) \\
\therefore \frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_j}(T - V)\right) - \frac{\partial}{\partial q_j}(T - V) &= \sum_{i=1}^{N_{NC}} \vec{F}_i^{NC} \cdot \frac{\partial \vec{r}_i}{\partial q_j}
\end{aligned} \tag{4.9}$$

which is a general version of D'Alembert's principle that includes non-conservative forces. This does require that the system be holonomic, that is, the generalized coordinates are independent of each other (since this

result depends on being able to take variations of the generalized coordinates independently). Introducing the Lagrangian, $L = T - V$, we get:

$$\therefore \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = \sum_{i=1}^{N_{NC}} \vec{F}_i^{NC} \cdot \frac{\partial \vec{r}_i}{\partial q_j} \tag{4.10}$$

EXAMPLE 4-1: Determine the equation of motion for a block of mass, m, that is sliding with friction down an inclined slope of angle θ. The friction force has a magnitude f.

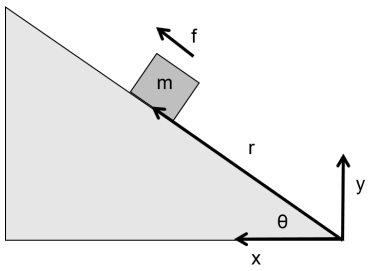

Figure 4.1: Block sliding with friction

4.1

4.1.2 The generalized potential

In the above, we considered the conservative generalized force given by equation 4.8:

$$Q_j^C = -\frac{\partial V}{\partial q_j}$$

When the potential did not depend on velocity, we could combine T and V conveniently into the Lagrangian:

$$\frac{\partial V}{\partial \dot{q}_j} = 0 \rightarrow \frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_j} T - V\right) = \frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_j} T\right)$$

$$\therefore \frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_j}(T-V)\right) - \frac{\partial}{\partial q_j}(T-V) = \sum_{i=1}^{N_{NC}} \vec{F}_i^{NC} \cdot \frac{\partial \vec{r}_i}{\partial q_j}$$

We can generalize the form of the generalized force to include forces that depend on a "generalized" potential, U, in the following manner:

$$Q_j^C = \frac{d}{dt}\frac{\partial U}{\partial \dot{q}_j} - \frac{\partial U}{\partial q_j}$$

Such a force is called "monogenic". If we construct the Lagrangian, $L = T - U$, we have, using D'Alembert's equation:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = \sum_{i=1}^{N_{NC}} \vec{F}_i^{NC} \cdot \frac{\partial \vec{r}_i}{\partial q_j}$$

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} - \frac{d}{dt}\frac{\partial U}{\partial \dot{q}_j} + \frac{\partial U}{\partial q_j} = \sum_{i=1}^{N_{NC}} \vec{F}_i^{NC} \cdot \frac{\partial \vec{r}_i}{\partial q_j}$$

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = Q_j^C + \sum_{i=1}^{N_{NC}} \vec{F}_i^{NC} \cdot \frac{\partial \vec{r}_i}{\partial q_j}$$

giving us back D'Alembert's equation in generalized coordinates. Without loss of generality, we will let V stand for any generalized potential and the Lagrangian L will include the potentials associated with any monogenic forces.

EXAMPLE 4-2: A charge, e, is moving with a velocity, $\vec{v}$, in a region of electric field, $\vec{E}$ and magnetic field $\vec{B}$ Both fields are free to vary in space and time. Show that a generalized potential given by $V = e\phi - e\vec{A}\cdot\vec{v}$ (where *phi* is the electric potential, $\vec{A}$ is the vector potential, and $\vec{v}$ is the velocity vector) results in the Lorentz force and write out the Lagrangian for the particle.

4.1.3 Multiple particles

So far the description in this chapter has been limited to a single particle in a monogenic field. We showed in the last chapter that D'Alembert's principle can be written generally as:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} = Q_j$$
$$Q_j \equiv \sum_{i=1}^{N} \vec{F}_i \cdot \frac{\partial \vec{r}_i}{\partial q_j} \tag{4.11}$$

where T is the sum of the kinetic energies of all the particles in the system, and Q_j is the sum of all the forces on all the particles in the system. In this chapter we treated Q_j as the sum of all the forces on one particle, then divided that up into monogenic and non-monogenic forces. We introduced V as the sum of all of the potential energies from all of the forces on that one particle. There is no loss in generalization if we claim that now V is the sum of all the potential energies related to all the monogenic forces acting on

all the particles in the system (the math is exactly the same and not worth repeating). Similarly, the sum of the monogenic forces on 1 particles is treated the same as the sum of the monogenic forces on all the particles.

EXAMPLE 4-3: Determine the equation of motion for two blocks of mass, m_1 and m_2, that are connected by a mass-less rod of length, L, and are sliding with friction down an inclined slope of angle θ. The friction forces have magnitudes f_1 and f_2.

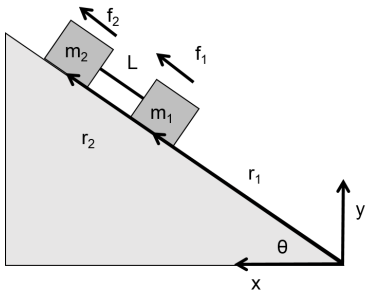

Figure 4.2: Two connected block sliding with friction

4.2

4.1.4 The connection to variational calculus and the use of auxiliary constraints

It should be clear now that systems can be described by the Lagrangian, $L = T - V$. Although we have shown abundantly that this is equivalent to D'Alembert's Principle (which itself is equivalent to Newton's Second Law), the approach is conceptually very different. The Lagrangian is only constructed using scalar quantities that are related to energy. The concept of a force is no longer fundamental to describing a system, and describing the energy is the fundamental concept. Generalized forces are in general not in the same units as those in Newton's Second Law, and, as we saw in the previous chapter, can also be related to torques.

In the case that all forces are monogenic, and the system is holonomic, the equations of motion, $q_i(t)$ for n degrees of freedom can be written as the requirement that the action, S, is stationary:

$$S \equiv \int_{t_1}^{t_2} L dt$$
$$\therefore \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0 \tag{4.12}$$

which are called the "Euler-Lagrange" equations from variational calculus. In the case where they are applied to mechanics, we often call them "Lagrange's equations of motion".

The requirement that the system be holonomic can be relaxed somewhat. We required that the system be holonomic so that the generalized coordinates can be varied independently. We however know from the calculus of variations that we can include "auxiliary conditions" between the generalized coordinates using Lagrange multipliers and treat the system as being holonomic. More precisely, if we have k equations of the form $f(q_1, \ldots q_n) = 0$, the Lagrangian can be modified to be:

$$\bar{L} \equiv L - \lambda_1 f_1 - \cdots - \lambda_k f_k \tag{4.13}$$

and the equations of motion are given by the Euler-Lagrange equations for the modified Lagrangian.

$$\frac{d}{dt}\left(\frac{\partial \bar{L}}{\partial \dot{q}_j}\right) - \frac{\partial \bar{L}}{\partial q_j} = 0 \tag{4.14}$$

Note that, mathematically, this is equivalent to introducing a series of "potential energies", V_i, of the form:

$$V_i = \lambda_i f_i \tag{4.15}$$

which correspond to monogenic forces given by:

$$Q_{ij} = \lambda_i \left(\frac{d}{dt}\frac{\partial f_i}{\partial \dot{q}_j} - \frac{\partial f_i}{\partial q_j}\right) \tag{4.16}$$

EXAMPLE 4-4: Determine the equations of motion for a hoop of mass M and radius R as is rolls without slipping down an incline of angle θ. Use the method of Lagrange multipliers to handle the constraint of rolling without slipping.

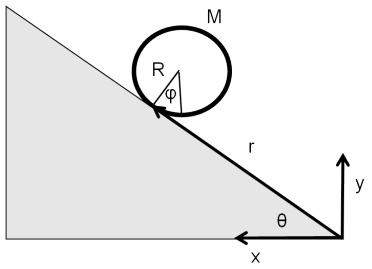

Figure 4.3: Hoop of mass, M and radius R rolling without slipping down an incline.

EXAMPLE 4-5: Determine the equations of motion for a hoop of mass M and radius R as is rolls without slipping down an incline of angle θ. Use the method of Lagrange multipliers to handle the constraint of rolling without slipping and that of the hoop being constrained to the incline.

It should become clear to the reader that the concept of force is getting more obscure as we arrive at a more general description of classical mechanics in the analytic framework. It is important to understand

that issues of constraint are often related to the forces from vectorial mechanics, although it is generally not necessary to use the concept of force to describe a system.

As we generalize further, we will also restrict ourselves to systems that are holonomic and monogenic. This is in fact a rather good approximation to nature. Most non-monogenic forces (for example, friction) arise from the fact that we limit ourselves to considering a restricted system (for example, we consider only a block sliding with friction). If we consider a system as a whole (for example, the block and the incline), then most forces become workless "internal" forces of constraint. Indeed, if we consider the Universe as a whole, then only four fundamental forces are responsible for all interactions (gravity, electromagnetic, strong nuclear, weak nuclear). Furthermore, in the classical macroscopic world, only the electromagntic and gravitational forces are of consideration, and these are both monogenic.

4.2 The Lagrangian from Hamilton's Principle

Hamilton's Principle is based on the idea that a physical system is fully described by a function of, $L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)$ that depends only on the generalized position and velocities of the particles in the system. The evolution of the system is determined by fixing the value of L at two positions in configuration time and requiring that the functional:

$$S = \int_{t_1}^{t_2} L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)dt \tag{4.17}$$

(called the action) is stationary. We have seen that if L is the Lagrangian, then Hamilton's principle is equivalent to D'Alembert's Principle and, thus, Newton's Second Law. From variational calculus, we know that Hamilton's principle leads to the Euler-Lagrange equations of motion for each coordinate q_i:

$$\delta S = 0 \rightarrow \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0 \tag{4.18}$$

4.2.1 The Lagrangian for a free particle

We can now ask ourselves what the function L can look like for describing a free particle. Our main requirement is that L should be a function that does not depend on the inertial frame of reference that we choose to describe the system. Given two frames of references, moving with a velocity $\vec{V}$ with respect to each other, the position, $\vec{r}$ and velocity of a particle, v, must transform according to:

$$\begin{aligned} \vec{r'} &= \vec{r} + \vec{V}t \\ \vec{v'} &= \vec{v} + \vec{V} \end{aligned} \tag{4.19}$$

Note that we implicitly assume that time is "absolute" and independent of reference frame. This of course is not true in Special Relativity.

The Lagrangian should be such that applying Hamilton's Principle is independent of the inertial frame of reference that is chosen. This means that the Lagrangian cannot depend on the absolute position of the particle. For a free particle, it should not matter where in space the particle is; it should always be described in the same way. We thus require:

$$\frac{\partial L}{\partial \vec{r}} = 0 \tag{4.20}$$

Similarly, the Lagrangian cannot depend on the direction of the velocity. Space is isotropic and the behaviour of the free particle will not depend on which direction it travels. The Lagrangian can thus only depend on its magnitude, v^2:

$$L = L(v^2) \tag{4.21}$$

$$\therefore \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0 \tag{4.22}$$

Lagrange's equations thus requires that $\frac{\partial L}{\partial \dot{q}_i}$ is a constant in time, thus requiring that the magnitude of v be constant in time. Of course, that is the result that we expect for a free particle. However, we don't know the form for the Lagrangian, in principle it could be any function of v^2, such as a polynomial, $L = av^2 + b(v^2)^2 + \ldots$.

Let's consider how things are affected when we transform the Lagrangian to a frame of reference that is moving with an infinitesimal velocity $\vec{\epsilon}$ with respect to the original inertial frame of reference. The particle thus has a velocity $\vec{v'} = \vec{v} + \vec{\epsilon}$

$$\begin{aligned} L(v'^2) &= L((\vec{v} + \vec{\epsilon}) \cdot (\vec{v} + \vec{\epsilon})) \\ &= L(v^2 + 2\vec{v} \cdot \vec{\epsilon} + \epsilon^2) \end{aligned} \tag{4.23}$$

We can expand this as a Taylor series of L centered at v^2, neglecting terms of order ϵ^2 and higher:

$$\begin{aligned} L(v'^2) &= L(v^2) + \frac{\partial L}{\partial v^2} 2\vec{v} \cdot \vec{\epsilon} + \ldots \\ &= L(v^2) + \frac{\partial L}{\partial v^2} x \end{aligned} \tag{4.24}$$

where $2\vec{v} \cdot \vec{\epsilon}$ is just some constant number that we call x (it is constant since $\vec{v}$ and $\vec{\epsilon}$ are constants). Applying Hamilton's Principle and using the Euler-Lagrange equations:

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L(v'^2)}{\partial \dot{q}_i}\right) - \frac{\partial L(v'^2)}{\partial q_i} &= 0 \\ \frac{d}{dt}\left(\frac{\partial L(v'^2)}{\partial \dot{q}_i}\right) &= 0 \\ \frac{d}{dt}\left(\frac{\partial L(v^2)}{\partial \dot{q}_i}\right) + \frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_i}\frac{\partial L}{\partial v^2} x\right) &= 0 \\ \frac{\partial}{\partial \dot{q}_i}\frac{\partial L}{\partial v^2} &= k \end{aligned} \tag{4.25}$$

where we used the fact that the original Lagrangian does not depend on position, that the velocity, $\vec{v}$ is constant, and we introduced k as another constant. For the last line to be true, L must be linear in the velocity squared:

$$L(v^2) = av^2 \tag{4.26}$$

Note that for a single particle, the Lagrangian can be multiplied by any number and remain invariant (the number a just factors out). In principle, if we have a system of N non-interacting particles, they will each have their own value of a, and only the ratios of the different values of a are relevant. We choose a new constant instead of a and call it $\frac{1}{2}m$, so that we have the Lagrangian:

$$L = \frac{1}{2}mv^2 \tag{4.27}$$

and we call m the "mass" of the particle. Note that we found this particular form of the Lagrangian simply by requiring that the Lagrangian be invariant under transformations from one inertial frame of reference to another. More fundamentally, we can see that for a system of non-interacting particles, it is the ratio of their masses that matters, not their absolute value.

4.3 Properties of the Lagrangian

The Lagrangian has several properties that we explore here.

4.3.1 Invariance with respect to point transformations

Coordinate transformations $q \rightarrow q'$ can generally be called "contact transformations" and written in the form:

$$q_i' = q_i'(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t) \tag{4.28}$$

The functional forms of the q_i' must satisfy certain conditions to be allowable (e.g. continuous, differentiable, non-zero Jacobian). When the transformation does not depend on the derivatives (velocities), the transformation is called a "point transformation". A point transformation is a mapping of a space onto a different space and requires that each point be mapped onto a unique point in the mapped spaced. Thus, points that are near each other remain near each other in the mapped space. A common mapping would be to change coordinates from cartesian to polar coordinates:

$$\begin{aligned} r &= r(x, y) = \sqrt{x^2 + y^2} \\ \theta &= \theta(x, y) = \tan^{-1}(\frac{y}{x}) \end{aligned} \tag{4.29}$$

In this case, a square in cartesian space does not map into a square in the space of polar coordinates, as seen in Figure 4.4. Other properties are however preserved. If two lines do not intersect in one space, then they won't intersect in the other space either. As one considers points that are closer together their geometrical properties in the mapped space become more similar. For example, a very small square in cartesian coordinates, maps into something that almost looks like a square in polar coordinates.

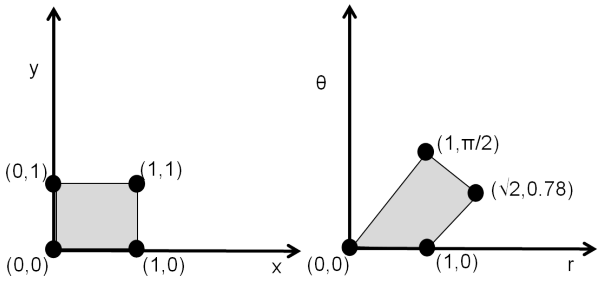

Figure 4.4: Mapping of a square in cartesian space to polar coordinate space.

Lagrange's equations determine the evolution of a system in n-dimensional configuration space (the space of the n generalized coordinates). One can extend the configuration space by one dimension, to include time as well. The evolution of a system is then a curve in the n+1 dimensional space, and the Lagrange equations determine the shape of that curve given that the end points are fixed. Through a point transformation, that curve will map to a curve that also follows the Lagrange equations in the mapped space (since infinitesimals are preserved through a point transformation). Thus, the Lagrangian description of motion is invariant under point transformations. That is, even though the actual functional form of the Lagrangian is different under a transformation of the coordinates, it invariably describes the same physical process.

This means that we are generally free to choose the set of generalized coordinates for which to describe the system (it makes sense that the Lagrangian description not depend on the choice of coordinates). This is one of the great powers of the Lagrangian method. One should be a little careful here, as describing a system using a moving frame of reference is still an acceptable point transformation. In Special Relativity, we know that the description of time depends on the relative speed of the two reference frames. We will see later that in the case of Special Relativity, we have to treat time the same way as one of the coordinates, and time will also be allowed to transform from one coordinate system to the other. With that modification, the Lagrangian will be truly independent of the coordinate system (provided they are related to an inertial system through a point transformation).

EXAMPLE 4-6: Describe a particle in a gravitational field using cartesian and polar coordinates

4.3.2 Addition of Lagrangians

The Lagrangians of different particles (or different systems) are additive. That is, given a Lagrangian L_A for a system A and a Lagrangian, L_B, for a system B, the entire system $A + B$ can be described by the Lagrangian $L = L_A + L_B$. This is easily shown by the additive properties of derivatives in the Euler-Lagrange equations:

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L_A}{\partial \dot{q}_i}\right) - \frac{\partial L_A}{\partial q_i} &= 0 \\ \frac{d}{dt}\left(\frac{\partial L_B}{\partial \dot{q}_i}\right) - \frac{\partial L_B}{\partial q_i} &= 0 \\ \therefore \frac{d}{dt}\left(\frac{\partial (L_A + L_B)}{\partial \dot{q}_i}\right) - \frac{\partial (L_A + L_B)}{\partial q_i} &= 0 \end{aligned} \tag{4.30}$$

4.3.3 Multiplication of the Lagrangian by a constant

The equations of motions are not affected if the Lagrangian is multiplied by an overall constant:

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} &= 0 \\ \therefore \frac{d}{dt}\left(\frac{\partial (kL)}{\partial \dot{q}_i}\right) - \frac{\partial (kL)}{\partial q_i} &= 0 \end{aligned} \tag{4.31}$$

For a free particle, multiplication of the Lagrangian by a constant is analogous to changing the units of mass.

4.3.4 Addition of a total time derivative

Similarly, the Lagrangian is un-affected if one adds to the Lagrangian a function that is a total time-derivative, $\frac{df}{dt}$:

$$\begin{aligned} S &= \int_{t_1}^{t_2} \left(L + \frac{df}{dt} \right) dt \\ &= \int_{t_1}^{t_2} L dt + \int_{t_1}^{t_2} \frac{df}{dt} dt \\ &= \int_{t_1}^{t_2} L dt + f(t_1) - f(t_2) \end{aligned} \tag{4.32}$$

When taking the variation of the action, the last two terms will vanish, since the end points of the L curve in configuration space are fixed. Of course, the Lagrangian is unchanged if a constant term is added to the Lagrangian (this can be thought of as a change in the absolute value of the potential energy).

4.3.5 Lagrangian does not explicitly depend on time

If the Lagrangian is independent of time, the Euler-Lagrange equations simplify, as we saw in the case of equation 2.57, when we looked at the calculus of variations. In the case where L does not explicitly depend on time, the Euler-Lagrange equations imply that:

$$\begin{aligned} \frac{\partial L}{\partial t} &= 0 \\ \therefore \left(\sum_{i=1}^{n} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} \right) - L &= h \end{aligned} \tag{4.33}$$

where h is a constant. In fact, we will see that h is related to the "Hamiltonian" of the system. h is called the "Jacobi integral" and can be calculated whether or not it is a constant by using the second equation. It is only a constant when L does not explicitly depend on time:

$$\begin{aligned} h(q_i, \dot{q}_i, t) &\equiv \left(\sum_{i=1}^{n} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} \right) - L \\ \frac{\partial L}{\partial t} = 0 &\rightarrow \frac{dh}{dt} = 0 \end{aligned} \tag{4.34}$$

One can in general re-write $h(q_i, \dot{q}_i, t)$ and replace $\dot{q}_i \rightarrow p_i$, and obtain a new function, $H(q_i, p_i, t)$ which is called the Hamiltonian.

Equation 4.33 is the general form for the quantity h that is conserved if L does not explicitly depend on time. If we make the following, further, assumptions:

1. The kinetic energy is of the form $T = \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} \dot{q}_j \dot{q}_k$, that is, quadratic in the velocities. Note that $a_{jk} = a_{kj}$, since these are partial derivatives of the generalized coordinates.

2. The potential energy, V, does not depend explicitly on the velocities ($\frac{\partial V}{\partial \dot{q}_i} = 0$).

The we have:

$$\begin{aligned}
\frac{\partial L}{\partial \dot{q}_i} &= \frac{\partial T}{\partial \dot{q}_i} \\
&= \frac{\partial}{\partial \dot{q}_i} \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} \dot{q}_j \dot{q}_k \\
&= \frac{1}{2} \left(\sum_{j=1}^{n} a_{ij} \dot{q}_j + \sum_{k=1}^{n} a_{ik} \dot{q}_k \right) \\
&= \sum_{k=1}^{n} a_{ik} \dot{q}_k \\
\therefore \left(\sum_{i=1}^{n} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} \right) &= \sum_{i=1}^{n} \sum_{k=1}^{n} \dot{q}_i a_{ik} \dot{q}_k \\
&= 2T \\
\therefore h &= \left(\sum_{i=1}^{n} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} \right) - L \\
&= 2T - T + V \\
&= T + V
\end{aligned} \tag{4.35}$$

and we see that, in the case where L does not explicitly depend on time, T is quadratic in the velocities, and V does not depend on velocities, the total energy of the system, $T + V$, is the conserved quantity h. This situation leads to a particularly simple treatment for problems with 1 degrees of freedom, as the solution for the equation of motion, $q(t)$ can always be written out as an integral:

$$\begin{aligned}
T(\dot{q}) &= h - V(q) \\
\therefore \dot{q} &= f(h, V) \\
dt &= \int_{q_a}^{q_b} \frac{dq}{f(h, V)}
\end{aligned} \tag{4.36}$$

which can then be inverted to obtain $q(t)$.

EXAMPLE 4-7: Calculate the period for a simple harmonic oscillator with spring constant k which has been released at $t = 0$ with starting position $x = x_0$ at rest.

4.36

4.3.6 Cyclic coordinates - Lagrangian does not depend on a specific coordinate

In some cases, the Lagrangian does not depend explicitly on all of the generalized coordinates. The Lagrangian will still depend implicitly on those coordinates through their velocities in the kinetic energy. We call those coordinates that do not explicitly appear in the Lagrangian "cyclic", or "ignorable", or "kinosthenic". In this case, the equations of motion also simplify for those coordinates:

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} &= 0 \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) &= 0 \\ \therefore \frac{\partial L}{\partial \dot{q}_i} &= p_i = \text{constant} \end{aligned} \tag{4.37}$$

where p_i are constants for each of the cyclic coordinates.

EXAMPLE 4-8: Determine the equations of motion for a particle in a gravitational field, where the Lagrangian does not depend on time, and find conserved quantities.

4.37

4.4 The generalized momentum

We saw in example 4-8, that in the case of a simple Lagrangian, the quantities $p_i \equiv \frac{\partial L}{\partial \dot{q}_i}$ are related to the linear momentum of a particle. For a more general system, the p_i, are called the generalized momenta of the system. We also saw that in the case where the Lagrangian does not explicitly depend on a coordinate, the corresponding generalized momentum is a constant of motion. We call the generalized momentum p_i the "conjugate momentum" of coordinate q_i. The generalized momenta are not necessarily related to the linear momentum of the system.

Inserting this into the Lagrange equations:

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i} \tag{4.38}$$

$$\frac{dp_i}{dt} - \frac{\partial L}{\partial q_i} = 0 \tag{4.39}$$

we see that we recover Newton's Second Law in the case where the p_i are linear momenta and the forces monogenic (thus given by the second term).

EXAMPLE 4-9: Determine the generalized momenta of a free particle described in polar coordinates

4.5 The Routhian and eliminating cyclic coordinates

In the case where we have cyclic coordinates, it is possible to use a method developed by Routh to eliminate them from the Lagrangian completely. Supposed that the Lagrangian has n degrees of freedom and the first k coordinates are cyclic. Thus:

$$\frac{\partial L}{\partial \dot{q}_i} = \beta_i \quad \text{(i=1...k)} \tag{4.40}$$

where the β_i are really just the generalizes momenta, but we use the β_i to distinguish the constant momenta. We first rewrite the Langrangian, L, by using the β_i to replace the k velocities corresponding to cylclic coordinates, $\dot{q}_i$. We then define the "Routhian" as:

$$\begin{aligned} R &\equiv L - \sum_{i=1}^{k} \beta_i \dot{q}_i \\ &= R(q_i, \dot{q}_i, \beta_i, t) \end{aligned} \tag{4.41}$$

The Routhian thus depends on $n-k$ of the q_i (and their velocities) in L, as well as k quantities β_i (in the sum), and time. Now consider the variation of the Routhian:

$$\delta R = \sum_{i=k+1}^{n} \frac{\partial R}{\partial q_i}\delta q_i + \sum_{i=k+1}^{n} \frac{\partial R}{\partial \dot{q}_i}\delta \dot{q}_i + \sum_{i=1}^{k} \frac{\partial R}{\partial \beta_i}\delta \beta_i + \frac{\partial R}{\partial t} \tag{4.42}$$

Using the definition of the Routhian, this must also equal:

$$\begin{aligned} \delta R =& \delta \left(L - \sum_{i=1}^{k} \beta_i \dot{q}_i \right) \\ =& \sum_{i=k+1}^{n} \frac{\partial L}{\partial q_i}\delta q_i + \sum_{i=k+1}^{n} \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i + \frac{\partial L}{\partial t} - \delta \left(\sum_{i=1}^{k} \beta_i \dot{q}_i \right) \\ =& \sum_{i=k+1}^{n} \frac{\partial L}{\partial q_i}\delta q_i + \sum_{i=k+1}^{n} \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i + \frac{\partial L}{\partial t} - \sum_{i=1}^{k} (\beta_i \delta \dot{q}_i + \dot{q}_i \delta \beta_i) \\ =& \sum_{i=k+1}^{n} \frac{\partial L}{\partial q_i}\delta q_i + \sum_{i=k+1}^{n} \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i + \frac{\partial L}{\partial t} - \sum_{i=1}^{k} \frac{\partial L}{\partial \dot{q}_i}\delta \dot{q}_i - \sum_{i=1}^{k} \dot{q}_i \delta \beta_i \end{aligned} \tag{4.43}$$

However, the second last term is zero, because we have explicitly rewritten the Lagrangian to remove the first k of the $\dot{q}_i$. We can thus compare the two versions of the equations for δR and match each term:

$$\begin{aligned} \frac{\partial L}{\partial q_i} &= \frac{\partial R}{\partial q_i} \\ \frac{\partial L}{\partial \dot{q}_i} &= \frac{\partial R}{\partial \dot{q}_i} \\ \dot{q}_i &= -\frac{\partial R}{\partial \beta_i} \quad \text{(i=1...k)} \\ \frac{\partial L}{\partial t} &= \frac{\partial R}{\partial t} \end{aligned} \tag{4.44}$$

Given these equivalences, it is clear that the Routhian will also follow the Lagrange equations:

$$\frac{d}{dt}\frac{\partial R}{\partial \dot{q}_i} - \frac{\partial R}{\partial q_i} = 0 \quad (i = k+1 \ldots n) \tag{4.45}$$

where we have now effectively reduced the number of degrees of freedom in the problem, since the Routhian only has $n-k$ coordinates. The equations of motion are thus given by:

$$\dot{q}_i = -\frac{\partial R}{\partial \beta_i} \quad \text{(i=1...k)}$$
$$\frac{d}{dt}\frac{\partial R}{\partial \dot{q}_i} - \frac{\partial R}{\partial q_i} = 0 \quad (i = k+1 \ldots n) \tag{4.46}$$

The procedure for solving a problem with the Routhian is:

1. Write the Lagrange and identify the cyclic coordinates
2. Calculate the β_i and use them to eliminate the corresponding $\dot{q}_i$ in the Lagrangian
3. Write the Routhian
4. The equations of motion for the k velocities from the cyclic variables are easily written in terms of the β_i
5. The equations of motion for the remaining $n-k$ coordinates are obtained by applying the usual Lagrange equations to R.

EXAMPLE 4-10: Use the Routhian to describe the motion of a particle in a gravitational field

4.6 Problems

Problem 4-1: The Atwood Machine
Figure 4.5 shows three point masses, m_1, m_2, and m_3 that are part of an Atwood machine with two pulleys. The pulleys are both solid disks of mass M and radius R. The ropes slide slide without slipping on the pulleys. The top rope has a total length L_1 and the bottom rope has a total length L_2.
a)Choose suitable generalized coordinates and write the Lagrangian for the system in terms of the generalized coordinates
b)Use the Lagrangian to obtain the equations of motion for the generalized coordinates

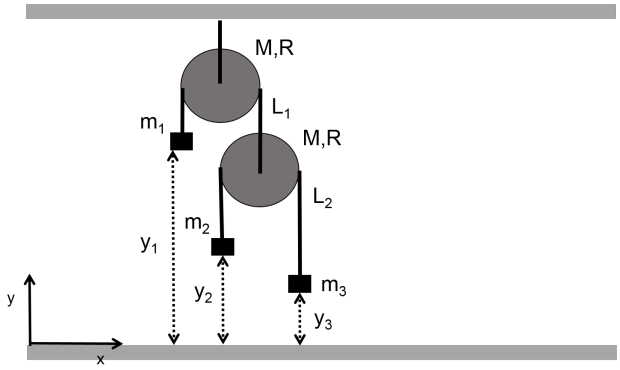

Figure 4.5: An Atwood machine.

Problem 4-2: Simple pendulum
The pendulum in Figure 4.6 is composed of a mass m attached to a mass-less rigid rod of the length L. The pendulum can swing in the xy-plane.

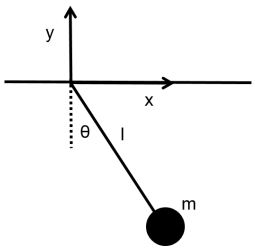

Figure 4.6: The mass m is attached by mass-less rigid rod of length L and free to swing in the xy-plane under the action of gravity. (Problem 4-2)

a)Choose suitable generalized coordinates and write the Lagrangian for the system in terms of the generalized coordinates
b)Use the Lagrangian to obtain the equations of motion for the generalized coordinates
c)Repeat parts a) and b) to obtain the equations of motion for the case where the rod has a mass M

Problem 4-3: Moving pendulum
The pendulum in Figure 4.7 is composed of a mass m attached to a mass-less rigid rod of the length L. The pendulum can swing in the xy-plane. The pivot point of the bar moves downwards at a fixed, known,

speed v.

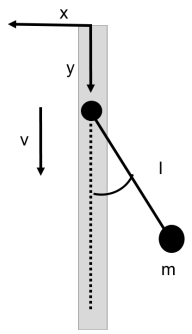

Figure 4.7: The mass m is attached by mass-less rigid rod of length L and free to swing in the xy-plane under the action of gravity. The pivot point moves with a fixed, known velocity, v, and was at the origin at time $t = 0$. (Problem 4-3)

a)Choose suitable generalized coordinates and write the Lagrangian for the system in terms of the generalized coordinates
b)Use the Lagrangian to obtain the equations of motion for the generalized coordinates.

Problem 4-4: Two masses and two springs
The figure shows two masses, m_1 and m_2, each connected to two springs with spring constants k_1 and k_2. Mass m_1 is constrained to slide without friction along the x-axis, whereas mass m_2 is constrained to move in the vertical direction, constrained by a massless frictionless vertical rod that is attached to m_1. Both springs have a resting length of L.

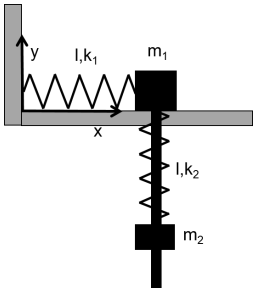

Figure 4.8: Two masses and two springs, problem 4-4

a)Choose suitable generalized coordinates and write the Lagrangian for the system in terms of the generalized coordinates
b)Use the Lagrangian to obtain the equations of motion for the generalized coordinates.

Problem 4-5: Pendulum with a spring
The figure shows a bead of mass, m, that can slide freely along a long massless rail which has one end fixed at the origin, forming a pendulum. The mass is connected to the pivot point at the origin by a massless spring of resting length, L, and spring constant k. The motion is constrained to be in the vertical plane (gravity pointing downwards in the figure).
a)Choose suitable generalized coordinates and write the Lagrangian for the system in terms of the generalized coordinates
b)Use the Lagrangian to obtain the equations of motion for the generalized coordinates.

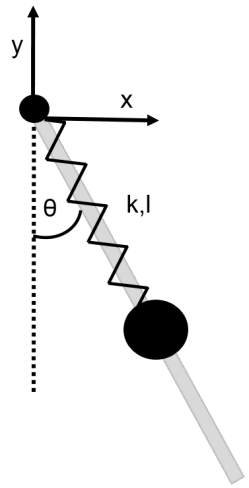

Figure 4.9: Pendulum with a spring, problem 4-5

Problem 4-6: Compound pendulum
The figure shows two beads of mass, m_1 and m_2, that form a compound pendulum. Mass m_1 is connected by a massless rigid rod of length L_1 to a fixed pivot point at the origin. Mass $m_@$ is connected to mass m_1 by a massless rigid rod of length L_2. The motion is constrained to be in the vertical plane

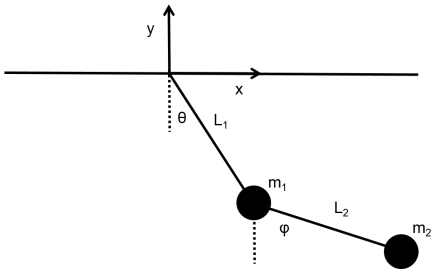

Figure 4.10: Pendulum with a spring (Problem 4-6).

a)Choose suitable generalized coordinates and write the Lagrangian for the system in terms of the generalized coordinates
b)Use the Lagrangian to obtain the equations of motion for the generalized coordinates.

Problem 4-7: Block on a hemisphere
Find the Lagrangian for a block of mass m sliding under the influence of gravity without friction along a hemisphere of radius R (Figure 4.11). Use the method of Lagrange multipliers to determine the normal force exerted by the hemisphere on the block. Find the point at which the block will fall off the hemisphere if it started at rest at the top.

Problem 4-8: Sphere on a hemisphere
Find the Lagrangian for a sphere of mass m and radius r that is rolling without slipping along a hemisphere of radius R (Figure 4.12). Use the method of Lagrange multipliers to determine the forces exerted by the hemisphere on the sphere. Find the point at which the sphere will fall off the hemisphere if it started at

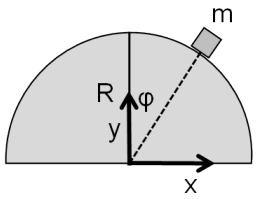

Figure 4.11: A block sliding without friction on a hemisphere (Problem 4-7)

rest at the top.

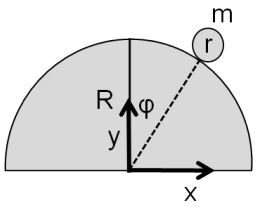

Figure 4.12: A ball rolling without slipping on a hemisphere (Problem 4-8).

Problem 4-9: Sliding blocks
The block of mass m in Figure 4.13 can slide without friction on the wedge of mass M that itself can slide without friction on the ground. The dimensions of the small block can be assumed to be negligible, the dimensions of the wedge are shown in the figure, and the motion is constrained to be in the xy-plane.

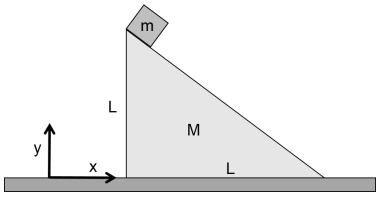

Figure 4.13: The mass m can slide without friction on the wedge of mass M which itself can slide with no friction on the ground (Problem 4-9)

a)Choose suitable generalized coordinates and write the Lagrangian for the system in terms of the generalized coordinates
b)Use the Lagrangian to obtain the equations of motion for the generalized coordinates.

Problem 4-10: Sphere on a wedge
The sphere of mass m and radius r in Figure 4.14 can roll without slipping on the wedge of mass M that itself can slide without friction on the ground. The dimensions of the wedge are shown in the figure, and the motion is constrained to be in the xy-plane.
a)Choose suitable generalized coordinates and write the Lagrangian for the system in terms of the generalized coordinates
b)Use the Lagrangian to obtain the equations of motion for the generalized coordinates.

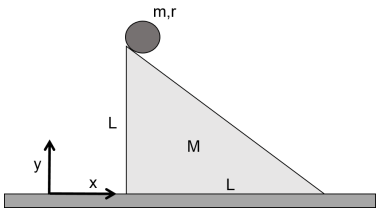

Figure 4.14: The sphere of mass m and radius r rolls without slipping on the wedge of mass M which itself can slide with no friction on the ground (Problem 4-10)

Problem 4-11: Sliding blocks with a spring

The block of mass m in Figure 4.15 can slide without friction on the wedge of mass M that itself can slide without friction on the ground. The dimensions of the small block can be assumed to be negligible, the dimensions of the wedge are shown in the figure, and the motion is constrained to be in the xy-plane. The small block is attached to a spring with rest length d and spring constant k.

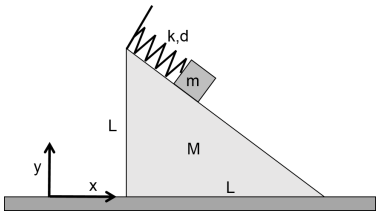

Figure 4.15: The mass m can slide without friction on the wedge of mass M which itself can slide with no friction on the ground. The mass m is connected to M with a spring with spring constant k and rest length d (Problem 4-11)

a)Choose suitable generalized coordinates and write the Lagrangian for the system in terms of the generalized coordinates

b)Use the Lagrangian to obtain the equations of motion for the generalized coordinates.

Problem 4-12: Routhian and a spherical pendulum

A spherical pendulum is constructed by using a mass m and a mass-less rod of length L as seen in in Figure 4.16 .

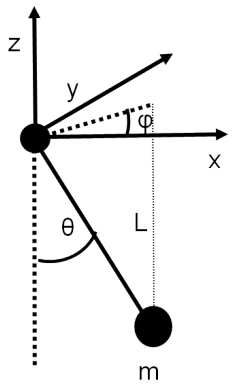

Figure 4.16: A spherical pendulum with mass m attached to a rigid mass-less rod of length L (Problem 4-12)

a)Write out the Lagrangian for the spherical pendulum using θ and ϕ as the generalized coordinates

b)Write out expressions for any conserved quantities.
c)Show that the system is described by the following Routhian:

$$R = \frac{1}{2}mL^2\dot{\theta}^2 - \frac{1}{2}\frac{\beta_\phi^2}{mL^2\sin^2\theta} + mgL\cos\theta$$

5 Conservations laws and symmetries

We saw in the previous chapter that cyclic coordinates lead to their conjugate momenta being constants of motion. We also saw that when the Lagrangian does not depend explicitly on time, the total energy of the system is conserved (or more precisely, the Jacobi integral is conserved). These conservation laws can be put in a more general form that states that for each symmetry in the action, there is a corresponding conserved quantity. This is called Noether's theorem.

5.1 Conserved generalized momenta

Given a Lagrangian, L, the generalized momenta are defined as:

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i} \tag{5.1}$$

The Lagrange equations of motion are thus:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0$$

$$\therefore \dot{p}_i = \frac{\partial L}{\partial q_i} \tag{5.2}$$

Thus, if the Lagrangian does not depend explicitly on q_i, the generalized momentum p_i is a constant (i.e. it does not change with time). The interesting point to note is that the conserved generalized momenta depend on the choice of generalized coordinates.

EXAMPLE 5-1: Compare the conserved quantities for a particle moving in a potential that depends on the distance from the origin for cartesian and spherical coordinates

---

The question then arises of how to determine if there are any conserved quantities or an optimal choice of generalized coordinates that uncover those conserved quantities. In this chapter, we will focus on uncovering the conserved quantities. In a later chapter, we will discuss "Canonical Transformations", which is the method of discovering a set of generalized coordinates where all generalized momenta are conserved (at the expense that the coordinate transformations are not straightforward).

5.2 Continuous Transformations

We start by considering "continuous transformations" of coordinates. Transformations of coordinates are continuous if they depend on a parameter, ϵ that can be varied continuously:

$$\begin{aligned} q_i' &= q_i'(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t, \epsilon) \\ t' &= t'(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t, \epsilon) \end{aligned} \tag{5.3}$$

where we have included the possibility that time can also be transformed. Since ϵ can be varied continuously, the transformations can be built up as the sum of infinitesimal transformations. It is intended that when $\epsilon = 0$ then $q_i' = q_i$. For an infinitesimal transformation, we can introduce functions f and g and write:

$$\begin{aligned} q_i' &= q_i + \delta q_i = q_i + f_i(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)\delta\epsilon \\ t' &= t + \delta t = t + g(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)\delta\epsilon \\ \delta q_i &= q_i' - q_i = f_i(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)\delta\epsilon \\ \delta t &= t' - t = g(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n, t)\delta\epsilon \end{aligned} \tag{5.4}$$

EXAMPLE 5-2: Determine the infinitesimal form of a continuous transformation corresponding to a translation along the x-axis of a distance ϵ

EXAMPLE 5-3: Determine the infinitesimal form of a continuous transformation corresponding to an infinitesimal rotation about the z-axis by angle ϵ.

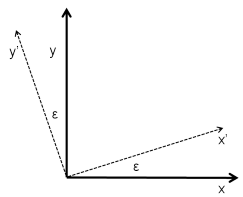

Figure 5.1: Infinitesimal rotation of the axes by an angle ϵ around the z-axis

5.3 Noether's theorem

Noether's theorem states that any continuous transformation that does not change the action is associated with a conserved quantity (we call the quantity a "charge", Q). We say that the action is invariant under such a continuous transformation, and we mean that:

$$S = \int_{t_a}^{t_b} L(q_i, \dot{q}_i, t)dt = \int_{t'_a}^{t'_b} L(q'_i, \dot{q}'_i, t)dt \tag{5.5}$$

where on the right hand side, since t is just a dummy integration variable we left it as t instead of t' (note however that we did need to change the limits of the integral to be in the transformed time). As a notation, we imply that $L(q_i, \dot{q}_i, t)$ depends on all the q_i.

Because the continuous transformations are continuous we can consider the case of the infinitesimal version of a given transformation. If the action is invariant under the infinitesimal transformation, it will also be invariant under a series of infinitesimal transformations that lead to a total transformation. Requiring invariance under the infinitesimal transformation:

$$\begin{aligned} S &= \int_{t_a}^{t_b} L(q_i, \dot{q}_i, t)dt = \int_{t_a+\delta t_a}^{t_b+\delta t_b} L(q_i + \delta q'_i, \dot{q}_i + \delta \dot{q}'_i, t)dt \\ \therefore \delta S &= \int_{t_a+\delta t_a}^{t_b+\delta t_b} L(q_i + \delta q'_i, \dot{q}_i + \delta \dot{q}'_i, t)dt - \int_{t_a}^{t_b} L(q_i, \dot{q}_i, t)dt \end{aligned} \tag{5.6}$$

where the $\delta t_{a(b)} = g(t_a(t_b))\delta\epsilon$ in the limits of the integral, can in principle depend on t (through g) and are thus different at the two end points. If the action is invariant under the transformation, then $\delta S = 0$.

Under the condition that $\delta S = 0$ the following quantity Q is conserved (does not vary with time):

$$\begin{aligned} Q &\equiv \sum_{i=1}^{n} p_i \delta q_i + \left(\sum_{i=1}^{n} p_i \dot{q}_i - L \right) \delta t \\ &= \sum_{i=1}^{n} p_i f_i + \left(\sum_{i=1}^{n} p_i \dot{q}_i - L \right) g \end{aligned} \tag{5.7}$$

where we have (re-)introduced $\delta q_i = f_i \delta\epsilon$ and $\delta t = g\delta\epsilon$. Note that there are two terms, one is associated with invariance in time, g, and one is associated with invariance in space, f. You may recognize the one related to time as the Jacobi integral.

Note that if a symmetry preserves the Lagrangian, the action will also be preserved. It is often more straightforward to verify if the Lagrangian is invariant under a transformation. It is also more straightforward to only consider variations in the position coordinates.

Consider the variation of the Lagrangian with respect to the coordinate transformation:

$$\begin{aligned} q_i' &= q_i + f_i(q_1, \ldots, q_n)\delta\epsilon \\ \delta q_i &= \delta\epsilon f_i(q_1, \ldots, q_n) \\ \delta L &= \sum_{i=1}^{n} \frac{\partial L}{\partial q_i}\delta q_i + \frac{\partial L}{\partial \dot{q}_i}\delta\dot{q}_i \end{aligned} \tag{5.8}$$

Recall the L must still satisfy Lagrange's equations, so we have:

$$\begin{aligned} p_i &= \frac{\partial L}{\partial \dot{q}_i} \\ \dot{p}_i &= \frac{\partial L}{\partial q_i} \end{aligned} \tag{5.9}$$

which we can substitute back into the variation of the Lagrangian:

$$\begin{aligned} \delta L &= \sum_{i=1}^{n} \dot{p}_i\delta q_i + p_i\delta\dot{q}_i \\ &= \frac{d}{dt}\sum_{i=1}^{n} p_i\delta q_i \\ &= \delta\epsilon\frac{d}{dt}\sum_{i=1}^{n} p_i f_i(q_1, \ldots, q_n) \end{aligned} \tag{5.10}$$

which must be true for any arbitrary ϵ. Thus, if the variation of the Lagrangian is zero then the quantity:

$$Q = \sum_{i=1}^{n} p_i f_i(q_1, \ldots, q_n) \tag{5.11}$$

is conserved, which is one of the quantities we had before. In principle, the variation of L can be equal to zero within a total time derivative, since adding a total time derivative to the Lagrangian does not change the equations of motion.

Note that Noether's theorem only applies to **continuous** transformations that leave the action or the Lagrangian unchanged. For example, a transformation that mirrors the coordinate system may preserve the Lagrangian, but it is not a continuous transformation, so does not have an associated conserved quantity. There may be other continuous symmetries that are not straightforward coordinate transforms (such as those that we considered in this proof) that may lead to less obvious conserved quantities.

5.4 Translation symmetry

We can find the conserved quantities if the Lagrangian is invariant to translation. For a translation, the transformation equations are:

$$q_i' = q_i + \delta\epsilon f_i \tag{5.12}$$

where the function f_i is independent of the coordinates. Consider a translation in the q_1 direction, where $f_1 = 1$ and all other $f_i = 0$. Since time is not involved in the transformation, we also set $g = 0$, or consider only the invariance of the Lagrangian. If the Lagrangian is invariant under translations in the q_1 direction, then the quantity:

$$Q = \sum_{i=1}^{n} p_i f_i(q_1, \ldots, q_n) = p_1 \tag{5.13}$$

is conserved. That is, we recover the observation that was made about cyclic coordinates: if a coordinate does not appear explicitly in the Lagrangian (thus the Lagrangian is invariant when that coordinate is transformed), the conjugate momentum for that coordinate is conserved.

More generally, for each direction in space that the Lagrangian is symmetric, momentum in that direction is a conserved quantity. In cartesian coordinates, this is the principle of conservation of momentum.

5.5 Rotational symmetry

Consider a Lagrangian that is invariant under a rotation about the z-axis. We saw in example 5-3 that for a rotation about the z-axis in cartesian coordinates, we have:

$$f_1 = -y \tag{5.14}$$
$$f_2 = x \tag{5.15}$$
$$f_3 = 0 \tag{5.16}$$
$$g = 0 \tag{5.17}$$

The conserved quantity is thus:

$$Q = \sum_{i=1}^{n} p_i f_i = -p_x y + p_y x \tag{5.18}$$

which is the z component of angular momentum:

$$\vec{L} = \vec{r} \times \vec{p} \tag{5.19}$$

In general, the angular momentum in a direction is conserved if the Lagrangian is invariant to rotations about an axis in that direction.

5.6 Time symmetry

In the case of a time symmetry, we have $f_i = 0$ and $g = 1$:

$$t' = t + \delta\epsilon \tag{5.20}$$

This gives the conserved quantity:

$$Q = \sum_{i=1}^{n} p_i \dot{q}_i - L \tag{5.21}$$

We have already seen that when the Lagrangian does not explicitly depend on time, the quantity:

$$h = \sum_{i=1}^{n} p_i \dot{q}_i - L \tag{5.22}$$

is conserved. Of course, if the Lagrangian does not depend on time, then the action will be invariant to a time translation.

5.7 Problems

Problem 5-1: Spherical pendulum
Write the Lagrangian for a spherical pendulum (a bob of mass m attached to a mass-less rigid rod of length l) using spherical coordinates and identify all conserved quantities. A spherical pendulum is the generalized case of the simple pendulum when the mass is not constrained to swing in a plane.

Problem 5-2: Two masses on a spring
Two masses, m_1 and m_2, are connected by a spring of rest length l and spring constant k. The two masses are constrained to move in one dimension, along the x-axis, on a friction-less surface.
a) Give the Lagrangian for the system and write the equations of motion for the two masses

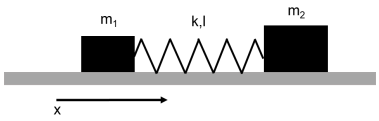

Figure 5.2: Two blocks connected by a spring slide on a friction-less surface. (Problem 5-2)

b) Show that the total linear momentum in the x-direction, P_x, is conserved ($P_x = m_1 v_1 + m v_2$)
c) List all conserved quantities for the system.

Problem 5-3: Arbitrary potential
Calculate the conserved quantities for the Lagrangian:

$$L = \frac{1}{2} m_1 \dot{q}_1^2 + \frac{1}{2} m_2 \dot{q}_2^2 - V(a q_1 - b q_2)$$

where a and b are constants, and V() is some unknown function of the linear combination $a q_1 - b q_2$.

Problem 5-4: Spring pendulum with two masses
The pendulum in figure 5.3 is constructed with two masses, m_1 and m_2, a spring of constant k and rest length d. Mass m_2 is fixed at the end of a mass-less rigid rod of length l, while m_1 can slide without friction along the rod and is connected to the spring.
a) Give the Lagrangian for the system and write the equations of motion for the two masses
b) List all conserved quantities for the system.

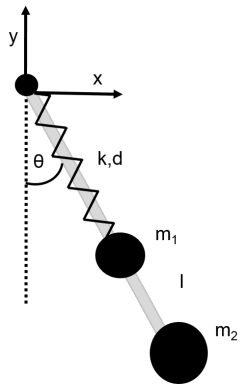

Figure 5.3: A pendulum with two masses, one of which is connected to the pivot point by a spring (Problem 5-4).

6 Applications of the Lagrangian formalism

In this chapter, we pause and look at some common applications of Lagrangian mechanics. In particular, we look at the central force problem and the problem of small oscillations. Both of these situations are of interest as many problems can be formulated in similar frameworks, and the techniques that are used here can often be applied in other situations.

6.1 Central Force Problem

6.1.1 One particle with 1 degree of freedom

Consider a simple system, with one (cartesian) degree of freedom and a potential of the form $V(q_1)$, so that the Lagrangian is:

$$L = \frac{1}{2}m\dot{q}_1^2 - V(q_1) \tag{6.1}$$

This is called a "central force problem", since it represents a particle of mass m with a force that is related to the distance from the origin and is directed towards or away from the origin. Example of such forces are Newtonian gravity, the Coulomb force, and the spring force. Since the Lagrangian does not explicitly depend on time, the total energy, E is conserved and is a constant. This allows the velocity to be determined from the position

$$\begin{aligned} E &= \frac{1}{2}m\dot{q}_1^2 + V(q_1) \\ \dot{q}_1 &= \sqrt{\frac{2}{m}(E - V(q_1))} \end{aligned} \tag{6.2}$$

Note that this implies that $E - V(q_1) > 0$, or the velocity would be imaginary.

The velocity equation can be integrated to get time as a function of position, and in principle, inverted to get position as a function of time:

$$\int_{t_0}^{t} dt = \int_{q_0}^{q_1(t)} \frac{dq_1}{\sqrt{\frac{2}{m}(E - V(q_1))}} \tag{6.3}$$

6.1.2 Two particles and two degrees of freedom

Now consider two particles of mass m_1, m_2, each having a single degree of freedom and interacting with each other through a force that is related to the distance between the two particles. The total potential energy for the system would thus have a form $V(q_2 - q_1)$ and the Lagrangian is:

$$L = \frac{1}{2}m_1\dot{q}_1^2 + \frac{1}{2}m_2\dot{q}_2^2 - V(q_2 - q_1) \tag{6.4}$$

Again, the total energy of the system is a constant, since the Lagrangian does not depend explicitly on time. Interpreting q_i as the position of each particle along a q-axis, then the Lagrangian is invariant along translations in the q-direction:

$$q_1 \to q_1 + \epsilon \tag{6.5}$$
$$q_2 \to q_2 + \epsilon \tag{6.6}$$

so there must be a conserved quantity:

$$\begin{aligned} Q &= \sum_i p_i \\ &= m_1\dot{q}_1 + m_2\dot{q}_2 \end{aligned} \tag{6.7}$$

which is the total momentum in the q-direction.

Consider the following change of variables from (q_1,q_2) to (r, R):

$$\begin{aligned} r &\equiv q_2 - q_1 \\ R &\equiv \frac{m_1}{m_1+m_2}q_1 + \frac{m_2}{m_1+m_2}q_2 \\ M &\equiv m_1 + m_2 \\ \mu &\equiv \frac{m_1 m_2}{m_1+m_2} \end{aligned} \tag{6.8}$$

where R is the coordinate of the center of mass, and μ is called the "reduced mass".

EXAMPLE 6-1: Invert the transformation equations 6.8 for q_1 and q_2 to obtain the kinetic energy in terms of r and R.

In terms of the new coordinates, the Lagrangian is then:

$$L = \frac{1}{2}M\dot{R}^2 + \frac{1}{2}\mu\dot{r}^2 - V(r) \tag{6.9}$$

The R coordinate is cyclic, and the corresponding momentum:

$$P = M\dot{R} \tag{6.10}$$

corresponding to the momentum of the whole system is conserved. The kinetic energy term from the center of mass is thus a constant added to the Lagrangian and does not influence the equations of motion; we can thus ignore it:

$$L = \frac{1}{2}\mu\dot{r}^2 - V(r) \tag{6.11}$$

where we have effectively reduced the problem to an equivalent problem with one degree of freedom and the reduced mass.

6.1.3 Two particles and three dimensional motion

Let us know consider the general problem of two particles of mass m_1 and m_2, with position vectors $\vec{r}_1$ and $\vec{r}_2$ and a potential energy of the form $V(\vec{r}_2 - \vec{r}_2)$. The Lagrangian is:

$$L = \frac{1}{2}m_1\dot{\vec{r}}_1^2 + \frac{1}{2}m_2\dot{\vec{r}}_2^2 - V(\vec{r}_2 - \vec{r}_1) \tag{6.12}$$

Again, we can introduce new coordinates and the reduced mass:

$$\begin{aligned}
\vec{r} &\equiv \vec{r}_2 - \vec{r}_1 \\
\vec{R} &\equiv \frac{m_1}{m_1 + m_2}\vec{r}_1 + \frac{m_2}{m_1 + m_2}\vec{r}_2 \\
\mu &\equiv \frac{m_1 m_2}{m_1 + m_2} \\
\therefore \vec{r}_1 &= \vec{R} - \frac{m_2}{m_1 + m_2}r \\
\therefore \vec{r}_2 &= \vec{R} + \frac{m_1}{m_1 + m_2}r
\end{aligned}$$

The algebra to obtain the kinetic energy in the new coordinates is the same as in the previous section, and the Lagrangian can be written as:

$$L = \frac{1}{2}\mu\dot{\vec{r}}^2 - V(\vec{r}) \tag{6.13}$$

where we have used the fact that the total momentum is a constant of the motion:

$$P = M\dot{\vec{R}} \tag{6.14}$$

Thus, in 3-dimensional space, the problem of two particles acting on each other through a central force (6 degrees of freedom) can be reduced to an equivalent problem with 1 particle acted on by a central force

directed at the origin (3 degrees of freedom). As origin, we use the center of mass of the two particles, with a coordinate system that is co-moving with the center of mass.

Since space is homogeneous, the potential cannot depend on the actual direction of $\vec{r}$ and should only depend on its magnitude. Indeed, if two particles acting on each other through a central potential (e.g. Coulomb force), the orientation in space of the particles will not matter. It is thus natural to describe the system using spherical coordinates, r, θ, ϕ, with the Lagrangian:

$$L = \frac{1}{2}\mu(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\dot{\phi}^2) - V(r) \tag{6.15}$$

The Lagrangian is invariant under rotations of the coordinate system, thus the angular momentum in each direction is conserved. Since the total angular momentum is conserved, the motion must be constrained to the place perpendicular to the total angular momentum. We thus choose to orient our coordinate system such that the polar axis points in the direction of the total angular momentum. The angle θ will thus be $\frac{\pi}{2}$ and constant. The Lagrangian thus simplifies even further to:

$$L = \frac{1}{2}\mu(\dot{r}^2 + r^2\dot{\phi}^2) - V(r) \tag{6.16}$$

The coordinate ϕ is cyclic, and the corresponding generalized momentum is conserved:

$$\begin{aligned} p_\phi &= \mu r^2\dot{\phi} = l \\ \therefore \dot{\phi} &= \frac{l}{r^2\mu} \end{aligned} \tag{6.17}$$

Recall, this does not mean that $\dot{\phi}$ is constant, only that it can be determined from a constant and from r. The constant, l, is obviously angular momentum. We can thus write the Lagrangian as:

$$L = \frac{1}{2}\mu\dot{r}^2 + \frac{l^2}{2\mu r^2} - V(r) \tag{6.18}$$

Since the middle term depends only r, we can introduce an "effective potential", V_{eff}

$$V_{eff} \equiv V(r) + \frac{l^2}{2\mu r^2} \tag{6.19}$$

and the Lagrangian is given by:

$$L = \frac{1}{2}\mu\dot{r}^2 - V_{eff}(r) \tag{6.20}$$

and we have again reduced the problem to a single degree of freedom!

Consider the force that arises from the effective potential:

$$\begin{aligned} F_r &= -\frac{\partial V_{eff}}{\partial r} \\ &= -\frac{\partial V(r)}{\partial r} - \frac{\partial}{\partial r}\frac{l^2}{2\mu r^2} \\ &= -\frac{\partial V(r)}{\partial r} - \frac{\partial}{\partial r}\frac{1}{2}\mu r^2\dot{\phi}^2 \\ &= -\frac{\partial V(r)}{\partial r} + \mu r\dot{\phi}^2 \end{aligned} \tag{6.21}$$

The second term (which we added) is in fact the apparent centrifugal force, trying to push the "particle" outwards (recall that we are actually modeling two particles in 3 dimensions!). This apparent force is zero if $\dot{\phi}$ is zero.

Finally, conservation of energy allows us to solve for $r(t)$ by inverting the solution to:

$$t = \int_{r_0}^{r(t)} \frac{dr}{\sqrt{\frac{2}{\mu}(E - V_{eff}(r))}} \tag{6.22}$$

Using analytic mechanics we have thus been able to understand a lot about a system with a central force, without solving any differential equations and without even knowing the form of the force (or potential)! We have found that:

1. A system of two particles interacting through a central force in three dimensions can be reduced to an equivalent system with 1 degree of freedom

2. Energy is conserved (since L does not depend on t)

3. Motion is confined to a plane (angular momentum is conserved)

6.1.4 Qualitative characteristics of orbits

Let us make a few more general observation on the dynamics of a particle subject to a central force. We start by considering a potential that gives an attractive inverse-square force law (such as gravity):

$$\begin{aligned} V(r) &= \frac{-a}{r} \\ \therefore V_{eff}(r) &= \frac{-a}{r} + \frac{l^2}{2\mu r^2} \end{aligned} \tag{6.23}$$

The corresponding $V_{eff}(r)$ is sketched in figure 6.1 for a particular value of the angular momentum and a. Three possible cases for the energy are also shown:

1. If the energy is larger than zero, all possible values of $r > r_{min}$, are allowable. r_{min} corresponds to the case where $V_{eff} = E$. If the particle arrives from infinity, it will drift towards $r = 0$ until it reaches r_{min} where it will experience a **repulsive** force that pushes it back out from the center of mass. This repulsive force is from the conservation of angular momentum. The only case that the particle can fall into the center of mass is if $l = 0$, that is, there is no angular momentum with respect to the center of mass (the impact parameter is zero).

2. If the energy is less than zero, but bigger that the minimum of V_{eff}, the particle is constrained to be between two circles of radius r_1 and r_2. This does not necessarily mean that the orbit is "closed", only that the trajectory is constrained between two circles.

3. If the energy is exactly V_{min} then the particle is constrained to a specific radius, r_c. In this case, the orbit is necessarily closed and circular. The minimum of V_{eff} occurs when:

$$\begin{aligned} \frac{dV_{eff}}{dr} &= 0 \\ \frac{a}{r^2} &= \frac{l^2}{\mu r^3} \\ \frac{a}{r^2} &= \mu r \dot{\phi}^2 \end{aligned} \tag{6.24}$$

which is precisely the requirement that the centripetal force ($\frac{a}{r^2}$) equal mass times centripetal acceleration. This corresponds to a radius r_c and energy $E = V_{min}$ given by:

$$r_c = \frac{l^2}{a\mu}$$
$$V_{min} = V_{eff}(r_c) = -\frac{a}{r_c} + \frac{l^2}{2\mu r_c^2} = -\frac{a^2\mu}{2l^2} \tag{6.25}$$

The value of the angular momentum and the initial value of the radius will determine the ultimate trajectory of the particles. Note that the kinetic energy of the particle (in the radial direction) is always equal to the distance between E and V_{eff}:

$$T_r(r) = \frac{1}{2}\mu\dot{r}^2 = E - V_{eff}(r) \tag{6.26}$$

which is zero for the circular orbit.

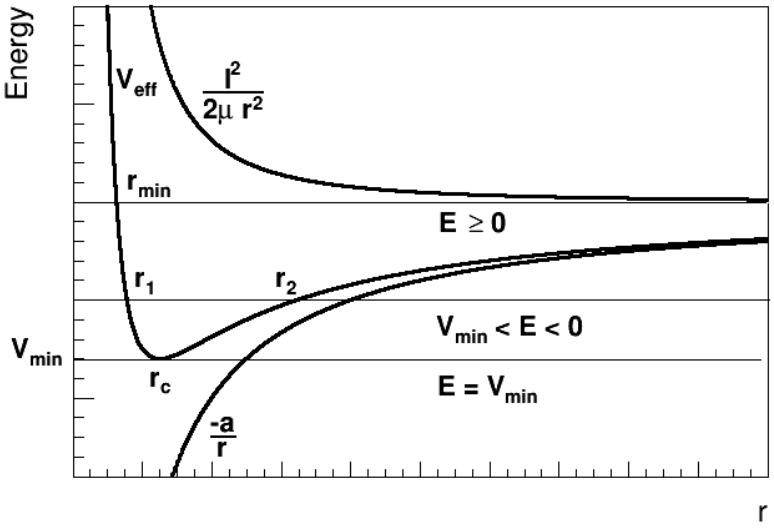

Figure 6.1: Effective potential for the case where $V(r) = -ar^{-1}$. The horizontal lines show three possible values for the energy.

Consider a potential that gives an attractive inverse-fourth force law:

$$V(r) = \frac{-a}{r^3}$$
$$\therefore V_{eff}(r) = \frac{-a}{r^3} + \frac{l^2}{2\mu r^2} \tag{6.27}$$

The effective potential is sketched in Figure 6.2, along with a possible energy E. For this particular case of energy choice, the motion will depend on the initial condition. If the particle started with a radius less than r_1 the motion will be bound and constrained to a radii smaller than r_1. These orbits will pass through the center of mass. In the case where the initial radius was bigger than r_2, the particle will experience a repulsive force and never access radii smaller than r_2. Of course, the repulsive force, is again the apparent centrifugal force.

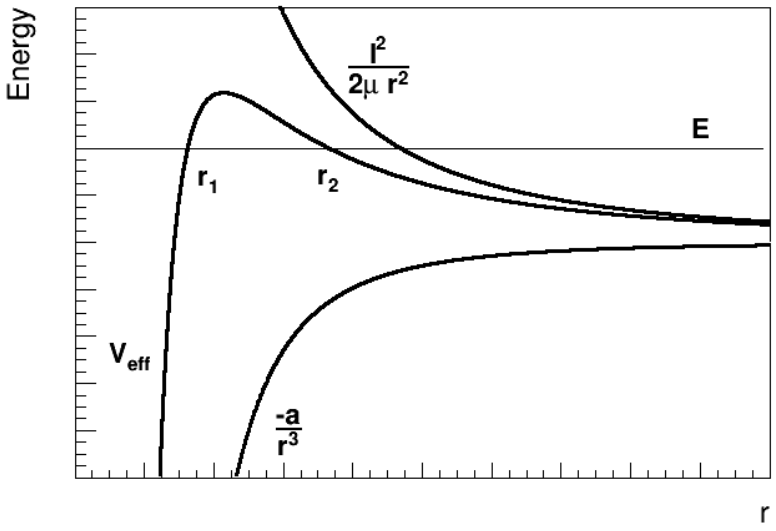

Figure 6.2: Effective potential for the case where $V(r) = -ar^{-3}$.

6.1.5 Determining the equation of the orbit

For certain potentials, it is possible to solve for the equation of the orbit. Starting from the Lagrangian:

$$L = \frac{1}{2}\mu\dot{r}^2 + \frac{l^2}{2\mu r^2} - V(r) \tag{6.28}$$

We know that conservation of energy gives

$$\frac{1}{2}\mu\dot{r}^2 + \frac{l^2}{2\mu r^2} + V(r) = E \tag{6.29}$$

This can be integrated to give r(t). However, we really are interested in obtaining $r(\phi)$. Using the Chain Rule, we can convert the above differential equation from an equation for $r(t)$ to an equation for $r(\phi)$. Consider:

$$\frac{dr}{dt} = \frac{dr}{d\phi}\dot{\phi} \tag{6.30}$$

We can use the conservation of angular momentum:

$$\begin{aligned} \dot{\phi} &= \frac{l}{r^2\mu} \\ \therefore \frac{dr}{dt} &= \frac{dr}{d\phi}\frac{l}{r^2\mu} \\ \therefore \frac{1}{2}\mu\left(\frac{dr}{d\phi}\frac{l}{r^2\mu}\right)^2 + \frac{l^2}{2\mu r^2} + V(r) &= 0 \\ \frac{l^2}{2\mu}\left[\frac{1}{r^4}\left(\frac{dr}{d\phi}\right)^2 + \frac{1}{r^2}\right] + V(r) &= E \end{aligned} \tag{6.31}$$

Let us now introduce a substitution of variables:

$$\begin{aligned} u &\equiv \frac{1}{r} \\ \therefore \frac{du}{d\phi} &= \frac{du}{dr}\frac{dr}{d\phi} = -\frac{1}{r^2}\frac{dr}{d\phi} \\ \therefore \frac{dr}{d\phi} &= -r^2\frac{du}{d\phi} \end{aligned} \tag{6.32}$$

The differential equation then becomes:

$$\frac{l^2}{2\mu}\left[\left(\frac{du}{d\phi}\right)^2 + u^2\right] + V(\frac{1}{u}) = E \tag{6.33}$$

Let us assume that the potential has the following form:

$$V(r) = ar^{-n} \tag{6.34}$$

Careful! This is a different definition than in the previous section where we had forced a to be positive. Here, we allow a to be positive (repulsive force) or negative (attractive force). The differential equation is then:

$$\left[\left(\frac{du}{d\phi}\right)^2 + u^2\right] + \frac{2\mu}{l^2}au^n = \frac{2\mu}{l^2}E \tag{6.35}$$

again, we can write this as an integral:

$$\begin{aligned} \int_{\phi_0}^{\phi} d\phi &= \int_{u_0}^{u} \frac{du}{\sqrt{\frac{2\mu}{l^2}E - \frac{2\mu}{l^2}au^n - u^2}} \\ &= -\int_{r_0}^{r} \frac{dr}{r^2\sqrt{\frac{2\mu}{l^2}E - \frac{2\mu}{l^2}ar^{-n} - \frac{1}{r^2}}} \end{aligned} \tag{6.36}$$

where we have used the fact that $du = -\frac{1}{r^2}dr$. This can be integrated analytically for certain values of n.

6.1.6 The case of gravity ("Kepler's problem")

Let us solve specifically the case of two masses m_1 and m_2 interacting with gravity:

$$\begin{aligned} V(r) = ar^{-n} &= -\frac{Gm_1m_2}{r} \\ a &= -Gm_1m_2 \\ n &= 1 \end{aligned} \tag{6.37}$$

The differential equation for $r(\phi)$ is thus:

$$\left[\left(\frac{du}{d\phi}\right)^2 + u^2\right] - \frac{2\mu}{l^2}au = \frac{2\mu}{l^2}E \tag{6.38}$$

Let us introduce a constant, α:

$$\begin{aligned} \alpha &\equiv \frac{l^2}{a\mu} \\ \therefore \left[\left(\frac{du}{d\phi}\right)^2 + u^2\right] - \frac{2}{\alpha}u &= \frac{2E}{a\alpha} \end{aligned} \tag{6.39}$$

We now add $\frac{1}{\alpha^2}$ to both sides:

$$\left[\left(\frac{du}{d\phi}\right)^2 + u^2\right] - \frac{2}{\alpha}u + \frac{1}{\alpha^2} = \frac{2E}{a\alpha} + \frac{1}{\alpha^2} \tag{6.40}$$

We introduce a new function, f, and a new constant K:

$$\begin{aligned} f(\phi) &\equiv u(\phi) - \frac{1}{\alpha} \\ \left(\frac{df}{d\phi}\right)^2 &= \left(\frac{du}{d\phi}\right)^2 \\ f^2 &= u^2 - 2u\frac{1}{\alpha} + \frac{1}{\alpha^2} \\ K^2 &\equiv \frac{1}{\frac{2E}{a\alpha} + \frac{1}{\alpha^2}} \end{aligned} \tag{6.41}$$

The differential equation becomes:

$$\begin{aligned} K^2\left[\left(\frac{df}{d\phi}\right)^2 + f^2\right] &= 1 \\ \left[\left(\frac{dKf}{d\phi}\right)^2 + (Kf)^2\right] &= 1 \\ \left[\left(\frac{dg}{d\phi}\right)^2 + g^2\right] &= 1 \end{aligned} \tag{6.42}$$

where we have introduced a new function $g(\phi) = Kf(\phi)$. We can identify g with a trigonometric function, let's choose:

$$g = \cos\phi \tag{6.43}$$

and the differential equation is obviously satisfied by the trigonometric identity $\cos^2 + \sin^2 = 1$. The solution is thus:

$$\begin{aligned} f(\phi) &= \frac{1}{K}\cos(\phi) = u(\phi) - \frac{1}{\alpha} \\ u(\phi) &= \frac{1}{K}\cos(\phi) + \frac{1}{\alpha} \\ &= \sqrt{\frac{2E}{a\alpha} + \frac{1}{\alpha^2}}\cos(\phi) + \frac{1}{\alpha} \\ &= \frac{1}{\alpha}\sqrt{\frac{2E\alpha}{a} + 1}\cos(\phi) + \frac{1}{\alpha} \end{aligned} \tag{6.44}$$

Introducing the quantity, e:

$$\begin{aligned} e^2 &\equiv \frac{2E\alpha}{a} + 1 \\ \frac{1}{r} &= \frac{e}{\alpha}\cos(\phi) + \frac{1}{\alpha} \end{aligned} \tag{6.45}$$

Finally, multiplying by $r\alpha$, we obtain the equation for a conical curve of eccentricity e in polar coordinates:

$$r(e\cos\phi + 1) = \alpha \tag{6.46}$$

which is what we expect from Kepler's laws. Note that we only considered the shape of the trajectory, so we ignored the initial condition on r or ϕ when performing the integral. If we need to obtain the equations for r as a function of time, it would be important to include the initial condition of ϕ and replace $\phi \to \phi - \phi_0$.

Recall the properties of the eccentricity, e, for conical curves:

1. $e = 0$ corresponds to a circle centered on the origin
2. $0 \leq e \leq 1$ corresponds to an ellipse with a focus at the origin
3. $e = 1$ corresponds to a parabola
4. $e > 1$ corresponds to a hyperbola

We can relate this to our earlier qualitative comments about the the orbits, based on the energy:

$$\begin{aligned} e^2 &= \frac{2E\alpha}{a} + 1 \\ &= E\frac{2l^2}{a^2\mu} + 1 \\ \therefore E &= \frac{\mu a^2}{2l^2}(e^2 - 1) \end{aligned} \tag{6.47}$$

An energy of zero thus results in a parabolic trajectory ($e = 1$). The circular orbit ($e = 0$) is given by:

$$E = -\frac{\mu a^2}{2l^2} \tag{6.48}$$

which is the result we obtained earlier. If e is bigger than 1, then the energy is positive and the orbit is unbound (a hyperbola). If e is less than 1, then the energy is negative and the orbit is bound (an ellipse).

6.2 Small oscillations and coupled oscillators

6.2.1 Equilibrium

Consider a simple system where all forces are monogenic and there is only one (cartesian) degree of freedom.

$$L = \frac{1}{2}m\dot{q}^2 - V(q) \tag{6.49}$$

The system will be in static equilibrium if the virtual work done by the monogenic forces is zero. This will be true if the generalized force is zero:

$$Q_1 = -\frac{\partial V}{\partial q} = 0 \tag{6.50}$$

Such a condition implies that we have an extremum of V. If we then expand the potential in a Taylor series near a point q_0, where the potential is at the extremum, we have:

$$\begin{aligned} V &= V(q_0) + \frac{\partial V}{\partial q}(q - q_0) + \frac{1}{2}\frac{\partial^2 V}{\partial^2 q}(q - q_0)^2 + \dots \\ &\sim \frac{1}{2}k(q - q_0)^2 \\ k &\equiv \frac{\partial^2 V}{\partial^2 q} \end{aligned} \tag{6.51}$$

which is the potential energy of a simple harmonic oscillator. We can always choose to define V such that it is zero at the point where the forces are zero (alternatively, the first term is a constant and does not affect the motion).

An equilibrium is "stable" if, when the system is displaced from equilibrium, it returns to that point (or oscillates around it) (e.g. a ball at the bottom of a bowl). The equilibrium is called "unstable" if a small perturbation of the system away from equilibrium will lead the system go further from the equilibrium point (e.g. a ball at the top of a hill).

The system will be in static equilibrium at q_0. The nature of the equilibrium (stable or unstable) will depend on the sign of k. If k is positive, then the point of equilibrium is a minimum of V and the particle will be bounded by points where $E = V$. If k is negative, then the particle will see a potential barrier at q_0, which it can only surmount if $E > 0$, otherwise, it will turn around when $E = V$.

6.2.2 Simple harmonic oscillator

The simple harmonic oscillator Lagrangian is given by:

$$L = \frac{1}{2}m\dot{q}^2 - \frac{1}{2}kq^2 \tag{6.52}$$

where the system is in equilibrium at $q = 0$. Since L does not depend explicitly on time, energy is conserved, and we can solve for the integral (assuming that k and E are positive):

$$\begin{aligned} E &= \frac{1}{2}m\dot{q}^2 + \frac{1}{2}kq^2 \\ \dot{q} &= \sqrt{\frac{2E - kq^2}{m}} = \sqrt{\frac{2E}{m}}\sqrt{1 - \frac{k}{2E}q^2} \\ \therefore t &= \sqrt{\frac{m}{2E}}\int \frac{dq}{\sqrt{1 - \frac{k}{2E}q^2}} \\ t &= \sqrt{\frac{m}{2E}}\sqrt{\frac{2E}{k}}\sin^{-1}\left(\sqrt{\frac{k}{2E}}q\right) + \text{const.} \\ &= \sqrt{\frac{m}{k}}\sin^{-1}\left(\sqrt{\frac{k}{2E}}q\right) + \text{const.} \\ \therefore q(t) &= A\sin\left(\sqrt{\frac{k}{m}}t + \phi_0\right) \\ &= A\sin\left(\omega t + \phi_0\right) \end{aligned} \tag{6.53}$$

where we have introduced the frequency ω and the constants A and ϕ_0 are determined from the initial conditions. Note that if k is positive, then E must also be positive. If k is negative, the solution is a hyperbolic sine function.

We could also solve this from the Lagrange equation of motion, by postulating a solution:

$$\begin{aligned} m\ddot{q} &= -kq \\ \therefore q(t) &= Ae^{-i\omega t} \\ \omega &= \sqrt{\frac{k}{m}} \end{aligned} \tag{6.54}$$

where $q(t)$ is the real part of the expression.

6.2.3 Many oscillators

Suppose that we have n degrees of freedom q_i and a Lagrangian of the form:

$$\begin{aligned} L &= T - V \\ T &= \frac{1}{2}\sum_{i,j} T_{ij}\dot{q}_i\dot{q}_j \end{aligned} \tag{6.55}$$

Suppose that the point $(q_{0,1}, \ldots, q_{0,n})$ corresponds to a static equilibrium:

$$\left(\frac{\partial V}{\partial q_i}\right)\bigg|_0 = 0 \tag{6.56}$$

Consider a point, q_i', that is slightly displaced from the equilibrium by a distance, η_i:

$$q_i' = q_{0,i} + \eta_i \tag{6.57}$$

The potential at q_i' can be expanded in a Taylor series near that point:

$$\begin{aligned} V(q_1', \ldots, q_n') &= V(q_{0,1}, \ldots, q_{0,n}) + \sum_i \left(\frac{\partial V}{\partial q_i}\right)\bigg|_0 \eta_i + \frac{1}{2}\sum_{i,j}\left(\frac{\partial V}{\partial q_i}\frac{\partial V}{\partial q_j}\right)\bigg|_0 \eta_i\eta_j + \ldots \\ &\sim \frac{1}{2}\sum_{i,j}\left(\frac{\partial V}{\partial q_i}\frac{\partial V}{\partial q_j}\right)\bigg|_0 \eta_i\eta_j \\ &= \frac{1}{2}\sum_{i,j} V_{ij}\eta_i\eta_j \end{aligned} \tag{6.58}$$

where we have dropped the first two terms, and introduced V_{ij} as the second order derivatives evaluated at the equilibrium. The first term is a constant and does not influence the motion, and the second is zero at the equilibrium. The kinetic energy can also be written in terms of the small displacements from equilibrium, η_i:

$$\begin{aligned} \eta_i &= q_i' - q_{0,i} \\ \therefore \dot{\eta}_i &= \dot{q}_i' \\ \therefore T &= \frac{1}{2}\sum_{i,j} T_{ij}\dot{\eta}_i\dot{\eta}_j \end{aligned} \tag{6.59}$$

and the Lagrangian, near equilibrium, can be written as:

$$L = \frac{1}{2}\sum_{i,j} T_{ij}\dot{\eta}_i\dot{\eta}_j - \frac{1}{2}\sum_{i,j} V_{ij}\eta_i\eta_j \tag{6.60}$$

Taking the Lagrange equation for η_i, we obtain:

$$\sum_j \left(T_{ij}\ddot{\eta}_j + V_{ij}\eta_j\right) = 0 \tag{6.61}$$

where we have used the fact that $V_{ij} = V_{ji}$. Again, we can postulate a solution of the form:

$$\eta_i(t) = C a_i e^{-i\omega t} \tag{6.62}$$

where the a_i correspond to each η_i, C is an overall (complex) constant, and only the real part is used to represent the motion. This solution is called a "normal mode", as all coordinates oscillate with the same frequency. If we substitute this back into the Lagrange equation:

$$\begin{aligned} \sum_j \left(-\omega^2 T_{ij}\eta_j + V_{ij}\eta_j\right) &= 0 \\ \sum_j a_j(V_{ij} - \omega^2 T_{ij}) &= 0 \end{aligned} \tag{6.63}$$

Consider this as a matrix equation:

$$\begin{pmatrix} V_{11} - \omega^2 T_{11} & \cdots & V_{1n} - \omega^2 T_{1n} \\ \vdots & \ddots & \vdots \\ V_{n1} - \omega^2 T_{n1} & \cdots & V_{nn} - \omega^2 T_{nn} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = 0 \tag{6.64}$$

We can re-write it as:

$$\begin{aligned} V\vec{a} - \omega^2 T\vec{a} &= 0 \\ V\vec{a} &= \omega^2 T\vec{a} \\ T^{-1}V\vec{a} &= \omega^2\vec{a} \end{aligned} \tag{6.65}$$

where we have treated T and V as matrices. The last step required T to be invertible. Note that V is a symmetric matrix. You should recognize that this is a characteristic eigenvalue equation ($\vec{a}$ is an eigenvector and ω^2 is the corresponding eigenvalue).

The particular solution that we postulated corresponds to the "normal modes" of the system. These normal modes have characteristic frequencies and characteristic displacement vectors.

EXAMPLE 6-2: Find the linear normal modes of vibration for three equal masse, m connected by two springs with spring constant k

---

Problem 6-1: Two masses connected by a spring
Two masses m_1 and m_2 are connected by a spring with spring constant k.
a)Write the Lagrangian for this system and show that it can be reduced to a single degree of freedom
b)Plot the effective potential for the system, and show that all orbits are bound
c)Determine the shape of the orbit, and plot it.

Problem 6-2: The Coulomb force
Two masses m_1 and m_2 with charges q_1 and q_2 interact through the Coulomb force.
a)Write the Lagrangian for this system and show that it can be reduced to a single degree of freedom
b)Plot the effective potential for the system when the force is attractive and describe the possible orbits
c)Plot the effective potential for the system when the force is repulsive and show that only hyperbolae are possible orbits
d)Plot the effective potential for the system when the force is repulsive but is constrained to act within a radius $r < R$ (for example, the Coulomb force from a nucleus that is screened by electrons at a large distance) and describe the possible orbits.

Problem 6-3: Kepler's problem
a) Show that the analytic treatment presented in this chapter is consistent with Kepler's three laws
b) Show that for elliptic orbits, the major axis only depends on the energy of the system. **c)**Make a plot showing how the effective potential depends on angular momentum (show several curves for different values

of angular momentum and comment).

Problem 6-4: Simple Harmonic Oscillator
a) Discuss the solutions for the simple harmonic oscillator for the cases when E and/or k are negative. If the solution exists, plot the position as a function of time.

Problem 6-5: Coupled oscillators
a) Show that in the case illustrated in example 6-2, the total momentum is conserved. In particular, show that the Lagrangian can be written with one less degree of freedom.

7 The Hamiltonian

In this chapter, we consider a different approach to analytical mechanics, using the Hamiltonian instead of the Lagrangian. In Lagrangian mechanics, we generally obtained n second order differential equations corresponding to the n degrees of freedom in configuration space. In the Hamiltonian formalism, we will show that we can obtain $2n$ first order differential equations for the n degrees of freedom and the n generalized momenta. The generalized momenta are "promoted" to variables that describe the motion, and we speak of describing a system in "phase space" by specifying q_i and p_i, instead of specifying only the q_i in configuration space. The main difference is that the position in phase space completely specifies the past and future motions of the system, while in configuration space, one also needs to specify the velocities.

7.1 The Legendre Transform

The Hamiltonian formalism can be formally derived through the use of the Legendre transforms, which we first introduce here. Consider a function, $f(u_1, \ldots, u_n)$, that depends on n variables, u_i. Now consider new variables, v_i, given by:

$$v_i \equiv \frac{\partial f}{\partial u_i} \tag{7.1}$$

The Legendre transformation takes the function $f(u_i)$ into a new function, $g(v_i)$, that only depends on the v_i. The transformation is given by:

$$g(v_i) = \sum_i u_i v_i - f \tag{7.2}$$

It is not immediately apparent that g does not depend on the u_i, but we can verify this by considering the variation of g:

$$\begin{aligned}
\delta g &= \delta \left(\sum_i u_i v_i - f \right) \\
&= \sum_i (u_i \delta v_i + v_i \delta u_i) - \sum_i \frac{\partial f}{\partial u_i} \delta u_i \\
&= \sum_i u_i \delta v_i + \sum_i \left(v_i - \frac{\partial f}{\partial u_i} \right) \delta u_i \\
&= \sum_i u_i \delta v_i
\end{aligned} \tag{7.3}$$

where in the last line, we used the definition of v_i, to set the last term equal to zero. Thus, the variation of g only depends on the variations of the v_i and not of the u_i. We can thus write $g = g(v_i)$ and the variation of g as:

$$\delta g = \sum_i \frac{\partial g}{\partial v_i} \delta v_i \tag{7.4}$$

and make the following identification:

$$u_i = \frac{\partial g}{\partial v_i} \tag{7.5}$$

The Legendre Transformation thus has a nice set of symmetries:

$$\begin{aligned} v_i &= \frac{\partial f}{\partial u_i} \\ u_i &= \frac{\partial g}{\partial v_i} \\ g &= \sum_i u_i v_i - f \\ f &= \sum_i u_i v_i - g \end{aligned} \tag{7.6}$$

Now consider the case when f depends on two sets of variables, u_i, and w_i, $f = f(u_i, w_i)$, and we again define the v_i as follows:

$$v_i \equiv \frac{\partial f(u_i, w_i)}{\partial u_i} \tag{7.7}$$

We thus call the u_i the "active" variables, and the w_i, the "passive" variables. Again, we define g:

$$g \equiv \sum_i u_i v_i - f(u_i, w_i) \tag{7.8}$$

The variation of g is given by:

$$\begin{aligned} \delta g &= \delta \left(\sum_i u_i v_i - f(u_i, w_i) \right) \\ &= \sum_i (u_i \delta v_i + v_i \delta u_i) - \sum_i \left(\frac{\partial f}{\partial u_i} \delta u_i + \frac{\partial f}{\partial w_i} \delta w_i \right) \\ &= \sum_i \left(u_i \delta v_i - \frac{\partial f}{\partial w_i} \delta w_i \right) + \sum_i \left(v_i - \frac{\partial f}{\partial u_i} \right) \delta u_i \\ &= \sum_i \left(u_i \delta v_i - \frac{\partial f}{\partial w_i} \delta w_i \right) \end{aligned} \tag{7.9}$$

which again is independent of the variation on the u_i. Since $g = g(v_i, w_i)$, we can also write the variation as:

$$\delta g_i = \sum_i \left(\frac{\partial g}{\partial v_i} \delta v_i + \frac{\partial g}{\partial w_i} \delta w_i \right) \tag{7.10}$$

and immediately identify:

$$\begin{aligned} u_i &= \frac{\partial g}{\partial v_i} \\ \frac{\partial g}{\partial w_i} &= -\frac{\partial f}{\partial w_i} \end{aligned} \tag{7.11}$$

Note that we used the case where there are the same number of w_i as there are u_i and v_i; it is easily seen that the results do not depend on this.

7.2 Legendre Transform of the Lagrangian

Consider now the Lagrangian, $L(q_i, \dot{q}_i, t)$, where we will consider the q_i as the passive variables, and the $\dot{q}_i$ as the active variables. We introduce a new set of variables, p_i, given by:

$$p_i \equiv \frac{\partial L}{\partial \dot{q}_i} \tag{7.12}$$

and we introduce a new function, $H(q_i, p_i, t)$, called the Hamiltonian, which is the Legendre transform of the Lagrangian:

$$H(q_i, p_i, t) \equiv \sum_i p_i \dot{q}_i - L \tag{7.13}$$

Again, consider the variation of the Hamiltonian:

$$\begin{aligned}
\delta H &= \delta \left(\sum_i p_i \dot{q}_i - L \right) \\
&= \sum_i \left(p_i \delta \dot{q}_i + \dot{q}_i \delta p_i \right) - \sum_i \left(\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right) - \frac{\partial L}{\partial t} \delta t \\
&= \sum_i \left(\dot{q}_i \delta p_i - \frac{\partial L}{\partial q_i} \delta q_i \right) - \frac{\partial L}{\partial t} \delta t
\end{aligned} \tag{7.14}$$

Again, we can identify this with the variation of $H(q_i, p_i, t)$:

$$\delta H = \sum_i \left(\frac{\partial H}{\partial p_i} \delta p_i + \frac{\partial H}{\partial q_i} \delta q_i \right) + \frac{\partial H}{\partial t} \delta t \tag{7.15}$$

to obtain:

$$\begin{aligned}
\dot{q}_i &= \frac{\partial H}{\partial p_i} \\
\frac{\partial H}{\partial q_i} &= -\frac{\partial L}{\partial q_i} \\
\frac{\partial H}{\partial t} &= -\frac{\partial L}{\partial t}
\end{aligned} \tag{7.16}$$

Note that the Hamiltonian can and must always be written explicitly only in terms of the generalized momenta and the coordinates (that is, the velocities should always be eliminated in favour of the momenta).

7.3 The Canonical equations

Again, consider the relations for transforming back and forth between the Lagrangian and the Hamiltonian:

$$\begin{aligned}
\dot{q}_i &= \frac{\partial H}{\partial p_i} \\
p_i &= \frac{\partial L}{\partial \dot{q}_i}
\end{aligned} \tag{7.17}$$

The second equation can be re-arranged using the Euler-Lagrange equation:

$$\begin{aligned}
\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} &= \frac{\partial L}{\partial q_i} \\
\therefore \dot{p} = \frac{\partial L}{\partial q_i} &= -\frac{\partial H}{\partial q_i}
\end{aligned} \tag{7.18}$$

We can write the transformation equations as:

$$\begin{aligned}
\dot{q}_i &= \frac{\partial H}{\partial p_i} \\
\dot{p}_i &= -\frac{\partial H}{\partial q_i}
\end{aligned} \tag{7.19}$$

These are called the "canonical equations of Hamilton". They are $2n$ first order differential equations that specify the location, p_i, q_i, of the system in phase space, and have the property that the total time derivatives are isolated on one side of the equation. Note that if a coordinate is cyclic in the Hamiltonian, conservation of the associated generalized momentum follows immediately. In order to solve these equations in configuration space, one can substitute the first equation into the second one and re-obtain the second-order equation for the coordinate as a function of time that one obtains from the Euler-Lagrange equations.

Consider now the total time derivative of the Hamiltonian, $H(p_i, q_i, t)$:

$$\begin{aligned}\frac{dH}{dt} &= \sum_i \left(\frac{\partial H}{\partial p_i}\dot{p}_i + \frac{\partial H}{\partial q_i}\dot{q}_i\right) + \frac{\partial H}{\partial t} \\ &= \sum_i \left(\dot{q}_i\dot{p}_i - \dot{p}_i\dot{q}_i\right) + \frac{\partial H}{\partial t} \\ &= \frac{\partial H}{\partial t}\left(= -\frac{\partial L}{\partial t}\right)\end{aligned} \tag{7.20}$$

where we used the canonical equations. We see that if the Lagrangian (or Hamiltonian) does not explicitly depend on time, then the total time derivative of H is zero (i.e. H is a constant). Often, the Hamiltonian is equal to the total energy, in particular, if it does not depend on time explicitly and if the potential does not depend on velocity. This is the result that we obtained for the Jacobian integral. In the case that the Hamiltonian is the total energy, it can be written as:

$$H = T + V \tag{7.21}$$

One should however be careful and keep in mind that this form is not the general definition of the Hamiltonian, which is given by the Legendre transformation.

7.4 Canonical equations from Hamilton's principle

Recall that we obtained the equations of motion by requiring that the action, S, is stationary under variations of the q_i:

$$S = \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t)dt \tag{7.22}$$

The Euler-Lagrange equations of motion were obtained by requiring that the variations of the q_i were zero at the end points. We can write the variation of S in terms of the Hamiltonian:

$$\delta S = \delta \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t)dt = \delta \int_{t_1}^{t_2} \left[\sum_i p_i\dot{q}_i - H(q_i, p_i, t)\right] dt \tag{7.23}$$

In the Lagrangian formalism, we found that variations of the action with respect to the q_i led to the equations of motion. The $\dot{q}_i$ did not vary independently from the q_i, since:

$$\delta\dot{q}_i = \delta\frac{dq_i}{dt} = \frac{d}{dt}\delta q_i \tag{7.24}$$

In the Hamiltonian formalism, we must treat the q_i and p_i on equal footing, that is, we must allow them to be varied independently. We know from the properties of the Legendre transformation that variations of the p_i do not affect the Lagrangian, and hence the variation of the action will not be affected by independently varying the p_i:

$$\begin{aligned}&p_i \to p_i + \delta p_i \\ &q_i \to q_i + \delta q_i \\ &\delta S = \int_{t_1}^{t_2} \left[\sum_i \left(p_i\delta\dot{q}_i + \dot{q}_i\delta p_i - \frac{\partial H}{\partial q_i}\delta q_i - \frac{\partial H}{\partial p_i}\delta p_i\right)\right] dt\end{aligned} \tag{7.25}$$

We can write:

$$p_i \delta \dot{q}_i = p_i \frac{d}{dt} \delta q_i = \frac{d}{dt}(p_i \delta q_i) - \dot{p}_i \delta q_i \tag{7.26}$$

The first term, being a total time derivative, will not contribute to the variation of the action (since it just adds a constant), so it can be dropped:

$$\begin{aligned} \delta S &= \int_{t_1}^{t_2} \left[\sum_i \left(-\dot{p}_i \delta q_i + \dot{q}_i \delta p_i - \frac{\partial H}{\partial q_i} \delta q_i - \frac{\partial H}{\partial p_i} \delta p_i \right) \right] dt \\ &= \int_{t_1}^{t_2} \left[\sum_i \left(-\dot{p}_i - \frac{\partial H}{\partial q_i} \right) \delta q_i + \left(\dot{q}_i - \frac{\partial H}{\partial p_i} \right) \delta p_i \right] dt \end{aligned} \tag{7.27}$$

For the variation of S to be zero when q_i and p_i are varied independently, then the terms in front of the δq_i and δp_i must always be zero, which is precisely the canonical equations.

7.5 Symplectic notation

Symplectic notation is a way to handle the Hamiltonian formalism using matrices. It is often implemented in computerized algorithms, for example for solving for the motion of astrophysical bodies (think of calculating the trajectory of a probe on its way to Mars). Such computerized algorithms are called "symplectic integrators".

One starts by defining a vector, $\vec{\eta}$, of dimension $2n$ where the first n coordinates are the generalized coordinates, q_i, and the next n coordinates are the generalized momenta, p_i.

$$\vec{\eta} \equiv \begin{pmatrix} q_1 \\ \vdots \\ q_n \\ p_1 \\ \vdots \\ p_n \end{pmatrix} \tag{7.28}$$

We define a second vector, $\frac{\partial \vec{H}}{\partial \eta}$:

$$\frac{\partial \vec{H}}{\partial \eta} \equiv \begin{pmatrix} \frac{\partial H}{\partial q_1} \\ \vdots \\ \frac{\partial H}{\partial q_n} \\ \frac{\partial H}{\partial p_1} \\ \vdots \\ \frac{\partial H}{\partial p_n} \end{pmatrix} \tag{7.29}$$

and a $2n \times 2n$ square matrix J:

$$J \equiv \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{7.30}$$

where I is the $n \times n$ identity matrix. Hamilton's equations in symplectic notation are thus written as:

$$\frac{d\vec{\eta}}{dt} = J \frac{\partial \vec{H}}{\partial \eta} \tag{7.31}$$

7.6 Phase space, the phase space fluid and Liouville's Theorem

Recall that in the Lagrangian formalism, one can describe a system by specifying the value of its generalized coordinates in configuration space. In order to know the future (or past) development of the system, it is also necessary to specify the velocities. For example, a canon ball's trajectory through (say, Cartesian) configuration space, depends on its velocity. In general, for a given starting position in configuration space, different initial velocities can lead to intersecting trajectories in configuration space (see Figure 7.1).

In the Hamiltonian formalism, the system is described by the generalized coordinates **and** by the generalized momenta. This can be viewed as a set of coordinates in "phase space". The trajectory of the system through phase space is completely specified by the equations of motion. If the Hamiltonian does not change with time, the paths of different systems through phase space cannot intersect (this would mean that two systems with different positions (in configuration space) and momentum, following the same equations of motion, could end up at the same position and momentum). One can make the analogy with fluid dynamics, where the trajectories through phase space, for a system following a given Hamiltonian, are similar to the flow lines for a fluid (see Figure 7.1). Each trajectory corresponds to a particular value of the Hamiltonian. This fluid is called the "phase space fluid".

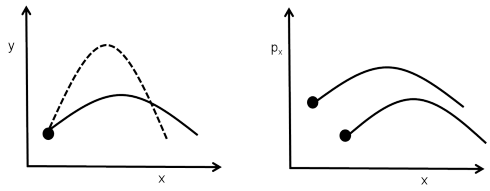

Figure 7.1: Paths of a systems through configuration space (left) and phase space (right). In configuration space, the trajectories depend on the initial velocities and the equations of motion, so they can intersect. In phase space, the trajectories are completely determined by the equations of motion and the starting position in phase space;the trajectories for different starting points cannot intersect.

EXAMPLE 7-1: Use the Hamiltonian formalism to describe the simple harmonic oscillator of mass m and spring constant k, and describe the motion in phase space.

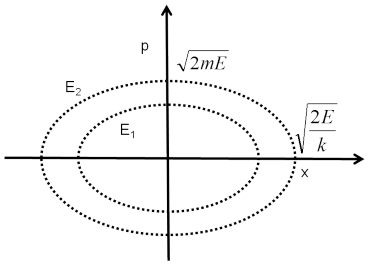

Figure 7.2: Motion of the simple harmonic oscillator in phase space for two different values of energy.

Liouville's theorem (although it was formulate by Gibbs and does not have much to do with Liouville) states that the phase space fluid acts like an incompressible fluid. That is, if you select a closed volume and let it evolve as the fluid moves in time, that total volume does not change with time. We know from fluid dynamics, that the divergence of the velocity field must be zero for an incompressible fluid. Indeed, consider the continuity equation for a fluid of density ρ and velocity field $\vec{v}$:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0 \tag{7.33}$$

If the density is constant (the fluid incompressible), then the divergence of the velocity vector is zero.

Treating the generalized momenta and the generalized coordinates as "regular" coordinates of a particle in configuration space, the velocity of that particle is given by a vector of dimension $2n$:

$$\vec{v} = \begin{pmatrix} \dot{q}_1 \\ \vdots \\ \dot{q}_n \\ \dot{p}_1 \\ \vdots \\ \dot{p}_n \end{pmatrix} = \begin{pmatrix} \frac{\partial H}{\partial p_1} \\ \vdots \\ \frac{\partial H}{\partial p_n} \\ -\frac{\partial H}{\partial q_1} \\ \vdots \\ -\frac{\partial H}{\partial q_n} \end{pmatrix} \tag{7.34}$$

where we have also substituted Hamilton's canonical equations in the second equal sign.

If the divergence of the corresponding velocity field is zero:

$$\begin{aligned} \nabla \cdot \vec{v} &= 0 \\ &= \sum_i^n \left(\frac{\partial v_i}{\partial q_i} + \frac{\partial v_{n+i-1}}{\partial p_i} \right) \\ &= \sum_i^n \left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right) \\ &= \sum_i^n \left(\frac{\partial}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial}{\partial p_i}\frac{\partial H}{\partial q_i} \right) = 0 \end{aligned} \tag{7.35}$$

which is equal to zero, since the partial derivatives commute. The velocity field in field space does indeed behave like an incompressible fluid.

7.7 Poisson Brackets

The Poisson Bracket between two functions of the canonical variable, $U(q_i, p_i, t)$, and $V(q_i, p_i, t)$ is defined as:

$$\{U, V\} \equiv \sum_i^n \left(\frac{\partial U}{\partial q_i} \frac{\partial V}{\partial p_i} - \frac{\partial U}{\partial p_i} \frac{\partial V}{\partial q_i} \right) \tag{7.36}$$

The Poisson Bracket has several properties that are easily verified (in the following, capital letters denote functions of q_i and p_i, whereas k is a constant):

$$\begin{aligned}
\{U, V\} &= -\{V, U\} \\
\{U, U\} &= 0 \\
\{kU, V\} &= k\{U, V\} \\
\{A + B, V\} &= \{A, V\} + \{B, V\} \\
\{AB, V\} &= A\{B, V\} + B\{A, V\} \\
\{q_i, q_j\} = \{p_i, p_j\} &= 0 \\
\{q_i, p_j\} &= \delta_{ij} \\
\{q_i^n, p_j\} &= n q_i^{n-1} \delta_{ij} \\
\{U, p_i\} &= \frac{\partial U}{\partial q_i} \\
\{U, q_i\} &= -\frac{\partial U}{\partial p_i} \\
\{U, \{V, W\}\} + \{V, \{W, U\}\} + \{W, \{U, V\}\} &= 0
\end{aligned} \tag{7.37}$$

This last relation is called "Jacobi's identity". δ_{ij} is the Kronecker delta, and is equal to zero unless $i = j$. It seems a little strange to introduce the Poisson Brackets as a mathematical artefact, but they are related to the commutators that appear in Quantum Mechanics, so it is worthwhile to explore them further. Consider for example the Poisson Bracket of a canonical variable with the Hamiltonian:

$$\begin{aligned}
\{q_j, H\} &= \sum_i^n \left(\frac{\partial q_j}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial q_j}{\partial p_i} \frac{\partial H}{\partial q_i} \right) \\
&= \frac{\partial H}{\partial p_j} \\
\{p_j, H\} &= \sum_i^n \left(\frac{\partial p_j}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial p_j}{\partial p_i} \frac{\partial H}{\partial q_i} \right) \\
&= -\frac{\partial H}{\partial q_j}
\end{aligned} \tag{7.38}$$

where the terms $\frac{\partial q_j}{\partial q_i}$ and $\frac{\partial p_j}{\partial p_i}$ are zero unless $i = j$, and the terms $\frac{\partial q_j}{\partial p_i}$ and $\frac{\partial p_j}{\partial q_i}$ are zero. Identifying this with Hamilton's canonical equations, we can re-write the canonical equations as:

$$\begin{aligned}
\dot{q}_i &= \{q_i, H\} \\
\dot{p}_i &= \{p_i, H\}
\end{aligned} \tag{7.39}$$

More generally, consider the total time derivative of a function, $F(q_i, p_i, t)$, of the canonical variables:

$$\begin{aligned}\frac{dF}{dt} &= \sum_i \left(\frac{\partial F}{\partial q_i}\dot{q}_i + \frac{\partial F}{\partial p_i}\dot{p}_i\right) + \frac{\partial F}{\partial t} \\ &= \sum_i \left(\frac{\partial F}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial H}{\partial q_i}\right) + \frac{\partial F}{\partial t} \\ &= \{F, H\} + \frac{\partial F}{\partial t}\end{aligned} \tag{7.40}$$

We can see that the time derivative of a quantity is given by its Poisson Bracket with the Hamiltonian. If F does not depend explicitly on time, it is a constant of motion if its Poisson Bracket with the Hamiltonian is zero:

$$\{F(q_i, p_i), H\} = 0 \rightarrow F = \text{const.} \tag{7.41}$$

7.7.1 Poisson Brackets and symmetries

Recall that we showed in Chapter 5 that there exists a conserved quantity, Q, for each axis of rotation about which the Lagrangian was invariant. For an infinitesimal rotation of angle $\delta\epsilon$ about the z-axis:

$$\begin{aligned}x' &= x + f_x\delta\epsilon = x - y\delta\epsilon \\ y &= y + f_y\delta\epsilon = y + x\delta\epsilon \\ z &= z + f_z\delta\epsilon = z \\ Q &= \sum_i f_i p_i = (xp_y - yp_x) = L_z\end{aligned} \tag{7.42}$$

and the conserved quantity was the z-component of angular momentum, L_z.

Now consider the Poisson Bracket:

$$\begin{aligned}\{x, L_z\} &= \{x, L_z\} \\ &= \{x, xp_y - yp_x\} \\ &= \{x, xp_y\} - \{x, yp_x\} \\ &= x\{x, p_y\} + p_y\{x, x\} - y\{x, p_x\} - p_x\{x, y\} \\ &= -y\end{aligned} \tag{7.43}$$

where we have made use of the properties from equations 7.37. One can easily show that:

$$\begin{aligned}\{x, L_z\} &= -y = f_x \\ \{y, L_z\} &= x = f_y \\ \{z, L_z\} &= 0 = f_z\end{aligned} \tag{7.44}$$

Thus, the Poisson Bracket of a coordinate with a component of angular momentum gives the coefficient (f_i) corresponding to the transformation of that coordinate with respect to infinitesimal rotations.

Now consider the Poisson Bracket with the components of momentum:

$$\begin{aligned}\{p_x, L_z\} &= -p_y = f_x \\ \{p_y, L_z\} &= p_x = f_y \\ \{p_z, L_z\} &= 0 = f_z\end{aligned} \tag{7.45}$$

which are easily demonstrated. The momentum vector transforms the same way as coordinates under a rotation, so the Poisson Brackets also correspond to the correct coefficients for rotations of the momentum vector. We find that the f_i corresponding to how a particular quantity is rotated about the z-axis are given by the Poisson Bracket of the quantity with the z-component of angular momentum. We say that angular momentum is the "generator" of rotations.

EXAMPLE 7-2: Show that a charged sphere rotating about some arbitrary axis precesses about an axis parallel to a uniform magnetic field. Assume the magnetic field is in the z-direction

---

Let's examine the relation between Poisson Brackets, symmetries and conserved quantities further. Recall that we showed that if the Lagrangian is invariant under translation in a direction, then momentum in that direction is conserved. The infinitesimal translation about the x-axis transformation equations and the associated conserved quantity are given by:

$$\begin{aligned} x' &= x + f_x\delta\epsilon = x + \delta\epsilon \\ y' &= y + f_y\delta\epsilon = y \\ z' &= z + f_z\delta\epsilon = z \\ Q &= p_x \end{aligned} \tag{7.49}$$

And we find that the Poisson Brackets are evaluated trivially:

$$\begin{aligned} \{x, p_x\} &= 1 = f_x \\ \{y, p_x\} &= 0 = f_y \\ \{z, p_x\} &= 0 = f_z \end{aligned} \tag{7.50}$$

More generally, given a function, F, we have:

$$\{F, p_x\} = \frac{\partial F}{\partial x} \tag{7.51}$$

Under an infinitesimal translation a distance $\delta x = \delta\epsilon$ in the x-direction, the variation of F is:

$$\begin{aligned} F' &= F + \frac{\partial F}{\partial x}\delta\epsilon = F + \{F, p_x\}\delta\epsilon \\ \delta F &= \{F, p_x\}\delta\epsilon \end{aligned} \tag{7.52}$$

Thus, taking the Poisson Bracket of a conserved quantity and a coordinate, tells us how that coordinate transforms under the symmetry that corresponds to that conserved quantity.

Recall that if the Hamiltonian does not depend on time, then energy is conserved. Energy in this case is the Hamiltonian (so the statement is a little redundant). We can thus find how a quantity transforms under a translation in time by taking the Poisson Bracket with the Hamiltonian. Under a time translation, $\delta t = \delta\epsilon$:

$$F' = F + \{F, H\}\delta\epsilon \tag{7.53}$$

Now, consider a general function, $G(q_i, p_i)$ of the $2n$ coordinates. Let's define a transformation of the coordinates given by:

$$\begin{aligned} q_i' &= q_i + \{q_i, G\}\delta\epsilon = q_i + \frac{\partial G}{\partial p_i}\delta\epsilon \\ p_i' &= p_i + \{p_i, G\}\delta\epsilon = p_i - \frac{\partial G}{\partial q_i}\delta\epsilon \end{aligned} \tag{7.54}$$

We will call G the "generator" for this transformation. The Hamiltonian will transform just as any other function of q_i and p_i. If the Hamiltonian is invariant under the transformation generated by G:

$$\begin{aligned} H' &= H + \{H, G\}\delta\epsilon \\ &= H \\ \therefore \{H, G\} &= 0 \end{aligned} \tag{7.55}$$

That is, if the Poisson Bracket $\{H, G\} = 0$, then the Hamiltonian in invariant under the coordinate transformation generated by G. But this also means that $\{G, H\} = 0$, hence that G does not change with time (or is a constant). Here, the Poisson Brackets give us this insight into how a conserved quantity is related to the coordinates transformations under which the Hamiltonian is invariant.

7.8 Problems

Problem 7-1: Simple pendulum
The pendulum in Figure 7.3 is composed of a mass m attached to a mass-less rigid rod of the length l. The pendulum can swing in the xy-plane.

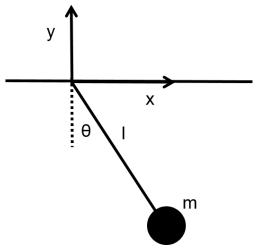

Figure 7.3: The mass m is attached by mass-less rigid rod of length l and free to swing in the xy-plane under the action of gravity. (Problem 7-1)

a)Choose suitable generalized coordinates and write the Hamiltonian for the system in terms of the generalized coordinates and their conjugate momenta
b)Use the canonical equations to obtain the equations of motion for the generalized coordinates and their conjugate momenta
c)Repeat parts a) and b) to obtain the equations of motion for the case where the rod has a mass M

Problem 7-2: Moving pendulum
The pendulum in Figure 7.4 is composed of a mass m attached to a mass-less rigid rod of the length l. The pendulum can swing in the xy-plane. The pivot point of the bar moves downwards at a fixed, known, speed v.

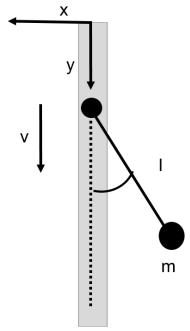

Figure 7.4: The mass m is attached by mass-less rigid rod of length l and free to swing in the xy-plane under the action of gravity. The pivot point moves with a fixed, known velocity, v, and was at the origin at time $t = 0$. (Problem 7-2)

a)Choose suitable generalized coordinates and write the Hamiltonian for the system in terms of the generalized coordinates and their conjugate momenta
b)Use the canonical equations to obtain the equations of motion for the generalized coordinates and their conjugate momenta

Problem 7-3: Two masses and two springs
Figure 7.5 shows two masses, m_1 and m_2, each connected to two springs with spring constants k_1 and k_2. Mass m_1 is constrained to slide without friction along the x-axis, whereas mass m_2 is constrained to move in the vertical direction, constrained by a massless frictionless vertical rod that is attached to m_1. Both springs have a resting length of l.

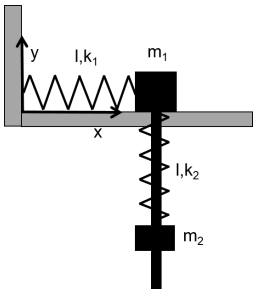

Figure 7.5: Two masses and two springs, problem 7-3

a)Choose suitable generalized coordinates and write the Hamiltonian for the system in terms of the generalized coordinates and their conjugate momenta
b)Use the canonical equations to obtain the equations of motion for the generalized coordinates and their conjugate momenta

Problem 7-4: Pendulum with a spring
Figure 7.6 shows a bead of mass, m, that can slide freely along a long mass-less rail which has one end fixed at the origin, forming a pendulum. The mass is connected to the pivot point at the origin by a mass-less spring of resting length, l, and spring constant k. The motion is constrained to be in the vertical plane (gravity pointing downwards in the figure).

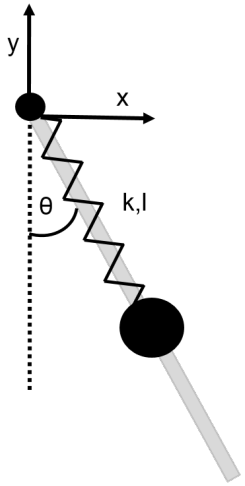

Figure 7.6: Pendulum with a spring, problem 7-4

a)Choose suitable generalized coordinates and write the Hamiltonian for the system in terms of the generalized coordinates and their conjugate momenta
b)Use the canonical equations to obtain the equations of motion for the generalized coordinates and their

conjugate momenta

Problem 7-5: Phase space trajectory of a vertically thrown ball
a) Show that in phase space, the trajectory of a ball thrown vertically is a parabola. Assume that there is only one degree of freedom (in the vertical direction) and that gravity acts downwards with acceleration g.
b) Sketch the velocity lines of the corresponding phase space fluid (i.e. sketch the trajectory in phase space for different initial positions in phase space).

Problem 7-6: Harmonic oscillator trajectory in phase space
In phase space, the one-dimensional simple harmonic oscillator trajectory is an ellipse. The size of the ellipse depends on the initial conditions, or alternatively on the energy of the system. Show that that the time for a system to go around the ellipse in phase space is independent of energy.

Problem 7-7: Poisson Bracket Properties
Prove the following relations from equation 7.37:

a) $\{AB, V\} = A\{B, V\} + B\{A, V\}$
b) $\{q_i, q_j\} = \{p_i, p_j\} = 0$
c) $\{q_i, p_j\} = \delta_{ij}$
d) $\{q_i^n, p_j\} = nq_i^{n-1}\delta_{ij}$
e) $\{U, p_i\} = \dfrac{\partial U}{\partial q_i}$
f) $\{U, q_i\} = -\dfrac{\partial U}{\partial p_i}$
g) $\{U, \{V, W\}\} + \{V, \{W, U\}\} + \{W, \{U, V\}\} = 0$

Problem 7-8: Angular momentum and rotation
a) Show that the Poisson Brackets between the cartesian coordinates and the components of angular momentum give the coefficients corresponding to the infinitesimal rotation about that axis of angular momentum.
b) Show that the Poisson Brackets of angular momentum components satisfy:

$$\{L_i, L_j\} = \epsilon_{ijk} L_k$$

where ϵ_{ijk} is the Levi-Cevita symbol ($+1$ for even permutations of the indices, -1 for odd permutations, and zero if two indices or more are equal).

Problem 7-9: Transformation of the Hamiltonian
Show that under the transformation:

$$q_i' = q_i + \{q_i, G\}\delta\epsilon$$
$$p_i' = p_i + \{p_i, G\}\delta\epsilon$$

The Hamiltonian transforms to:

$$H' = H + \{H, G\}\delta\epsilon$$

Problem 7-10: Hamiltonian for a charged particle
Given the Lagrangian for a charged particle:

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - e\phi + e\vec{A} \cdot \vec{v} \tag{7.56}$$

find equations for the generalized momenta and write the Hamiltonian for the system. Show that the canonical equations for Hamilton give the expected result.

Problem 7-11: Poisson's theorem
Show that if $F(q_i, p_i, t)$ and $G(q_i, p_i, t)$ are two constants of motion, then the Poisson Bracket:

$$\{F, G\}$$

is also a constant of motion.

Problem 7-12: Poisson brackets and conserved quantities
Use Poisson Brackets to show that, given the following Hamiltonian:

$$H = q_1 p_1 - q_2 p_2 - a q_1^2 + b q_2^2$$

the product of $q_1 q_2$ is conserved.

8 Canonical Transformations

We found that certain choices of coordinates give rise to cyclic coordinates leading to their corresponding conjugate momenta being constants. We will see that it is possible to find a set of coordinate transformations that lead to all coordinates being cyclic.

In the Lagrangian formalism, it is difficult to determine these optimal transformations; this is because both the coordinates and their time derivatives appear in the Lagrangian. In the Hamiltonian formalism, we can treat the coordinates and their conjugate momenta as independent, and we thus have more liberty to seek transformations that make the coordinates cyclic. In general, we call transformations that preserve Hamilton's canonical equations, "canonical transformations". The transformed coordinates are denoted by Q_i and P_i, and the transformed Hamiltonian by K. The conditions satisfied by a canonical transformation are thus:

$$\begin{aligned} H(q_i, p_i, t) &\to K(Q_i, P_i, t) \\ \dot{Q}_i &= \frac{\partial K}{\partial P_i} \\ \dot{P}_i &= -\frac{\partial K}{\partial Q_i} \end{aligned} \tag{8.1}$$

8.1 Types of transformations

8.1.1 Point transformations

The simplest type of transformation, "point transformations" simply change the q_i to a new set of coordinates, Q_i. We know from our liberty in choosing a set of generalized coordinates that the Lagrangian is invariant under point transformation. Naturally, Hamilton's equations are also preserved, and point transformations of the type:

$$q_i = q_i(Q_i, t) \tag{8.2}$$

are canonical.

8.1.2 General transformations

Consider the general invertible canonical transformation of the form:

$$\begin{aligned} q_i &= q_i(Q_i, P_i, t) \\ p_i &= p_i(Q_i, P_i, t) \end{aligned} \tag{8.3}$$

The equations of motion will be preserved if the integrand of the action (i.e. the Lagrangian) is modified at most by a total time differential of some function, F:

$$\sum_i \dot{q}_i p_i - H = \sum_i \dot{Q}_i P_i - K + \frac{d}{dt} F(q_i, Q_i, t) \tag{8.4}$$

where we have, arbitrarily, chosen that F depends on the pair of variables q_i, Q_i. We could have just as well chosen any pair of variables that mixes the new and old coordinates, $F(p_i, P_i)$, $F(p_i, Q_i)$, or $F(q_i, P_i)$. We assume (impose) that the transformation is canonical, so that:

$$\begin{aligned}\dot{Q}_i &= \frac{\partial K}{\partial P_i} \\ \dot{P}_i &= -\frac{\partial K}{\partial Q_i}\end{aligned} \tag{8.5}$$

We will require that Equation 8.4 is true so that we can make the transformation canonical. Consider the time derivative of F:

$$\frac{dF}{dt} = \sum_i \left(\frac{\partial F}{\partial q_i}\dot{q}_i + \frac{\partial F}{\partial Q_i}\dot{Q}_i \right) + \frac{\partial F}{\partial t} \tag{8.6}$$

which we can substitute back into 8.4:

$$\begin{aligned}\sum_i \dot{q}_i p_i - H &= \sum_i \dot{Q}_i P_i - K + \sum_i \left(\frac{\partial F}{\partial q_i}\dot{q}_i + \frac{\partial F}{\partial Q_i}\dot{Q}_i \right) + \frac{\partial F}{\partial t} \\ \sum_i \left(p_i - \frac{\partial F}{\partial q_i} \right) \dot{q}_i - H &= \sum_i \left(P_i + \frac{\partial F}{\partial Q_i} \right) \dot{Q}_i - K + \frac{\partial F}{\partial t}\end{aligned} \tag{8.7}$$

We can guarantee the validity of this equation by setting the coefficients of $\dot{q}_i$ and $\dot{Q}_i$ to zero:

$$\begin{aligned}p_i &= \frac{\partial}{\partial q_i} F(q_i, Q_i, t) \\ P_i &= -\frac{\partial}{\partial Q_i} F(q_i, Q_i, t)\end{aligned} \tag{8.8}$$

and by requiring that the new Hamiltonian is then given by:

$$K(Q_i, P_i, t) = H(p_i, q_i, t) + \frac{\partial F}{\partial t} \tag{8.9}$$

We call F the "generator" of the canonical transformation, since it tells how to define the new Hamiltonian and the new coordinates. Given F and the above equations, one can always invert the transformation equations to get:

$$\begin{aligned}q_i &= q_i(Q_i, P_i, t) \\ p_i &= p_i(Q_i, P_i, t) \\ Q_i &= Q_i(q_i, p_i, t) \\ P_i &= P_i(q_i, p_i, t)\end{aligned} \tag{8.10}$$

and Hamilton's canonical equations for $K(Q_i, P_i, t)$.

We have four possible "types" of canonical transformations depending on the variables that F depends on:

1. $F = F_1(q_i, Q_i, t)$ are transformations of the first type
2. $F = F_2(q_i, P_i, t) - \sum_i Q_i P_i$ are transformations of the second type
3. $F = F_3(p_i, Q_i, t) + \sum_i p_i q_i$ are transformations of the third type
4. $F = F_4(p_i, P_i, t) + \sum_i p_i q_i - \sum_i Q_i P_i$ are transformations of the fourth type

These relations between the generating functions are similar to Legendre transformations, although it is not always true that, for a given situation, one can use any of the coordinate transformations (this may lead to expressions that are singular, ill-defined, etc.). It should also be noted that one can use generating functions that mix the above possibilities for different indices. For example, a valid generating function could be of one type for $i = 1$ and of another for $i = 2$.

One can use the same formalism as above to determine how the various generating functions give different coordinate transformations. These are summarized in table 8.1.

Type	**Transformation equations**
$F = F_1(q_i, Q_i, t)$	$p_i = \frac{\partial F_1}{\partial q_i}\ \ P_i = -\frac{\partial F_1}{\partial Q_i}$
$F = F_2(q_i, P_i, t) - \sum_i Q_i P_i$	$p_i = \frac{\partial F_2}{\partial q_i}\ \ Q_i = \frac{\partial F_2}{\partial P_i}$
$F = F_3(p_i, Q_i, t) + \sum_i p_i q_i$	$q_i = -\frac{\partial F_3}{\partial p_i}\ \ P_i = -\frac{\partial F_3}{\partial Q_i}$
$F = F_4(p_i, P_i, t) + \sum_i p_i q_i - \sum_i Q_i P_i$	$q_i = -\frac{\partial F_4}{\partial p_i}\ \ Q_i = \frac{\partial F_4}{\partial P_i}$

Table 8.1: Summary of the four types of canonical transformations

It is interesting to note that one can choose an arbitrary function $F(q_i, Q_i, t)$ and generate a canonical transformation, giving us a new set of arbitrary coordinates and Hamiltonian that will automatically satisfy the equations of motion. Being able to re-write the equations of motion in an arbitrary coordinate system is a powerful tool.

EXAMPLE 8-1: Find the transformation equations $Q_i(q_i, p_i, t)$ and $P_i(q_i, p_i, t)$ for the generating function given by $F = \sum_i Q_i q_i$

EXAMPLE 8-2: Find the transformation equations $Q_i(q_i, p_i, t)$ and $P_i(q_i, p_i, t)$ for the generating function given by $F = \sum_i q_i P_i$

EXAMPLE 8-3: Show that the harmonic oscillator problem can be solved easily using transformations from the generating function $F = \frac{m\omega}{2} q^2 \cot Q$, where $\omega = \sqrt{\frac{k}{m}}$ is the angular frequency.

8.2 Poisson Brackets and canonical transformations

Consider time-invariant canonical transformations to new coordinates $Q_i(q_i, p_i)$, $P_i(q_i, p_i)$, which preserve the Hamiltonian, H. By definition, Hamilton's canonical equations are satisfied:

$$\begin{aligned}\dot{Q}_i &= \frac{\partial H}{\partial P_i} = \sum_j \left(\frac{\partial H}{\partial q_j}\frac{\partial q_j}{\partial P_i} + \frac{\partial H}{\partial p_j}\frac{\partial p_j}{\partial P_i} \right) \\ \dot{P}_i &= -\frac{\partial H}{\partial Q_i} = -\sum_j \left(\frac{\partial H}{\partial q_j}\frac{\partial q_j}{\partial Q_i} + \frac{\partial H}{\partial p_j}\frac{\partial p_j}{\partial Q_i} \right)\end{aligned} \tag{8.11}$$

Consider the time derivative of the Q_i and P_i:

$$\begin{aligned}\dot{Q}_i &= \sum_j \left(\frac{\partial Q_i}{\partial q_j}\dot{q}_j + \frac{\partial Q_i}{\partial p_j}\dot{p}_j \right) = \sum_j \left(\frac{\partial Q_i}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial Q_i}{\partial p_j}\frac{\partial H}{\partial q_j} \right) \\ \dot{P}_i &= \sum_j \left(\frac{\partial P_i}{\partial q_j}\frac{\partial H}{\partial p_j} + \frac{\partial P_i}{\partial p_j}\dot{p}_j \right) = \sum_j \left(\frac{\partial P_i}{\partial q_j}\frac{\partial H}{\partial p_j} - \frac{\partial P_i}{\partial p_j}\frac{\partial H}{\partial q_j} \right)\end{aligned} \tag{8.12}$$

By comparing terms, we can get the "Direct Conditions" for the transformation to be canonical:

$$\begin{aligned}\frac{\partial p_j}{\partial P_i} &= \frac{\partial Q_i}{\partial q_j} \\ \frac{\partial q_j}{\partial P_i} &= -\frac{\partial Q_i}{\partial p_j} \\ \frac{\partial q_j}{\partial Q_i} &= \frac{\partial P_i}{\partial p_j} \\ \frac{\partial p_j}{\partial Q_i} &= -\frac{\partial P_i}{\partial q_j}\end{aligned} \tag{8.13}$$

Here, we obtained these relations by assuming that the transformation is canonical (that is, that we could get the time derivative of Q and P from the Hamiltonian). It is also possible to show that the Direct Conditions also hold if the transformations are time-dependent.

Now, consider the Poisson Brackets of the q_i and p_i evaluated in the coordinates Q_i and P_i, where we use the Direct conditions to simplify:

$$\begin{aligned}\{q_i, q_j\}_{Q,P} &= \sum_k \left(\frac{\partial q_i}{\partial Q_k}\frac{\partial q_j}{\partial P_k} - \frac{\partial q_i}{\partial P_k}\frac{\partial q_j}{\partial Q_k} \right) = \sum_k \left(-\frac{\partial q_i}{\partial Q_k}\frac{\partial Q_k}{\partial p_j} - \frac{\partial q_i}{\partial P_k}\frac{\partial P_k}{\partial p_j} \right) \\ &= -\frac{\partial q_i}{\partial p_j} = 0 \\ \{p_i, p_j\}_{Q,P} &= \sum_k \left(\frac{\partial p_i}{\partial Q_k}\frac{\partial p_j}{\partial P_k} - \frac{\partial p_i}{\partial P_k}\frac{\partial p_j}{\partial Q_k} \right) = \sum_k \left(\frac{\partial p_i}{\partial Q_k}\frac{\partial Q_k}{\partial q_j} + \frac{\partial p_i}{\partial P_k}\frac{\partial P_k}{\partial q_j} \right) \\ &= \frac{\partial p_i}{\partial q_j} = 0 \\ \{q_i, p_j\}_{Q,P} &= \sum_k \left(\frac{\partial q_i}{\partial Q_k}\frac{\partial p_j}{\partial P_k} - \frac{\partial q_i}{\partial P_k}\frac{\partial p_j}{\partial Q_k} \right) = \sum_k \left(\frac{\partial q_i}{\partial Q_k}\frac{\partial Q_k}{\partial q_j} + \frac{\partial q_i}{\partial P_k}\frac{\partial P_k}{\partial q_j} \right) \\ &= \frac{\partial q_i}{\partial q_j} = \delta_{ij}\end{aligned} \tag{8.14}$$

Thus it is clear that the Poisson Brackets between the coordinates is independent of the coordinate system. From the exact same algebra, it follows that:

$$\begin{aligned}\{Q_i, P_j\} &= \delta_{ij} \\ \{Q_i, Q_j\} &= \{P_i, P_j\} = 0\end{aligned} \tag{8.15}$$

which are equivalent to the Direct Conditions for checking that a transformation is canonical.

EXAMPLE 8-4: Show that the transformation $Q = p$, $P = -q$ is canonical

---

Now, consider two functions in the old coordinates, $U(q_i, p_i)$ and $V(q_i, p_i)$, with Poisson Bracket:

$$\{U, V\}_{q,p} = \sum_i^n \left(\frac{\partial U}{\partial q_i}\frac{\partial V}{\partial p_i} - \frac{\partial U}{\partial p_i}\frac{\partial V}{\partial q_i} \right) \tag{8.16}$$

Consider the same Poisson Bracket, expressed in terms of the new coordinates, Q_i, P_i, where the functions are expressed as $U(Q_i, P_i)$ and $V(Q_i, P_i)$:

$$\begin{aligned}\{U, V\}_{Q,P} =& \sum_i^n \left(\frac{\partial U}{\partial Q_i}\frac{\partial V}{\partial P_i} - \frac{\partial U}{\partial P_i}\frac{\partial V}{\partial Q_i} \right) \\ =& \sum_i^n \left[\left(\sum_j \frac{\partial U}{\partial q_j}\frac{\partial q_j}{\partial Q_i} + \frac{\partial U}{\partial p_j}\frac{\partial p_j}{\partial Q_i} \right) \left(\sum_j \frac{\partial V}{\partial q_j}\frac{\partial q_j}{\partial P_i} + \frac{\partial V}{\partial p_j}\frac{\partial p_j}{\partial P_i} \right) \right. \\ & \left. - \left(\sum_j \frac{\partial U}{\partial q_j}\frac{\partial q_j}{\partial P_i} + \frac{\partial U}{\partial p_j}\frac{\partial p_j}{\partial P_i} \right) \left(\sum_j \frac{\partial V}{\partial q_j}\frac{\partial q_j}{\partial Q_i} + \frac{\partial V}{\partial p_j}\frac{\partial p_j}{\partial Q_i} \right) \right] \\ =& \sum_j^n \left[\frac{\partial U}{\partial q_j}\frac{\partial V}{\partial q_j}\{q_j, q_j\}_{Q,P} + \frac{\partial U}{\partial q_j}\frac{\partial V}{\partial p_j}\{q_j, p_j\}_{Q,P} \right. \\ & \left. + \frac{\partial U}{\partial p_j}\frac{\partial V}{\partial p_j}\{p_j, p_j\}_{Q,P} + \frac{\partial U}{\partial p_j}\frac{\partial V}{\partial q_j}\{p_j, q_j\}_{Q,P} \right] \\ =& \sum_j^n \left[\frac{\partial U}{\partial q_j}\frac{\partial V}{\partial p_j} - \frac{\partial U}{\partial p_j}\frac{\partial V}{\partial q_j} \right] \\ =& \{U, V\}_{q,p}\end{aligned} \tag{8.17}$$

We thus find that the Poisson Bracket of two quantities is preserved in a canonical transformation. Poisson Brackets are canonical invariants.

8.3 Infinitesimal canonical transformations

In an infinitesimal canonical transformation, we have:

$$Q_i = q_i + \delta q_i \tag{8.18}$$
$$P_i = p_i + \delta p_i \tag{8.19}$$

where the δq_i and δp_i are small changes in the variables that are consistent with a canonical transformation (rather than an arbitrary virtual displacement). This transformation can be related to the identity transformation by:

$$F = \sum_i q_i P_i + \epsilon G(q_i, P_i) \tag{8.20}$$

where G is some arbitrary function and ϵ is small. The transformation for a canonical transformation of the second type are:

$$\begin{aligned} p_i &= \frac{\partial F}{\partial q_i} = P_i + \epsilon \frac{\partial G}{\partial q_i} \\ Q_i &= \frac{\partial F}{\partial P_i} = q_i + \epsilon \frac{\partial G}{\partial P_i} \end{aligned} \tag{8.21}$$

We can thus identify:

$$\begin{aligned} \delta q_i &= \epsilon \frac{\partial G}{\partial P_i} \\ \delta p_i &= -\epsilon \frac{\partial G}{\partial q_i} \end{aligned} \tag{8.22}$$

The first term can be written as:

$$\delta q_i = \epsilon \frac{\partial G}{\partial P_i} = \epsilon \frac{\partial G}{\partial (p_i + \delta p_i)} = \epsilon \frac{\partial G}{\partial p_i} \tag{8.23}$$

to first order in ϵ. That is, for a small value of ϵ, the derivative with respect to P_i is almost the same as with respect to p_i. From the properties of Poisson Brackets, this can also be written as:

$$\begin{aligned} \delta q_i &= \epsilon \frac{\partial G}{\partial p_i} = \epsilon \{q_i, G\} \\ \delta p_i &= -\epsilon \frac{\partial G}{\partial q_i} = \epsilon \{p_i, G\} \end{aligned} \tag{8.24}$$

if G is taken to be the Hamiltonian, then we have:

$$\begin{aligned} \delta q_i &= \epsilon \{q_i, H\} = \epsilon \frac{dq_i}{dt} \\ \delta p_i &= \epsilon \{q_i, H\} = \epsilon \frac{dp_i}{dt} \end{aligned} \tag{8.25}$$

and the displacement, $\epsilon \frac{dq(p)}{dt}$ are in the direction of time. That is, the Hamiltonian is the generator of the transformation that takes the variable q_i and p_i to variables Q_i and P_i an infinitesimal time later.

8.4 Hamilton-Jacobi Equation

Consider the case of a canonical transformation of the second type

$$\begin{aligned} F &= F_2(q_i, P_i, t) - \sum_i Q_i P_i \\ \frac{dF}{dt} &= \sum_i \left(\frac{\partial F_2}{\partial q_i} \dot{q}_i + \frac{\partial F_2}{\partial P_i} \dot{P}_i - Q_i \dot{P}_i - P_i \dot{Q}_i \right) + \frac{\partial F_2}{\partial t} \end{aligned} \tag{8.26}$$

Substituting into Equation 8.4:

$$\sum_i \dot{q}_i p_i - H = \sum_i \dot{Q}_i P_i - K + \sum_i \left(\frac{\partial F_2}{\partial q_i}\dot{q}_i + \frac{\partial F_2}{\partial P_i}\dot{P}_i - Q_i\dot{P}_i - P_i\dot{Q}_i \right) + \frac{\partial F_2}{\partial t}$$

$$\sum_i \left(p_i - \frac{\partial F_2}{\partial q_i} \right) \dot{q}_i - H = \sum_i \left(\frac{\partial F_2}{\partial P_i} - Q_i \right) \dot{P}_i + \frac{\partial F_2}{\partial t} - K \tag{8.27}$$

And we obtain the transformation equations:

$$\begin{aligned} p_i &= \frac{\partial F_2}{\partial q_i} \\ Q_i &= \frac{\partial F_2}{\partial P_i} \\ K &= H + \frac{\partial F_2}{\partial t} \end{aligned} \tag{8.28}$$

As a notational convention, we will use $S(q_i, P_i, t)$ instead of F_2 (as will see that it is equal to the action). We can also take $S(q_i, P_i, t)$ as the generator of a canonical transformation with the above equations, since this will guarantee that the canonical equations are satisfied. In fact, we can choose any form that we wish for S, since the above equations guarantee that the canonical equations are satisfied. Let's then choose the particular case that gives $K = 0$. This is a nice choice, since K will then be independent of all the variables, and all Q_i and P_i are cyclic:

$$\begin{aligned} \dot{Q}_i &= 0 \\ \dot{P}_i &= 0 \end{aligned} \tag{8.29}$$

We have thus changed the problem of dynamics to one of finding a transformation that gives all cyclic coordinates. Note that this is not necessarily more straightforward mathematically (in fact, it is usually much harder).

Imposing that $K = 0$ gives us:

$$\begin{aligned} H(q_i, p_i, t) + \frac{\partial S}{\partial t} &= 0 \\ H(q_i, \frac{\partial S}{\partial q_i}, t) + \frac{\partial S}{\partial t} &= 0 \end{aligned} \tag{8.30}$$

which is a partial differential equation for S and is called the "Hamilton-Jacobi equation". The function S is called "Hamilton's principal function". In the transformed coordinates, the solutions are trivial constants:

$$\begin{aligned} P_i &= \alpha_i \\ Q_i &= \beta_i = \frac{\partial S}{\partial \alpha_i} \end{aligned} \tag{8.31}$$

Is is clear that by combining this with the transformation equations, the problem is solved if all of the α_i and β_i are known. Typically, one does not need to refer to P_i and Q_i in the Hamilton-Jacobi formalism, since these are constants of motion. To highlight this, they are usually called α_i and β_i:

$$\begin{aligned} S &= S(q_i, \alpha_i, t) \\ H(q_i, \frac{\partial}{\partial q_i} S(q_i, \alpha_i, t), t) + \frac{\partial}{\partial t} S(q_i, \alpha_i, t) &= 0 \\ \alpha_i &= \text{const.} \\ \beta_i &= \frac{\partial S}{\partial \alpha_i} \end{aligned} \tag{8.32}$$

8.4.1 Connection to the Lagrangian and action

Noting that S depends on the q_i and the constants α_i, we can ask how S changes with time along the path of the system (the path where the P_i are constant and equal to α_i):

$$\begin{aligned} S &= S(q_i, \alpha_i, t) \\ \therefore \frac{dS}{dt} &= \sum_i \frac{\partial S}{\partial q_i}\dot{q}_i + \frac{\partial S}{\partial t} \end{aligned} \tag{8.33}$$

Using the transformation equations:

$$\begin{aligned} p_i &= \frac{\partial S}{\partial q_i} \\ K &= H + \frac{\partial S}{\partial t} = 0 \\ \frac{dS}{dt} &= \sum_i p_i \dot{q}_i - H = L \\ S &= \int L dt + \text{const} \end{aligned} \tag{8.34}$$

It is thus clear that S is equal to the action within an additive constant.

If the Hamiltonian does not depend on time, then H is a constant of the motion, call it E. If this is the case, one can assume a form for S where the time variable is separated out:

$$\begin{aligned} H(q_i, \frac{\partial S}{\partial q_i}, t) + \frac{\partial S}{\partial t} &= 0 \\ S(q_i, \alpha_i, t) &= W(q_i, \alpha_i) - Et \\ \therefore H(q_i, \frac{\partial W}{\partial q_i}) &= E \end{aligned} \tag{8.35}$$

Where W is called "Hamilton's characteristic function".

EXAMPLE 8-5: Use the Hamilton-Jacobi formalism to solve the simple harmonic oscillator problem

8.5 Action-angle variables

Action-angle variables are an extension to the Hamilton-Jacobi method, when the Hamiltonian does not depend on time. This formalism will allow us to determine certain fundamental frequencies of the system without explicitly solving the equations of motion. This method also allows one to transition from classical to quantum mechanics. Recall that in the case where the Hamiltonian does not depend on time, we can use Hamilton's Characteristic function, W in lieu of S:

$$\begin{aligned} S(q_i, \alpha_i, t) &= W(q_i, \alpha_i) - Et \\ p_i &= \frac{\partial W}{\partial q_i} \\ H(q_i, \frac{\partial W}{\partial q_i}) &= E \end{aligned} \tag{8.36}$$

We will consider the case where W is "separable" (at least in some of the coordinates):

$$W(q_1, \ldots, q_n, \alpha_i) = W_1(q_1, \alpha_i) + \cdots + W_n(q_n, \alpha_i) \tag{8.37}$$

that is, where the motion in each separable coordinate can be treated as independent from the others coordinates. The conjugate momentum for the separable coordinates are given by:

$$p_i = \frac{\partial W_i(q_i, \alpha_1, \ldots, \alpha_n)}{\partial q_i} \tag{8.38}$$

Furthermore, we will consider the case of "periodic' motion. Two such cases are considered. First, "vibrations" where the coordinates and their momenta, p_i, q_i, will, after a certain period of time, return to their original values. This is the case, for example, for a simple harmonic oscillator. Second, "rotations" where the momentum, p_i, is periodic as a function of its generalized coordinate, q_i. This is the case, for example, for a pendulum that has enough energy to "go over the top" (the angular position of the pendulum increases indefinitely with time and the momentum is periodic). The two cases are illustrated in phase space in Figure 8.1.

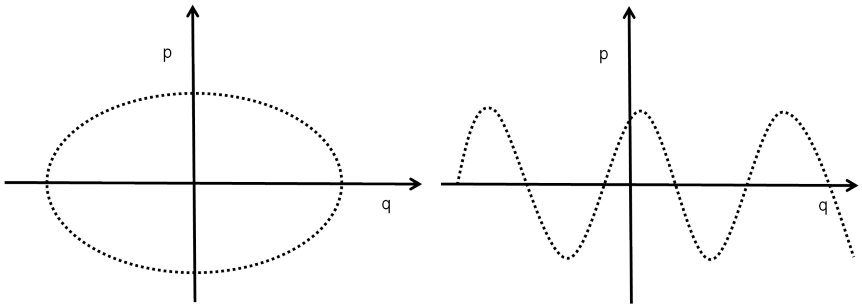

Figure 8.1: Phase space diagram of a vibration (left) and a rotation (right), both examples of periodic motion.

We start by introducing the "action-variables", J_i, obtained by integration over 1 period of the generalized coordinate:

$$J_i \equiv \oint p_i dq_i \tag{8.39}$$

There is one action variable per separable coordinate. Given the relations for the conjugate momenta for separable variables, we have:

$$\begin{aligned} J_i &= \oint p_i dq_i \\ &= \oint \frac{\partial W_i(q_i, \alpha_1, \ldots, \alpha_n)}{\partial q_i} dq_i \end{aligned} \tag{8.40}$$

and the J_i thus only depend on the α, which are constants of the motion. The J_i are thus constants of the motion as well. We assume that the equations relating the J_i and α_i are invertible:

$$\begin{aligned} J_i &= J_i(\alpha_1, \ldots, \alpha_n) \\ \alpha_i &= \alpha_i(J_1, \ldots, J_n) \end{aligned} \tag{8.41}$$

We can thus re-write Hamilton's characteristic function, W, and the Hamiltonian in terms of the J instead of the α:

$$\begin{aligned} W(q_i, \alpha_i) &\to W(q_i, J_i) \\ H(q_i, \frac{\partial W}{\partial \alpha_i}) &\to H(q_i, \frac{\partial W}{\partial J_i}) \\ S(q_i, \frac{\partial W}{\partial \alpha_i}) &\to W(q_i, J_i) - Et = W(q_i, J_i) - Ht \end{aligned} \tag{8.42}$$

In effect, we have chosen a new canonical transformation where J_i are the momenta, P_i. In terms of the generating function S, we have:

$$\begin{aligned} p_i &= \frac{\partial S}{\partial q_i} \\ Q_i &= \frac{\partial S}{\partial J_i} \end{aligned} \tag{8.43}$$

where S still satisfies the Hamilton-Jacobi equation. The new canonical momenta, J_i, are thus still constants of motion, and we have:

$$\begin{aligned} P_i &\equiv J_i = \text{const.} \\ Q_i &\equiv \beta_i = \text{const.} \end{aligned} \tag{8.44}$$

We introduce the "angle variables", w_i:

$$w_i \equiv \frac{\partial W(q_1, \ldots, q_n, J_1, \ldots, J_n)}{\partial J_i} \tag{8.45}$$

and consider the constants β_i in terms of the angle variables:

$$\begin{aligned} \beta_i &= \frac{\partial S}{\partial J_i} \\ &= \frac{\partial}{\partial J_i}(W(q_i, J_i) - Ht) \\ &= w_i - \frac{\partial H}{\partial J_i} t \end{aligned} \tag{8.46}$$

We write this as:

$$w_i = \beta_i + \nu_i t \tag{8.47}$$

where we have introduced the "frequency", ν_i (we call it frequency only because of its units for the moment):

$$\nu_i \equiv \frac{\partial H}{\partial J_i} \tag{8.48}$$

Now, assume that the system has a period of T_i for the ith degree of freedom (recall that we imposed that the system was periodic). In the amount of time T_i, the angle variable will have changed by an amount:

$$\Delta w_i \equiv w_i(t = T_i) - w_i(t = 0) = \nu_i T_i \tag{8.49}$$

We can also calculate the amount that w_i changes as the system goes through one period:

$$\begin{aligned}
\Delta w_i &= \oint dw_i \\
&= \oint \sum_j \frac{\partial w_i}{\partial q_j} dq_j \\
&= \oint \sum_j \frac{\partial}{\partial q_j} \frac{\partial W}{\partial J_i} dq_j \\
&= \frac{\partial}{\partial J_i} \oint \sum_j \frac{\partial W}{\partial q_j} dq_j \\
&= \frac{\partial}{\partial J_i} \oint \sum_j p_j dq_j \\
&= \frac{\partial}{\partial J_i} \sum_j \oint p_j dq_j \\
&= \frac{\partial}{\partial J_i} \sum_j J_j \\
&= 1
\end{aligned} \tag{8.50}$$

where we have used the definition of the angle variables, the fact that the conjugate momenta can be obtained from W, and the definition of the action variables. Equating the two expressions for Δw_i, it is clear that ν_i is indeed the frequency with which the system is periodic in the ith degree of freedom:

$$\nu_i = \frac{1}{T_i} \tag{8.51}$$

Although the derivation was not particularly intuitive, we have in fact derived a rather elegant result. Namely that for a periodic system, one can re-write the Hamiltonian in terms of the action variable and the derivative of the Hamiltonian with respect to that action variable is the frequency of the motion in the associated coordinate.

EXAMPLE 8-6: Determine the frequencies of a two-dimensional simple harmonic oscillator with different spring constants in the two dimensions

8.5.1 Connection to Quantum Mechanics

The Hamiltonian description of classical mechanics is important as it is connected with Quantum Mechanics. In this section, we look at a few similarities between the two formalisms and how they are connected.

Recall a few key points about Quantum Mechanics:

- In Quantum Mechanics, the state of a system is described by a wave function: $|\psi>$
- Observables are obtained by operating on the wave function: $\hat{X}|\psi>=x|\psi>$
- Operators do not, in general, commute: $\hat{X}\hat{Y}|\psi>\neq\hat{Y}\hat{X}|\psi>$
- In the Heisenberg representation, the state vector (wave function) is constant in time and the operators change with time
- In the time dependent Schrödinger representation, the operators are constant in time and the wave-function changes with time

Poisson brackets and commutators

You may have noticed a similarity between the use of Poisson Brackets in Classical Mechanics and Commutator relations in Quantum Mechanics. In fact, the connection between Classical and Quantum Mechanics can be made by the prescription that the Poisson Brackets be replaced with commutators:

$$\{U,V\}\rightarrow\frac{1}{i\hbar}[\hat{U},\hat{V}] \tag{8.52}$$

where $\hbar$ is Planck's constant divided by 2π. Now, consider the evolution of a system in classical mechanics; in particular, recall how the time-variation of some quantity, $F(q_i,p_i,t)$, is given by its Poisson Bracket with the Hamiltonian:

$$\frac{dF}{dt}=\{F,H\}+\frac{\partial F}{\partial t} \tag{8.53}$$

Now consider the prescription from equation 8.52 to go to Quantum Mechanics:

$$\frac{d\hat{F}}{dt}=\frac{1}{i\hbar}[\hat{F},\hat{H}]+\frac{\partial\hat{F}}{\partial t} \tag{8.54}$$

This gives precisely the time evolution of a quantum mechanical operator in the Heisenberg formulation. This "Canonical Prescription" is the most general for going from Classical to Quantum Mechanics.

The Hamilton-Jacobi equation and the Schrödinger equation

Consider the time-dependent Schrödinger equation:

$$\begin{aligned} H(q_i,p_i,t)\psi&=i\hbar\frac{\partial\psi}{\partial t}\\ p_i&=\frac{\hbar}{i}\frac{\partial}{\partial q_i}\end{aligned} \tag{8.55}$$

and consider a solution of the form:

$$\begin{aligned} \psi&=Ae^{\frac{i}{\hbar}S(q_i,t)}\\ \therefore\frac{\hbar}{i}\frac{\partial\psi}{\partial q_i}&=\hbar\psi\frac{\partial S}{\partial q_i}\\ \therefore i\hbar\frac{\partial\psi}{\partial t}&=-\hbar\psi\frac{\partial S}{\partial t}\end{aligned} \tag{8.56}$$

Substituting into the Schrödinger equation:

$$\left(H(q_i,\frac{\partial S}{\partial q_i},t)+\frac{\partial S}{\partial t}\right)\hbar\psi=0 \tag{8.57}$$

Since ψ cannot vanish, we obtain the Hamilton-Jacobi equation:

$$H(q_i, \frac{\partial S}{\partial q_i}, t) + \frac{\partial S}{\partial t} = 0 \tag{8.58}$$

There is thus a connection between the Hamilton-Jacobi equation and the Schrödinger equation. Quantum and classical mechanics meet in the limit where $\hbar \to 0$. The function S is the phase of the wave-function. In the case where we separated out the energy term from S, we are equivalently searching for stationary states in the quantum mechanics formulation.

Sommerfeld and Wilson prescription

The earliest prescription for going from Classical to Quantum Mechanics is due to Sommerfeld and Wilson who postulated that requiring that the action variables be quantized is a sufficient condition to obtain a quantum mechanical description:

$$J_i = \oint p_i dq_i = nh \tag{8.59}$$

where n is an integer and h is Planck's constant. Consider for example the simple harmonic oscillator, where we have explicitly calculated the action variable in Example 8-6:

$$J = 2\pi\alpha\sqrt{\frac{m}{k}} = nh \tag{8.60}$$

where we can identify α with the energy. Writing this in terms of the energy, we have:

$$\alpha = E = n\frac{h}{2\pi}\sqrt{\frac{k}{m}} = n\hbar\omega \tag{8.61}$$

which is the correct quantization for a simple harmonic oscillator (apart for the ground state).

Consider a particle in a central force field, described in polar coordinates, where we have seen that the momentum p_ϕ is conserved and equal to the z-component of angular momentum, L_z. The action variable is easily determined:

$$J_\phi = \oint p_\phi d\phi = \int_{)}^{2\pi} L_z d\phi = 2\pi L_z \tag{8.62}$$

Applying the Sommerfeld-Wilson quantization prescription, we find that:

$$L_z = n\hbar \tag{8.63}$$

which is the correct quantization for the angular momentum along a specific direction. One should however note that in Classical Mechanics, the total momentum is also equal to L_z and one could be wrongly tempted to conclude that the total angular momentum is quantized in units of $\hbar$.

Finally, consider the fact that the energy can be written in terms of the quantized action variables:

$$E = H(J_1, \ldots, J_n) = H(n_1 h, \ldots, n_n h) \tag{8.64}$$

And consider the change in energy when one of the action variables changes by one unit:

$$J_i \to J_i + h \tag{8.65}$$

The change in energy of the system is:

$$\Delta E = H(J_1, \ldots, J_i + h, \ldots, J_n) - H(J_1, \ldots, J_i, \ldots, J_n) \tag{8.66}$$

If we write the first term as a Taylor series (since h is small):

$$H(J_1, \ldots, J_i + h, \ldots, J_n) = H(J_1, \ldots, J_i, \ldots, J_n) + \frac{\partial H}{\partial J_i} h \tag{8.67}$$

hence:

$$\Delta E = \frac{\partial H}{\partial J_i} h = \nu_i h \tag{8.68}$$

where ν_i is the frequency of the corresponding degree of freedom. It should be clear that this is exactly equivalent to the prescription for the Bohr atom.

It should be noted that the Sommerfeld-Wilson prescription is part of the "old quantum mechanics" and was derived in an ad-hoc fashion before a more self-consistent formulation was obtained by Schrödinger, Heisenberg, Dirac and others.

8.6 Problems

Problem 8-1: Invariance of Poisson Brackets
Show that the Poisson Bracket of two functions, U, V, is invariant even under time-dependent canonical transformations, $Q_i(q_i, p_i, t)$, $P_i(q_i, p_i, t)$:

$$\{U, V\}_{q,p} = \{U, V\}_{Q,P}$$

Problem 8-2: Canonical prescription and angular momentum
Use the canonical prescription for quantization to show that the angular momentum in the z-direction is quantized.

Problem 8-3: Canonical transformation types
Show that the transformation equations for type 3 and type 4 canonical transformations are given by the following relations, respectively:

$$F = F_3(p_i, Q_i, t) + \sum_i p_i q_i$$
$$q_i = -\frac{\partial F_3}{\partial p_i}$$
$$P_i = -\frac{\partial F_3}{\partial Q_i}$$

and:

$$F = F_4(p_i, P_i, t) + \sum_i p_i q_i - \sum_i Q_i P_i$$
$$q_i = -\frac{\partial F_4}{\partial p_i}$$
$$Q_i = \frac{\partial F_4}{\partial P_i}$$

Problem 8-4: Identifying canonical transformations
Determine and show which of the following transformations are canonical:
a):

$$Q = \frac{1}{2}(q^2 + p^2)$$
$$P = -\tan^{-1}(\frac{q}{p})$$

b):

$$Q = \sqrt{2q}e^t \cos(p)$$
$$P = \sqrt{2q}e^{-t} \sin(p)$$

c):

$$Q = \ln \frac{\sin(p)}{q}$$
$$P = q \cot(p)$$

Problem 8-5: Canonical transformation of a Hamiltonian
a) Show that the following transformation is canonical:

$$Q = \frac{1}{2}(q^2 + p^2)$$
$$P = -\tan^{-1}(\frac{q}{p})$$

b) If the Hamiltonian, $H(q, p)$, is given by:

$$H = \frac{1}{2}(q^2 + p^2)$$

Write an expression of the new Hamiltonian, $K(Q, P)$, and write the equations of motion for the new canonical variables (Q, P).

c) Show that the transformed Hamiltonian, $K(Q, P)$, is a constant of the motion
d) Write an expression for the Lagrangian of this system in terms of q and $\dot{q}$
e) Write out the equation of motion for $\ddot{q}$, and give an example of a physical system that is described by this Lagrangian/Hamiltonian

9 Continuous Media and the Theory of Fields

In this chapter, we develop the formalism to deal with continuous media instead of a finite set of particles.

9.1 The stretched string

We first consider the one-dimensional longitudinal vibrations that occur in a stretched string because of its elasticity. We start by modelling a set of n beads of mass m that are held together by springs with spring constant k, as in Figure 9.1.

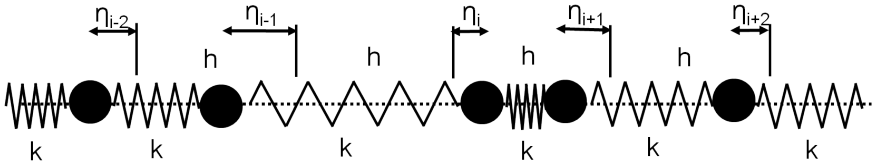

Figure 9.1: A set of beads held together by springs to model longitudinal vibrations in a stretched string. At rest, the beads are separated by a distance h.

The position of each bead is given by η_i and the string is fully described by the set of η_i. The Lagrangian for the system is given by:

$$L = T - V = \sum_i^n \frac{1}{2}\left[m\dot{\eta}_i^2 - k(\eta_{i+1} - \eta_i)^2\right] \tag{9.1}$$

and the conditions that $\eta_0 = 0$ and $\eta_n = 0$ (the ends of the string are fixed). We can apply the Euler-Lagrange equation to determine the equation of motion for η_i:

$$\begin{aligned} \frac{d}{dt}\frac{\partial L}{\partial \dot{\eta}_i} - \frac{\partial L}{\partial \eta_i} &= 0 \\ \therefore m\ddot{\eta}_i - k(\eta_{i+1} - \eta_i) + k(\eta_i - \eta_{i-1}) &= 0 \end{aligned} \tag{9.2}$$

and this can be solved more generally with the method for coupled oscillators from Chapter 6.

In going from a discrete medium to a continuous medium, one can imagine making the masses as well as the distances between them infinitesimally smaller. As we move to a continuous system, the meaning of the η_i changes from the position of mass i to the amount of distance an element of the string at position i has been displaced. In fact, we can no longer use a discrete index, i, to label the "particle" in the system. Rather, we should use a continuous variable, say x, to label position along the string. We thus describe the position of the various mass elements along the string with a continuous function $\eta(x)$. Recall, η is the generalized coordinate, not x! x is just a "label" for the generalized coordinate. Formally, $\eta(x)$ is called a "field".

Let us introduce, h, as the distance between the masses on the string when they are at rest. To go from a

discrete to a continuous system, we will let h go to zero. Introducing h, we can re-write the Lagrangian as:

$$L = \sum_i^n h \frac{1}{2} \left[\frac{m}{h} \dot{\eta}_i^2 - hk \left(\frac{\eta_{i+1} - \eta_i}{h} \right)^2 \right] \tag{9.3}$$

It is clear that the term $\frac{m}{h}$ will become the mass per unit length of the string, μ. The term hk is Young's modulus for the string, Y. Recall Hooke's Law for a continuous rod/string:

$$F = Y\lambda \tag{9.4}$$

where F is the force required to stretch the rod by an amount λ per unit length (or conversely, for a given force F, Hooke's Law indicates how much the rod will stretch/contract per unit length). In the case of a discrete system, the contraction per unit length is $\frac{\eta_{i+1} - \eta_i}{h}$, and so Hooke's law would read:

$$F = k(\eta_{i+1} - \eta_i) = hk \left(\frac{\eta_{i+1} - \eta_i}{h} \right) = Y\lambda \tag{9.5}$$

so that kh can indeed be identified with Young's modulus. We thus have the Lagrangian:

$$L = \sum_i^n h \frac{1}{2} \left[\mu \dot{\eta}_i^2 - Y \left(\frac{\eta_{i+1} - \eta_i}{h} \right)^2 \right] \tag{9.6}$$

In going to the limit of a continuous system, we have the following conditions:

$$\begin{aligned} i &\to x \\ \eta_i &\to \eta(x) \\ \eta_{i+1} &\to \eta(x+h) \\ \eta_{i+1} - \eta_i &\to \eta(x+h) - \eta(x) \to d\eta \\ h &\to dx \\ \sum_i^n h &\to \int dx \end{aligned} \tag{9.7}$$

So that the Lagrangian is given by:

$$\begin{aligned} L &= \int \frac{1}{2} \left[\mu \left(\frac{d\eta}{dt} \right)^2 - Y \left(\frac{d\eta}{dx} \right)^2 \right] dx \\ &= \int \mathcal{L}\, dx \end{aligned} \tag{9.8}$$

where we have introduced the "Lagrangian density":

$$\mathcal{L} \equiv= \frac{1}{2} \left[\mu \left(\frac{d\eta}{dt} \right)^2 - Y \left(\frac{d\eta}{dx} \right)^2 \right] \tag{9.9}$$

and we note that the field, $\eta(x,t)$, is a function of x and t. Rather than apply the variational principle to the Lagrangian, we can look at the equation of motion by converting the discrete version of the Euler-Lagrange

equation that we had to the continuous version with the same replacements:

$$\begin{aligned}
m\ddot{\eta}_i - k(\eta_{i+1} - \eta_i) + k(\eta_i - \eta_{i-1}) &= 0 \\
\mu\ddot{\eta}_i - kh\frac{(\eta_{i+1} - \eta_i) - (\eta_i + \eta_{i-1})}{h^2} &= 0 \\
\mu\ddot{\eta}_i - Y\frac{1}{dx}\frac{(\eta_{i+1} - \eta_i) - (\eta_i - \eta_{i-1})}{dx} &= 0 \\
\mu\ddot{\eta}_i - Y\frac{1}{dx}\frac{(\eta(x+dx) - \eta(x)) - (\eta(x) - \eta(x-dx))}{dx} &= 0 \\
\mu\ddot{\eta}_i - Y\frac{1}{dx}\left(\frac{d\eta}{dx}\Big|_{x+h} - \frac{d\eta}{dx}\Big|_{x}\right) &= 0 \\
\mu\frac{d^2\eta}{dt^2} - Y\frac{d^2\eta}{dx^2} &= 0
\end{aligned} \tag{9.10}$$

which is the wave equation, with propagation speed:

$$v = \sqrt{\frac{Y}{\mu}} \tag{9.11}$$

In this derivation, we did not actually apply the variational principles of mechanics. Rather, we took the Lagrangian and the result from the discrete case and made both equations continuous by taking the limit of small h. Of course, we expect that the variational principle should give us the correct equation of motion from the Lagrangian density.

9.2 The variational principle applied to a one dimensional Lagrangian density

In the stretched string example, we saw that the system is described by a Lagrangian density, $\mathcal{L}$, and the generalized coordinates are replaced by a field, $\eta(x,t)$ that depends on both position and time. The equations of motion are those that completely specify the description of the field in space and time. We also saw that the integral of the Lagrangian density over space gives the Lagrangian (hence the name Lagrangian density). In general, the Lagrangian density depends on:

$$\mathcal{L} = \mathcal{L}(\eta, \frac{d\eta}{dt}, t, \frac{d\eta}{dx}, x) \tag{9.12}$$

where the first three variables are similar to those that determine the Lagrangian, $L(q, \dot{q}, t)$, and the last two variables $\frac{d\eta}{dx}, x$ are related to the fact that η is a field. Again, we stress the point that the field plays the role of the generalized coordinate and that it depends on position in space and time. When we apply a variation to the "coordinate" η, we do not vary x and t.

Before proceeding, we tidy up the notation slightly by introducing $\eta' = \frac{d\eta}{dx}$ and $\dot{\eta} = \frac{d\eta}{dt}$, so that the Lagrangian density is written as:

$$\mathcal{L} = \mathcal{L}(\eta, \dot{\eta}, \eta', x, t) \tag{9.13}$$

The action integral is given by:

$$S = \int_{t_a}^{t_b}\int_{x_1}^{x_2} \mathcal{L}\, dx\, dt \tag{9.14}$$

and we want to find the condition under which S is stationary (Hamilton's variational principle). We proceed in a similar fashion as we did in Chapter 2. Let us introduce a small parameter, ϵ and the varied

field, $\bar{\eta}$:

$$\begin{aligned}\bar{\eta}(x,t) &= \eta(x,t) + \epsilon\phi(x,t) \\ \delta\eta \equiv &\bar{\eta}(x,t) - \eta(x,t) = \epsilon\phi(x,t)\end{aligned} \tag{9.15}$$

where ϕ is a continuous and differentiable function over the intervals in x and t where the unvaried field, $\eta(x,t)$, is continuous and differentiable. Note that $\phi(x,t)$ is identically zero at the end points of the integral, since the variation vanishes there.

The variation of the action integral is thus:

$$\delta S = \int_{t_a}^{t_b}\int_{x_1}^{x_2} \delta\mathcal{L}\, dx\, dt \tag{9.16}$$

The variation of the Lagrangian density is:

$$\delta\mathcal{L} = \mathcal{L}(\eta + \delta\eta, \dot{\eta} + \delta\dot{\eta}, \eta' + \delta\eta', x, t) - \mathcal{L}(\eta, \dot{\eta}, \eta', x, t) \tag{9.17}$$

As usual, we expand the first term using a Taylor series near the unvaried Lagrangian density:

$$\mathcal{L}(\eta + \delta\eta, \dot{\eta} + \delta\dot{\eta}, \eta' + \delta\eta', x, t) = \mathcal{L}(\eta, \dot{\eta}, \eta', x, t) + \frac{\partial\mathcal{L}}{\partial\eta}\delta\eta + \frac{\partial\mathcal{L}}{\partial\dot{\eta}}\delta\dot{\eta} + \frac{\partial\mathcal{L}}{\partial\eta'}\delta\eta' + \ldots \tag{9.18}$$

and neglect terms that are second order in the variations. This gives us the variational integral:

$$\delta S = \int_{t_a}^{t_b}\int_{x_1}^{x_2} \left(\frac{\partial\mathcal{L}}{\partial\eta}\delta\eta + \frac{\partial\mathcal{L}}{\partial\dot{\eta}}\delta\dot{\eta} + \frac{\partial\mathcal{L}}{\partial\eta'}\delta\eta'\right) dx\, dt \tag{9.19}$$

Because variation and differentiation commute, we have the following:

$$\begin{aligned}\delta\eta &= \epsilon\phi \\ \delta\dot{\eta} &= \epsilon\dot{\phi} \\ \delta\eta' &= \epsilon\phi'\end{aligned} \tag{9.20}$$

So that the variational integral becomes:

$$\delta S = \epsilon\int_{t_a}^{t_b}\int_{x_1}^{x_2} \left(\frac{\partial\mathcal{L}}{\partial\eta}\phi + \frac{\partial\mathcal{L}}{\partial\dot{\eta}}\dot{\phi} + \frac{\partial\mathcal{L}}{\partial\eta'}\phi'\right) dx\, dt \tag{9.21}$$

and we require that the rate of change of δS with respect to ϵ vanish (since S must be stationary):

$$\frac{d\delta S}{d\epsilon} = \frac{\delta S}{\epsilon} = 0 \tag{9.22}$$

Now consider the third term in the integrand, which we integrate by parts over x:

$$\begin{aligned}\int_{t_a}^{t_b}\int_{x_1}^{x_2} \frac{\partial\mathcal{L}}{\partial\eta'}\phi'\, dx\, dt &= \int_{t_a}^{t_b}\left[\phi\frac{\partial\mathcal{L}}{\partial\eta'}\right]_{x_1}^{x_2} dt - \int_{t_a}^{t_b}\int_{x_1}^{x_2} \phi\frac{d}{dx}\frac{\partial\mathcal{L}}{\partial\eta'}\, dx\, dt \\ &= \int_{t_a}^{t_b}\int_{x_1}^{x_2} \phi\frac{d}{dx}\frac{\partial\mathcal{L}}{\partial\eta'}\, dx\, dt\end{aligned} \tag{9.23}$$

however, the first term is zero since ϕ is identically zero at x_1 and x_2. Similarly, we integrate the second term by parts over t, where ϕ is identically zero at t_a and t_b:

$$\begin{aligned}\int_{t_a}^{t_b}\int_{x_1}^{x_2} \frac{\partial\mathcal{L}}{\partial\dot{\eta}}\dot{\phi}\, dx\, dt &= \int_{x_1}^{x_2}\left[\phi\frac{\partial\mathcal{L}}{\partial\dot{\eta}}\right]_{t_a}^{t_b} dt - \int_{t_a}^{t_b}\int_{x_1}^{x_2} \phi\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\eta}}\, dx\, dt \\ &= \int_{t_a}^{t_b}\int_{x_1}^{x_2} \phi\frac{d}{dt}\frac{\partial\mathcal{L}}{\partial\dot{\eta}}\, dx\, dt\end{aligned} \tag{9.24}$$

The variation integral thus becomes:

$$\frac{\delta S}{\epsilon} = \int_{t_a}^{t_b} \int_{x_1}^{x_2} \phi \left(\frac{\partial \mathcal{L}}{\partial \eta} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\eta}} - \frac{d}{dx}\frac{\partial \mathcal{L}}{\partial \eta'} \right) dx\, dt \tag{9.25}$$

Since this must equal zero for any choice of ϕ, the term in parenthesis must be zero over the entire region of integration. We obtain the Euler-Lagrange equation for a one dimensional field:

$$\therefore \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\eta}} + \frac{d}{dx}\frac{\partial \mathcal{L}}{\partial \eta'} - \frac{\partial \mathcal{L}}{\partial \eta} = 0 \tag{9.26}$$

In the case that the Lagrangian density depends on multiple one-dimension fields, it is straightforward to show that one obtains a Euler-Lagrange equation for each field (recall, the fields are similar to generalized coordinates).

EXAMPLE 9-1: Show that the Euler-Lagrange for a one-dimensional field gives the wave equation for the Lagrangian density of a stretched string, $\mathcal{L} = \frac{1}{2}\left[\mu\left(\frac{d\eta}{dt}\right)^2 - Y\left(\frac{d\eta}{dx}\right)^2\right]$

9.3 The variational principle with three-dimensional fields

We now generalize to the case where the Lagrangian density depends on a field that is defined in three dimensional space, $\eta(x, y, z, t)$. The Lagrangian density, in general, will be expressed as:

$$\mathcal{L} = \mathcal{L}(\eta, \frac{d\eta}{dt}, \frac{d\eta}{dx}, \frac{d\eta}{dy}, \frac{d\eta}{dz}, t, x, y, z) \tag{9.27}$$

It should be immediately apparent that the notation is going to be cumbersome if keep all three space coordinates and time as we apply the variational principle. One particularly interesting observation in the previous derivation for the one-dimensional field, is that the variables t and x were treated in the same way. There was nothing special about t; it was just another variable that the field depended on. With that in mind, we introduce a new notation where t is treated exactly as another one of the space coordinates. Let the 4-dimensional vector, x^μ, have the following components:

$$x^\mu \equiv (ct, x, y, z) \tag{9.28}$$

where the index μ goes from 0 to 3. We have multiplied t by a constant, c, that has dimension of speed so that ct has dimensions of length along with the other three components. Additionally, we introduce the convention that when x is indexed with a roman letter, the indices run from 1 to 3 (the three space

coordinates), and when x is indexed with a greek letter, the indices run from 0 to 3 (all four "coordinates"). The Lagrangian is thus given by:

$$L = \int \mathcal{L}\, dx^i \tag{9.29}$$

where dx^i, with the roman index, stands for $dx\, dy\, dz$. The action is then given by:

$$S = \int L dt = \int\int \mathcal{L}\, dx^i\, dt = \frac{1}{c}\int \mathcal{L}\, dx^\mu \tag{9.30}$$

where μ now runs over all four coordinates and we had to divide by c since $dt = cdx^0$. Again, note that the Lagrangian does not treat time and the space coordinates in the same way (dx^i instead of dx^μ). However, the equations of motion are obtained from the action, where all four coordinates can be combined.

We thus write the Lagrangian density as:

$$\mathcal{L}(\eta, \frac{d\eta}{dt}, \frac{d\eta}{dx}, \frac{d\eta}{dy}, \frac{d\eta}{dz}, t, x, y, z) \to \mathcal{L}(\eta, \frac{d\eta}{dx^\mu}, x^\mu) \tag{9.31}$$

We also introduce the "Einstein summation convention", where repeated indices occurring as a product must be summed over. For example, $x_\mu x^\mu \equiv x_0 x^0 + x_1 x^1 + x_2 x^2 + x_3 x^3$, where, at this point, there is no distinction between an index being a subscript or a superscript.

EXAMPLE 9-2: Use the Einstein summation convention to simplify $x_\nu y^\mu \delta^\nu_\mu$,where δ^ν_μ is the "Kronecker delta" (equal to 1 if $\nu = \mu$ and 0 otherwise).

We introduce one last notation, equivalent to the dots and apostrophes that we had for the t and x derivatives:

$$\eta_{,\mu} \equiv \frac{d\eta}{dx^\mu} \tag{9.32}$$

The comma is not a mistake! We have the comma to allow for the case when the field, η, is a vector field (or more generally a "tensor field"). For example, if η is the electric field, you can think of η as having "components" in each space direction, η_i. For example, if η were the electric field, $\vec{E}(\vec{r}, t)$, we would have:

$$\eta_{1,2} \equiv \frac{dE_x}{dy} \tag{9.33}$$

In fact, in this formalism, electro-magnetism is indeed handled by a four-dimension field, with the first component related to the electric potential, ϕ, and the other three components related to the vector potential, $\vec{A}$.

Thus, the variational principle requires us to consider:

$$\delta S = \delta \frac{1}{c} \int \mathcal{L}(\eta, \eta_{,\mu}, x^\mu)\, dx^\nu \tag{9.34}$$

Since c is constant it can be ignored as it will not influence the variation. We again consider the case where the varied field, $\bar{\eta}(x^\mu)$ is given by:

$$\begin{aligned} \bar{\eta}(x^\mu) &\equiv \eta(x^\mu) + \epsilon\phi(x^\mu) \\ \delta\eta &\equiv \epsilon\phi(x^\mu) \\ \delta\eta_{,\mu} &= \epsilon\phi_{,\mu}(x^\mu) \end{aligned} \tag{9.35}$$

The variation of the Lagrangian density is:

$$\begin{aligned} \delta\mathcal{L} &\equiv \mathcal{L}(\eta + \delta\eta, \eta_{,\mu} + \delta\eta_{,\mu}, x^\mu) - \mathcal{L}(\eta, \eta_{,\mu}, x^\mu) \\ &= \frac{\partial\mathcal{L}}{\partial\eta}\delta\eta + \frac{\partial\mathcal{L}}{\partial\eta_{,\mu}}\delta\eta_{,\mu} \end{aligned} \tag{9.36}$$

where we have used a Taylor series to expand the varied Lagrangian density around the unvaried value and dropped terms that are higher than first order in the variations. Note that the second term contains the product of two terms with the same index, which according to the Einstein summation convention must be summed over. To be explicit:

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial\eta}\delta\eta + \frac{\partial\mathcal{L}}{\partial\eta_{,\mu}}\delta\eta_{,\mu} = \frac{\partial\mathcal{L}}{\partial\eta}\delta\eta \\ + \frac{\partial\mathcal{L}}{\partial\left(\frac{d\eta}{cdt}\right)}\delta\left(\frac{d\eta}{cdt}\right) + \frac{\partial\mathcal{L}}{\partial\left(\frac{d\eta}{dx}\right)}\delta\left(\frac{d\eta}{dx}\right) + \frac{\partial\mathcal{L}}{\partial\left(\frac{d\eta}{dy}\right)}\delta\left(\frac{d\eta}{dy}\right) + \frac{\partial\mathcal{L}}{\partial\left(\frac{d\eta}{dz}\right)}\delta\left(\frac{d\eta}{dz}\right) \end{aligned} \tag{9.37}$$

which correctly reduces to the form we had earlier for the one-dimensional field. Also note how the c's cancel! Of course, the reason we introduced the notation was to avoid having to expand it out, so we will let the reader convince themselves that the math is correct by expanding out future formulas.

The variation of the action is thus:

$$\begin{aligned} \delta S &= \int \left(\frac{\partial\mathcal{L}}{\partial\eta}\delta\eta + \frac{\partial\mathcal{L}}{\partial\eta_{,\mu}}\delta\eta_{,\mu}\right) dx^\nu \\ &= \epsilon \int \left(\frac{\partial\mathcal{L}}{\partial\eta}\phi + \frac{\partial\mathcal{L}}{\partial\eta_{,\mu}}\phi_{,\mu}\right) dx^\nu \end{aligned} \tag{9.38}$$

Setting that the rate of change of the action with respect to ϵ be zero, and then integrating by parts, we have:

$$\begin{aligned} \frac{\delta S}{\epsilon} &= \int \left(\frac{\partial\mathcal{L}}{\partial\eta}\phi + \frac{\partial\mathcal{L}}{\partial\eta_{,\mu}}\phi_{,\mu}\right) dx^\nu \\ &= \int \phi \left(\frac{\partial\mathcal{L}}{\partial\eta} - \frac{d}{dx^\mu}\frac{\partial\mathcal{L}}{\partial\eta_{,\mu}}\right) dx^\nu \end{aligned}$$

Again, this must be true for any choice of ϕ, so we obtain the Euler-Lagrange equation for a scalar field in three dimensional space:

$$\frac{d}{dx^\mu}\frac{\partial\mathcal{L}}{\partial\eta_{,\mu}} - \frac{\partial\mathcal{L}}{\partial\eta} = 0 \tag{9.39}$$

This result easily generates to the case of multiple fields or multiple components of a single field (i.e. a tensor or vector field).

9.4 Problems

Problem 9-1: Transverse vibrations of a string
Derive the equation of motion for transverse vibrations in a string of mass density μ with a tension T. Show that the Lagrangian density is given by:

$$\mathcal{L} = \frac{1}{2}\left[\mu\left(\frac{d\eta}{dt}\right)^2 - T\left(\frac{d\eta}{dx}\right)^2\right]$$

where the string is stretched in the x directions and $\eta(x,t)$ correspond to displacement along y of the point on the string at position x.

Problem 9-2: Transverse and longitudinal vibrations of a string
Derive the equation of motion for a string of mass density μ with a tension T, and Young's modulus, Y, undergoing both transverse and longitudinal vibrations.