PHY321: Variational calculus

Morten Hjorth-Jensen[1,2]

[1]Department of Physics and Astronomy and Facility for Rare Ion Beams (FRIB), Michigan State University, USA
[2]Department of Physics, University of Oslo, Norway

Apr 14, 2021

Aims and Overarching Motivation

Monday.

- Euler-Lagrange equations and the Lagrangian with examples

- Principle of Least Action, watch Feynman Lecture.

- Discussions of second midterm

Reading suggestion: Taylor sections 6.3-6.4

Wednesday.

- Lagrangian formalism, constrained and unconstrained motion

- Discussions of second midterm

Reading suggestion: Taylor sections 7.1-7.4

Friday.

- Variational Calculus with examples

- Discussions of second midterm

Reading suggestion: Taylor sections 7.5-7.8

Second midterm, Angular Momentum and Kepler's Laws

```
%matplotlib inline

# let's start by importing useful packages we are familiar with
import numpy as np
from math import *
import matplotlib.pyplot as plt
import seaborn as sns
import math

#Velocity-Verlet Method
newDeltaT = 0.001
#set up arrays
tfinal = 10 # in years
n = ceil(tfinal/newDeltaT)
# set up arrays for time t, velocity v, and position r
t = np.zeros(n)
v = np.zeros((n,2))
r = np.zeros((n,2))
#newa = np.zeros((n,2))
# Initial conditions as compact 2-dimensional arrays
r0 = np.array([1.0,0.0])
v0 = np.array([0.0,2*pi])
#newa0 = np.array(-Fourpi2*newr0)
r[0] = r0
v[0] = v0
#newa[0] = newa0
Fourpi2 = 4*pi*pi
# Start integrating using Euler's method
for i in range(n-1):
    # Set up the accelerationn
    # Here you could have defined your own function for this
    rabs = sqrt(sum(r[i]*r[i]))
    a = -Fourpi2*r[i]/(rabs**3)
    # update velocity, time and position using Euler's forward method
    r[i+1] = r[i] + newDeltaT*v[i] + ((newDeltaT**2)/2)*(a)
    rabs = sqrt(sum(r[i+1]*r[i+1]))
    agh_a = -4*(pi**2)*r[i+1]/(rabs**3)
    v[i+1] = v[i] + newDeltaT*(0.5)*(a + agh_a)
    t[i+1] = t[i] + newDeltaT
sns.set()
plt.plot(r[:,0], r[:,1])
```

The Angular momentum Part

```
def AreaCalc(rad1,rad2):
    rad1n = np.linalg.norm(rad1)
    rad2n = np.linalg.norm(rad2)
    theta1 = math.atan(abs(rad1[1]/rad1[0]))
    theta2 = math.atan(abs(rad2[1]/rad2[0]))
    radn = 0.5*(rad1n+rad2n)
    delta_theta = np.abs(theta1 - theta2)
    return 0.5*delta_theta*radn**2

def AngMomentum(rad,vel):
    radn = np.linalg.norm(rad)
    veln = np.linalg.norm(vel)
    rad = rad/radn
    vel = vel/veln
```

```
        dotprod = rad[0]*vel[0]+rad[1]*vel[1]
        theta = math.acos(dotprod)
        return radn*veln*np.sin(theta)

    AreaVal = np.zeros(len(t))
    AreaVal[0] = 0
    AngMo = np.zeros(len(t))
    AngMo[0] = AngMomentum(r[0,:],v[0,:])
    for i in range(0,len(t)-1):
        AreaVal[i+1] = AreaVal[i] + AreaCalc(r[i,:],r[i+1,:])
        AngMo[i+1] = AngMomentum(r[i+1,:],v[i+1,:])

    fig, ax = plt.subplots(2,1,figsize=(12,6))
    ax[0].plot(t,AreaVal,label='Area')
    ax[0].set_title('Sweeped Area vs time')
    ax[0].set_xlabel('t [yr]')
    ax[0].set_ylabel('Area Sweeped in AU^2')

    ax[1].plot(t,AngMo,label='Angular Momentum')
    ax[1].set_title('Angular Momentum vs time')
    ax[1].set_xlabel('t [yr]')
    plt.tight_layout()
```

Building a code for the solar system, gravitational force and constants

We start with a simpler case first, the Earth-Sun system in two dimensions only. The gravitational force F_G is

$$F = \frac{GM_\odot M_E}{r^2},$$

where G is the gravitational constant,

$$M_E = 6 \times 10^{24}\text{Kg},$$

the mass of Earth,

$$M_\odot = 2 \times 10^{30}\text{Kg},$$

the mass of the Sun and

$$r = 1.5 \times 10^{11}\text{m},$$

is the distance between Earth and the Sun. The latter defines what we call an astronomical unit **AU**. From Newton's second law we have then for the x direction

$$\frac{d^2x}{dt^2} = \frac{F_x}{M_E},$$

and

$$\frac{d^2y}{dt^2} = \frac{F_y}{M_E},$$

for the y direction.

Building a code for the solar system, force equations

Introducing $x = r\cos(\theta)$, $y = r\sin(\theta)$ and

$$r = \sqrt{x^2 + y^2},$$

we can rewrite

$$F_x = -\frac{GM_\odot M_E}{r^2}\cos(\theta) = -\frac{GM_\odot M_E}{r^3}x,$$

and

$$F_y = -\frac{GM_\odot M_E}{r^2}\sin(\theta) = -\frac{GM_\odot M_E}{r^3}y,$$

for the y direction.

Building a code for the solar system, coupled equations

We can rewrite these two equations

$$F_x = -\frac{GM_\odot M_E}{r^2}\cos(\theta) = -\frac{GM_\odot M_E}{r^3}x,$$

and

$$F_y = -\frac{GM_\odot M_E}{r^2}\sin(\theta) = -\frac{GM_\odot M_E}{r^3}y,$$

as four first-order coupled differential equations

$$\frac{dv_x}{dt} = -\frac{GM_\odot}{r^3}x,$$

$$\frac{dx}{dt} = v_x,$$

$$\frac{dv_y}{dt} = -\frac{GM_\odot}{r^3}y,$$

$$\frac{dy}{dt} = v_y.$$

Building a code for the solar system, final coupled equations

The four coupled differential equations

$$\frac{dv_x}{dt} = -\frac{GM_\odot}{r^3}x,$$

$$\frac{dx}{dt} = v_x,$$

$$\frac{dv_y}{dt} = -\frac{GM_\odot}{r^3}y,$$

$$\frac{dy}{dt} = v_y,$$

can be turned into dimensionless equations (as we did in project 2) or we can introduce astronomical units with $1 \text{ AU} = 1.5 \times 10^{11}$.

Using the equations from circular motion (with $r = 1\text{AU}$)

$$\frac{M_E v^2}{r} = F = \frac{GM_\odot M_E}{r^2},$$

we have

$$GM_\odot = v^2 r,$$

and using that the velocity of Earth (assuming circular motion) is $v = 2\pi r/\text{yr} = 2\pi \text{AU}/\text{yr}$, we have

$$GM_\odot = v^2 r = 4\pi^2 \frac{(\text{AU})^3}{\text{yr}^2}.$$

Building a code for the solar system, discretized equations

The four coupled differential equations can then be discretized using the Euler-Cromer method as (with step length h)

$$v_{x,i+1} = v_{x,i} - h\frac{4\pi^2}{r_i^3}x_i,$$

$$x_{i+1} = x_i + hv_{x,i+1},$$

$$v_{y,i+1} = v_{y,i} - h\frac{4\pi^2}{r_i^3}y_i,$$

$$y_{i+1} = y_i + hv_{y,i+1},$$

Building a code for the solar system, adding Jupiter

It is rather straightforward to add a new planet, say Jupiter. Jupiter has mass

$$M_J = 1.9 \times 10^{27}\text{kg},$$

and distance to the Sun of 5.2 AU. The additional gravitational force the Earth feels from Jupiter in the x-direction is

$$F_x^{EJ} = -\frac{GM_J M_E}{r_{EJ}^3}(x_E - x_J),$$

where E stands for Earth, J for Jupiter, r_{EJ} is distance between Earth and Jupiter

$$r_{EJ} = \sqrt{(x_E - x_J)^2 + (y_E - y_J)^2},$$

and x_E and y_E are the x and y coordinates of Earth, respectively, and x_J and y_J are the x and y coordinates of Jupiter, respectively. The x-component of the velocity of Earth changes thus to

$$\frac{dv_x^E}{dt} = -\frac{GM_\odot}{r^3}x_E - \frac{GM_J}{r_{EJ}^3}(x_E - x_J).$$

Building a code for the solar system, adding Jupiter

We can rewrite

$$\frac{dv_x^E}{dt} = -\frac{GM_\odot}{r^3}x_E - \frac{GM_J}{r_{EJ}^3}(x_E - x_J).$$

to

$$\frac{dv_x^E}{dt} = -\frac{4\pi^2}{r^3}x_E - \frac{4\pi^2 M_J/M_\odot}{r_{EJ}^3}(x_E - x_J),$$

where we used

$$GM_J = GM_\odot \left(\frac{M_J}{M_\odot}\right) = 4\pi^2 \frac{M_J}{M_\odot}.$$

Similarly, for the velocity in y-direction we have

$$\frac{dv_y^E}{dt} = -\frac{4\pi^2}{r^3}y_E - \frac{4\pi^2 M_J/M_\odot}{r_{EJ}^3}(y_E - y_J).$$

Similar expressions apply for Jupiter. The equations for x and y derivatives are unchanged. This equations are similar for all other planets and as we will see later, it will be convenient to object orient this part when we program the full solar system.

How can I get the initial velocities and positions of the planets

NASA has an excellent site at `http://ssd.jpl.nasa.gov/horizons.cgi#top`. From there you can extract initial conditions in order to start your differential equation solver. At the above website you need to change from **OBSERVER** to **VECTOR** and then write in the planet you are interested in. The generated data contain the x, y and z values as well as their corresponding velocities. The velocities are in units of AU per day. Alternatively they can be obtained in terms of km and km/s.

For the first simple system involving the Earth and the Sun, you could just initialize the position with say $x = 1$ AU and $y = 0$ AU.

Variational Calculus

The calculus of variations involves problems where the quantity to be minimized or maximized is an integral.

The usual minimization problem one faces involves taking a function $\mathcal{L}(x)$, then finding the single value x for which $\mathcal{L}$ is either a maximum or minimum. In multivariate calculus one also learns to solve problems where you minimize for multiple variables, $\mathcal{L}(x_1, x_2, \cdots x_n)$, and finding the points $(x_1 \cdots y_n)$ in an n-dimensional space that maximize or minimize the function. Here, we consider what seems to be a much more ambitious problem. Imagine you have a function $\mathcal{L}(x(t), \dot{x}(t), t)$, and you wish to find the extrema for an infinite number of values

of x, i.e. x at each point t. The function $\mathcal{L}$ will not only depend on x at each point t, but also on the slope at each point, plus an additional dependence on t. Note we are NOT finding an optimum value of t, we are finding the set of optimum values of x at each point t, or equivalently, finding the function $x(t)$.

Variational Calculus, introducing the action

One treats the function $x(t)$ as being unknown while minimizing the action

$$S = \int_{t_1}^{t_2} dt \ \mathcal{L}(x(t), \dot{x}(t), t).$$

Thus, we are minimizing S with respect to an infinite number of values of $x(t_i)$ at points t_i. As an additional criteria, we will assume that $x(t_1)$ and $x(t_2)$ are fixed, and that that we will only consider variations of x between the boundaries. The dependence on the derivative, $\dot{x} = dx/dt$, is crucial because otherwise the solution would involve simply finding the one value of x that minimized $\mathcal{L}$, and $x(t)$ would equal a constant if there were no explicit t dependence. Furthermore, x wouldn't need to be continuous at the boundary.

Variational Calculus, general Action

In the general case we have an integral of the type

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

where S is the quantity which is sought minimized or maximized. The problem is that although $\mathcal{L}$ is a function of the general variables $q(t), \dot{q}(t), t$ (note our change of variables), the exact dependence of q on t is not known. This means again that even though the integral has fixed limits t_1 and t_2, the path of integration is not known. In our case the unknown quantities are the positions and general velocities of a given number of objects and we wish to choose an integration path which makes the functional $S[q]$ stationary. This means that we want to find minima, or maxima or saddle points. In physics we search normally for minima. Our task is therefore to find the minimum of $S[q]$ so that its variation δS is zero subject to specific constraints. The constraints can be treated via the technique of Lagrangian multipliers as we will see below.

Variational Calculus, Optimal Path

We assume the existence of an optimum path, that is a path for which $S[q]$ is stationary. There are infinitely many such paths. The difference between two paths δq is called the variation of q.

We call the variation $\eta(t)$ and it is scaled by a factor α. The function $\eta(t)$ is arbitrary except for

$$\eta(t_1) = \eta(t_2) = 0,$$

and we assume that we can model the change in q as

$$q(t, \alpha) = q(t) + \alpha\eta(t),$$

and

$$\delta q = q(t, \alpha) - q(t, 0) = \alpha\eta(t).$$

Variational Calculus, Condition for an Extreme Value

We choose $q(t, \alpha = 0)$ as the unkonwn path that will minimize S. The value $q(t, \alpha \neq 0)$ describes a neighbouring path.

We have

$$S[q(\alpha)] = \int_{t_1}^{t_2} \mathcal{L}(q(t, \alpha), \dot{q}(t, \alpha), t)dt.$$

The condition for an extreme of

$$S[q(\alpha)] = \int_{t_1}^{t_2} \mathcal{L}(q(t, \alpha), \dot{q}(t, \alpha), t)dt,$$

is

$$\left[\frac{\partial S[q(\alpha)]}{\partial t}\right]_{\alpha=0} = 0.$$

Variational Calculus. α Dependence

The α dependence is contained in $q(t, \alpha)$ and $\dot{q}(t, \alpha)$ meaning that

$$\left[\frac{\partial E[q(\alpha)]}{\partial \alpha}\right] = \int_{t_1}^{t_2} \left(\frac{\partial \updownarrow}{\partial q}\frac{\partial q}{\partial \alpha} + \frac{\partial \mathcal{L}}{\partial \dot{q}}\frac{\partial \dot{q}}{\partial \alpha}\right)dt.$$

We have defined

$$\frac{\partial q(x, \alpha)}{\partial \alpha} = \eta(x)$$

and thereby

$$\frac{\partial \dot{q}(t, \alpha)}{\partial \alpha} = \frac{d(\eta(t))}{dt}.$$

INtegrating by Parts

Using

$$\frac{\partial q(t, \alpha)}{\partial \alpha} = \eta(t),$$

and

$$\frac{\partial \dot{q}(t, \alpha)}{\partial \alpha} = \frac{d(\eta(t))}{dt},$$

in the integral gives

$$\left[\frac{\partial S[q(\alpha)]}{\partial \alpha}\right] = \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial q}\eta(t) + \frac{\partial \mathcal{L}}{\partial \dot{q}}\frac{d(\eta(t))}{dt}\right) dt.$$

Integrating the second term by parts

$$\int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial \dot{q}}\frac{d(\eta(t))}{dt}dt = \eta(t)\frac{\partial \mathcal{L}}{\partial \dot{q}}\Big|_{t_1}^{t_2} - \int_{a}^{b} \eta(t)\frac{d}{dx}\frac{\partial \mathcal{L}}{\partial \dot{q}}dt,$$

and since the first term dissappears due to $\eta(a) = \eta(b) = 0$, we obtain

$$\left[\frac{\partial S[q(\alpha)]}{\partial \alpha}\right] = \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dx}\frac{\partial \mathcal{L}}{\partial \dot{q}}\right)\eta(t)dt = 0.$$

Euler-Lagrange Equations

The latter can be written as

$$\left[\frac{\partial S[q(\alpha)]}{\partial \alpha}\right]_{\alpha=0} = \int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dx}\frac{\partial \mathcal{L}}{\partial \dot{q}}\right)\delta q(t)dt = \delta S = 0.$$

The condition for a stationary value is thus a partial differential equation

$$\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dx}\frac{\partial \mathcal{L}}{\partial \dot{q}} = 0,$$

known as the **Euler-Lagrange** equation.

Constrained motion

Sometimes an auxiliary constraint is added to the problem (beyond fixing the end poits y_1 and y_2). Just ahead, we will work on the example of a hanging chain. The shape of the curve minimizes the potential energy, under the constraint of a fixed length of chain. Before presenting such an example we first review the method of Lagrange multipliers as a method for finding minima or maxima under constraints.

Imagine a function $f(x_1, x_2 \cdots x_n)$ for which you wish to find the minima. Additionally, you are given a constraint

$$C(x_1 \cdots x_n) = 0 \tag{1}$$

Constrained motion, Condition for a Minimum

The usual condition for a a minimum is

$$\frac{\partial f}{\partial x_i} = 0, \quad \text{or } \nabla f = 0. \tag{2}$$

which would be n equations for the n variables. The gradient of a scalar is a vector, so you should think of ∇ as $\boldsymbol{\nabla}$. However, the solution will likely not satisfy the constraint, i.e. the point at which $f(x_1 \cdots x_n)$ has an extrema, may not be a point where $C(x_1 \cdots x_n) = 0$.

Constrained motion, Necessary Conditions

A necessary condition for the solution is that

$$\nabla f \cdot \boldsymbol{\epsilon} = 0, \tag{3}$$

for any infinitesimal vector $\boldsymbol{\epsilon}$ if $\boldsymbol{\epsilon}$ satisfies the condition

$$\delta C = \nabla C \cdot \boldsymbol{\epsilon} = 0. \tag{4}$$

Constrained motion

That is to say if I take a small step in a direction that doesn't change the constraint, then f must not change if it is an extrema. Not changing the constraint implies the step is orthogonal to ∇C. As there are n dimensions of x, the vector ∇C defines one direction, and $\boldsymbol{\epsilon}$ can be in any of the $n-1$ directions orthogonal to ∇C. If $\nabla f \cdot \boldsymbol{\epsilon} = 0$ for ANY of the $n-1$ directions of $\boldsymbol{\epsilon}$ orthogonal to ∇C, then

$$\nabla f \parallel \nabla C. \tag{5}$$

Because the two vectors are parallel you can say there must exist some constant λ such that

$$\nabla(f - \lambda C) = 0. \tag{6}$$

Constrained motionM, Lagrange Multiplier

Here, λ is known as a Lagrange multiplier. Satisfying this equation is a necessary, but not a sufficient condition. One could add a constant to the constraint and the gradient would not change. One must find the correct value of λ that satisfies the constraint $C = 0$, rather than $C =$ some other constant. The strategy is then to solve thw above equation then adjust λ until one finds the $x_1 \cdots x_n$ that gives $C(x_1 \cdots x_n) = 0$.

The method of Lagrange multipliers is counter-intuitive to one's intuition to use the constraint to reduce the dimensionality of the problem. Normally, minimizing a function of n variables, leads to n equations and n unknowns. A constraint could be used, by substitution, to replace the n variables with $n-1$ variables. Instead, we add an unknown parameter, λ, and change the equation to $n+1$ equations with $n+1$ unknowns, with the extra unknown being the Lagrange multiplier λ. Often, it is rather easy to solve for $x_1 \cdots x_n$. Then one is left with the usually difficult problem of finding λ, often requiring the solution of a transcendental equation.

Lagrange Multipliers

Let us try to formalize this. We consider a function of three independent variables $f(x, y, z)$. For the function f to be an extreme we have

$$df = 0.$$

A necessary and sufficient condition is

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = \frac{\partial f}{\partial z} = 0,$$

due to

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz.$$

Independent Variables

In physical problems the variables x, y, z are often subject to constraints (in our case q and the orthogonality constraint) so that they are no longer all independent. It is possible at least in principle to use each constraint to eliminate one variable and to proceed with a new and smaller set of independent varables.

The use of so-called Lagrangian multipliers is an alternative technique when the elimination of of variables is incovenient or undesirable. Assume that we have an equation of constraint on the variables x, y, z

$$\phi(x, y, z) = 0,$$

resulting in

$$d\phi = \frac{\partial \phi}{\partial x}dx + \frac{\partial \phi}{\partial y}dy + \frac{\partial \phi}{\partial z}dz = 0.$$

More on Independent Variables

Now we cannot set anymore

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = \frac{\partial f}{\partial z} = 0,$$

if $df = 0$ is wanted because there are now only two independent variables! Assume x and y are the independent variables. Then dz is no longer arbitrary.

However, we can add to

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz,$$

a multiplum of $d\phi$, viz. $\lambda d\phi$, resulting in

$$df + \lambda d\phi = (\frac{\partial f}{\partial z} + \lambda\frac{\partial \phi}{\partial x})dx + (\frac{\partial f}{\partial y} + \lambda\frac{\partial \phi}{\partial y})dy + (\frac{\partial f}{\partial z} + \lambda\frac{\partial \phi}{\partial z})dz = 0.$$

Choice of Multiplier

Our multiplier is chosen so that

$$\frac{\partial f}{\partial z} + \lambda\frac{\partial \phi}{\partial z} = 0.$$

However, we took dx and dy as to be arbitrary and thus we must have

$$\frac{\partial f}{\partial x} + \lambda\frac{\partial \phi}{\partial x} = 0,$$

and

$$\frac{\partial f}{\partial y} + \lambda\frac{\partial \phi}{\partial y} = 0.$$

When all these equations are satisfied, $df = 0$. We have four unknowns, x, y, z and λ. Actually we want only x, y, z, λ need not to be determined, it is therefore often called Lagrange's undetermined multiplier. If we have a set of constraints ϕ_k we have the equations

$$\frac{\partial f}{\partial x_i} + \sum_k \lambda_k \frac{\partial \phi_k}{\partial x_i} = 0.$$

Example: brachiostone I

Consider a particle constrained to move along a path (like a bead moving without friction on a wire) and you need to design a path from $x = y = 0$ to some final point x_f, y_f. Assume there is a constant force in the x direction, $F_x = mg$. Design the path so that the time the bead travels is a minimum.

The net time is

$$T = \int \frac{d\ell}{v} = \int_0^{x_f} dx \, \frac{\sqrt{1 + y'^2}}{\sqrt{2gx}} = \text{minimum}.$$

Example: brachiostone II

Here we made use of the fact that $d\ell = \sqrt{dx^2 + dy^2}$ and that the velocity is determined by $KE = mv^2/2 = mgx$. The Euler equations can be applied if you first define the function as

$$f(y, y'; x) \;\; = \;\; \frac{\sqrt{1 + y'^2}}{\sqrt{x}}.$$

The equations are then

$$\frac{d}{dx} \frac{\partial f}{\partial y'} \;\; = \;\; 0.$$

Example: brachiostone III

The simplification ensued from f not having any dependence on y. This yields the differential equation

$$\frac{y'}{x^{1/2}(1 + y'^2)^{1/2}} \;\; = \;\; (2a)^{-1/2}, \tag{7}$$

because $\partial f / \partial y'$ must be a constant, which with some foresight we label $(2a)^{-1/2}$. One can now solve for y',

$$\begin{aligned}
(y')^2 &= 2ax(1 + y'^2) \\
y' &= \sqrt{\frac{x}{2a-x}}, \\
y(t) &= \int_0^x dx' \, \frac{\sqrt{x'}dx'}{\sqrt{2a-x'}} = \int_0^x dx' \, \frac{x'dx'}{\sqrt{2ax'-x'^2}} \\
&= \frac{1}{2}\int_0^x \frac{(2x'-2a)dx'}{(2ax'-x'^2)^{1/2}} + a\int_0^x \frac{dx'}{\sqrt{2ax'-x'^2}} \\
&= \frac{-1}{2}\int_0^{2ax-x^2} \frac{du}{\sqrt{u}} + a\int_0^x \frac{dx'}{\sqrt{a^2-(x'-a)^2}} \\
&= -\sqrt{2ax-x^2} + a\cos^{-1}(1-x/a).
\end{aligned}$$

This turns out to be the equation for a *cycloid* or a *brachiostone*. If you rolled a wheel of radius a down the y axis and followed a point on the rim, it would trace out a cycloid. Here, the constant a must be chosen to match the boundary condition, $y_2 = y(x_2)$. You can see the textbook for more details, plus you get a chance to work with cycloids in the exercises at the end of this chapter.

Maximizing a Function

As an example of using Lagrange multipliers for a standard optimization formula we attempt to maximize the following function,

$$F(x_1 \cdots x_n) = -\sum_{i=1}^n x_i \ln(x_i),$$

with respect to the n variables x_i. With no constraints, each x_i would maximize the function for

$$\begin{aligned}
\frac{d}{dx_j}\left[-\sum_i x_i \ln(x_i)\right] &= 0 \\
-\ln(x_j) - 1 &= 0, \quad x_j = e^{-1}.
\end{aligned}$$

Two Constraints

Now, we repeat the problem but with two constraints,

$$\sum_i x_i = 1, \quad \sum_i x_i\epsilon_i = E.$$

Here, ϵ_i and E are fixed constants. We go forward by finding the extrema for

$$G(x_1\cdots x_n) = F - \alpha\sum_i x_i - \beta\sum_i \epsilon_i x_i = \sum_i \left\{-x_i\ln(x_i) - \alpha x_i - \beta\epsilon_i x_i\right\}.$$

Two Multipliers

There are two Lagrange multipliers, α and β, corresponding to the two constraints. One then solves for the extrema

$$
\begin{aligned}
\frac{d}{dx_j} G &= 0 \\
&= -\ln(x_j) - 1 - \alpha - \beta \epsilon_j, \\
x_j &= \exp\left\{-1 - \alpha - \beta \epsilon_j\right\}.
\end{aligned}
$$

Lagrange multipliers

For any given α and β this provides a solution for constraining $\sum_i x_i$ and $\sum_i \epsilon_i x_i$ to some values, just not the values of unity and E that you wish. One would then have to search for the correct values by adjusting α and β until the constraint are actually matched by solving a transcendental equation. Although this can be complicated, it is certainly less expensive than searching over all N values of x_i. This particular example corresponds to maximizing the entropy for a system, $S = -\sum_i x_i \ln(x_i)$, where x_i is the probability of the system being in a particular discrete level i that has energy ϵ_i. One wishes to maximize the entropy subject to the constraints that the probabilities sum to unity and the average energy has some given value. The result that $x_i \sim e^{-\beta \epsilon_i}$ demonstrates the origin of the Boltzmann factor, with the inverse temperature $\beta = 1/T$.

Lagrange multipliers

Lagrange multipliers also assist with the Euler-Lagrange equation. If one breaks an interval $x_1 < x < x_2$ into a large number $n \to \infty$ points separated by dx, the Euler-Lagrange equation involves finding the n values y_i at each point so that $\sum_i dx f \{y_i, y_i' = (y_{i+1} - y_{i-1})/(2dx)\}$ is maximized for some given function f. If an additional auxiliary constraint is added, also some function of the n values y_i, one can use the method of Lagrange multipliers. In the constraint can also be written as some function of $C(y_i, y_i')$, then one simply adds a term $\lambda C(y, y')$ to the function f and uses the Euler-Lagrange equation to find the extrema of.

$$
J = \int_{x_1}^{x_2} dx\ f\left\{y(t), y'(t), x\right\} - \lambda C\left\{y(t), y'(t), x\right\}, \tag{8}
$$

the one difference being that

$$
f\left\{y(t), y'(t), x\right\} \to f\left\{y(t), y'(t), x\right\} - \lambda C\left\{y(t), y'(t), x\right\} \tag{9}
$$

Example

Consider a chain of length L and mass per unit length κ that hangs from point $x = 0, y = 0$ to point x_f, y_f. The shape must minimize the potential energy.

Find general expressions for the shape in terms of three constants which must be chosen to match $y(0) = 0, y(x_f) = y_f$ and the fixed length. Equivalently, one finds the function $y(t)$ that provides an extrema for the integral,

One must minimize

$$\int d\ell \ \kappa g y - \lambda \int d\ell = \int_0^{x_f} dx \ \sqrt{1+y'^2}\kappa g y - \lambda \int_0^{x_f} dx\sqrt{1+y'^2}.$$

Here λ is the Lagrange multiplier associated with constraining the length of the chain. The constrained length L appears nowhere in the expression. Instead, one solves for form of the answer, then adjusts λ to give the correct length. For the purposes of the Euler-Lagrange minimization one considers the function

$$f(y, y'; x) \quad = \quad \kappa g y \sqrt{1+y'^2} - \lambda\sqrt{1+y'^2}. \tag{10}$$

Because λ is an unknown constant and because minimizing a function multiplied by a constant is the same as minimizing the function, we can equivlently minimize the integral using the function

$$\tilde{f}(y, y'; x) \quad = \quad y\sqrt{1+y'^2} - \tilde{\lambda}\sqrt{1+y'^2}, \tag{11}$$
$$\tilde{\lambda} \quad \equiv \quad \frac{\lambda}{\kappa g}.$$

The Euler-Lagrange equations then become

$$\frac{d}{dx}\left\{ \frac{y'}{\sqrt{1+y'^2}}y - \tilde{\lambda}\frac{y'}{\sqrt{1+y'^2}} \right\} \quad = \quad \sqrt{1+y'^2}.$$

Here, we will guess at the form of the solution,

$$y' \quad = \quad \sinh[(x-x_0)/a], \quad y = a\cosh[(x-x_0)/a] + y_0.$$

Plugging into the Euler-Lagange equations,

$$\frac{d}{dx}\left\{ (a\cosh[(x-x_0)/a] + y_0)\frac{\sinh[(x-x_0)/a]}{\cosh[(x-x_0)/a]} - \tilde{\lambda}\frac{\sinh[(x-x_0)/a]}{\cosh[(x-x_0)/a]} \right\} \quad = \quad \cosh[(x-x_0)/a],$$
$$\frac{d}{dx}\left\{ (y_0 - \tilde{\lambda})\tanh[(x-x_0)/a] \right\} = 0.$$

This solution works if $y_0 = \tilde{\lambda}$. So the general form of the solution is

$$y = \tilde{\lambda} + a\cosh[(x-x_0)/a].$$

One must find $\tilde{\lambda}$, x_0 and a to satisfy three conditions, $y(x = 0) = 0$, $y(x = x_f) = y_f$ and that the length is L. For a hanging chain a is positive. A solution with negative a would represent a maximum of the potential energy. A remarkable property of the solution is that once you define the length and the end-point positions y_1 and y_2, the solution does not depend on κ or g. Thus, the shape of the chain would be the same if you took it to the moon. These solutions are known as catenaries.

Lagrangians

Lagrangians represent a powerful method for solving problems that would be nearly impossible by direct application of Newton's third law, $\boldsymbol{F} = m\boldsymbol{a}$. The method works well for problems where a system is well described by a few *generalized coordinates.* A generalized coordinate might be the angle describing the position of a pendulum. This one angle takes the place of using x and y to describe the position of the pendulum, then applying a clumsy constraint.

The Lagrangian equations of motion can be derived from a principle of least action, where the action S is defined as

$$S = \int dt \; L(q, \dot{q}, t), \tag{12}$$

where q is some coordinate that describes the orientation of a system and the Lagrangian L is defined as

$$L = T - U, \tag{13}$$

the difference of the kinetic and potential energies. Minimizing the action through the Euler-Lagrange equations gives the Lagrangian equations of motion,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q}. \tag{14}$$

We begin with two simple examples, neither of which gains from the Lagrangian approach.

Consider a particle of mass m connected to a spring with stiffness k. Derive the Lagrangian equations of motion.

$$
\begin{aligned}
L &= \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2, \\
\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} &= \frac{\partial L}{\partial x}, \\
m\ddot{x} &= -kx.
\end{aligned}
$$

Derive the Lagrangian equations of motion for a pendulum of mass m and length ℓ.

$$L = \frac{m}{2}\ell^2\dot{\theta}^2 - mg\ell(1 - \cos\theta),$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} = \frac{\partial L}{\partial \theta},$$

$$m\ell^2\ddot{\theta} = -mg\ell\sin\theta,$$

$$\ddot{\theta} = -\frac{g}{\ell}\sin\theta,$$

$$\ddot{\theta} \approx -\frac{g}{\ell}\theta.$$

Proving Lagrange's Equations of Motion from Newton's Laws. Lagrange's equations of motion can only be applied for the following conditions:

- The potential energy is a function of the generalized coordinates q_i, but not of $\dot{q}_i$.

- The relation between the original coordinates $x, y, z \cdots$ and the generalized coordinates does not depend on $\dot{q}_i$, e.g. $x(q, t)$ not $x(q, \dot{q}, t)$.

- Any constraints used to reduce the number of degrees of freedom are functions of $\boldsymbol{q}$, but not of $\dot{\boldsymbol{q}}$.

- The motion is not dissipative (no damping or friction).

Going forward with the proof, consider $x_i(q_1, q_2 \cdots, t)$ and look at the l.h.s. of Lagrange's equations of motion.

$$
\begin{aligned}
\frac{\partial T}{\partial \dot{q}_j} &= \sum_i \frac{\partial T}{\partial \dot{x}_i}\frac{\partial \dot{x}_i}{\partial \dot{q}_j} + \sum_i \frac{\partial T}{\partial x_i}\frac{\partial x_i}{\partial \dot{q}_j} \qquad (15)\\
&= \sum_i m\dot{x}_i\frac{\partial \dot{x}_i}{\partial \dot{q}_j}\\
&= \sum_i m\dot{x}_i\frac{(\delta x_i/\delta t)|_{\text{fixed } q_{j'\neq j}}}{\delta q_j/\delta t}\\
&= \sum_i m\dot{x}_i\frac{\delta x_i|_{\text{fixed } q_{j'\neq j}}}{\delta q_j}\\
&= \sum_i m\dot{x}_i\frac{\partial x_i}{\partial q_j}.
\end{aligned}
$$

In the first line we used the fact that T does not depend on x. Continuing with taking the derivative of U,

$$-\frac{\partial U}{\partial \dot{q}_j} = -\sum_i \frac{\partial U}{\partial x_i}\frac{\partial x_i}{\partial \dot{q}_j} = 0. \qquad (16)$$

18

In the first line above we used the fact that U does not depend on $\dot{x}$ then we used the second condition that x does not depend on $\dot{q}$. Adding the two pieces together, then taking the derivative w.r.t. time,

$$\frac{d}{dt}\frac{\partial}{\partial \dot{q}}(T-U) \;=\; \sum_i m\ddot{x}_i \frac{\partial x_i}{\partial q_j} + \sum_i m\dot{x}_i \frac{\partial \dot{x}_i}{\partial q_j}.$$

Now, we consider the r.h.s. of Lagrange's equations. Because the kinetic energy depends only on $\dot{x}$ and not x, and because the potential depends on x but not $\dot{x}$,

$$
\begin{aligned}
\frac{\partial}{\partial q_j}(T-U) \;&=\; \sum_i \frac{\partial T}{\partial \dot{x}_i}\frac{\partial \dot{x}_i}{\partial q_j} - \sum_i \frac{\partial U}{\partial x_i}\frac{\partial x_i}{\partial q_j} \qquad (17)\\
&=\; \sum_i m\dot{x}_i \frac{\partial \dot{x}_i}{\partial q_j} - \sum_i \frac{\partial U}{\partial x_i}\frac{\partial x_i}{\partial q_j}
\end{aligned}
$$

Using the fact that $m\ddot{x}_i = -(\partial/\partial x_i)U$, one can see that the bottom expressions above are identical,

$$\frac{d}{dt}\frac{\partial}{\partial \dot{q}_i}(T-U) = \frac{\partial}{\partial q_i}(T-U). \qquad (18)$$

Lagrangian Examples.   Two examples are presented here. In the first, there are two generalized coordinates, but the two equations of motion can be reduced to one through conservation laws (angular momentum in this case). In the second, there is a time-dependent constraint.

Consider a cone of half angle α standing on its tip at the origin. The surface of the cone is defined as

$$r = \sqrt{x^2 + y^2} = z\tan\alpha.$$

Find the equations of motion for a particle of mass m moving along the surface under the influence of a constant gravitational force, $-mg\hat{z}$. For generalized coordinates use the azimuthal angle ϕ and r.

The kinetic energy is

$$
\begin{aligned}
T \;&=\; \frac{1}{2}mr^2\dot{\theta}^2 + \frac{1}{2}m(\dot{r}^2 + \dot{z}^2)\\
&=\; \frac{1}{2}mr^2\dot{\theta}^2 + \frac{1}{2}m\dot{r}^2\left(1 + \cot^2\alpha\right)\\
&=\; \frac{1}{2}mr^2\dot{\theta}^2 + \frac{1}{2}m\dot{r}^2\csc^2\alpha.
\end{aligned}
$$

The potential energy is

$$U = mgr \cot \alpha,$$

so Lagrange's equations give

$$\frac{d}{dt}\left(mr^2\dot\theta\right) = 0,$$

$$\frac{d}{dt}\left(m\csc^2\alpha\,\dot r\right) = mr\dot\theta^2 - mg\cot\alpha,$$

$$\ddot r = r\dot\theta^2\sin^2\alpha - g\cos\alpha\sin\alpha$$

The first equation is a statement of the conservation of angular momentum with $L = mr^2\dot\theta$, so the second equation can also be expressed as

$$\ddot r = \frac{L^2\sin^2\alpha}{m^2r^3} - g\sin\alpha\cos\alpha.$$

A bead slides along a wire bent in the shape of a parabola,

$$z = \frac{1}{2}kr^2, \quad r^2 = x^2 + y^2.$$

Also, the parabolic wire is rotating about the z axis with angular velocity ω. Derive the equations of motion. Are there any stable configurations?

Using the fact that

$$\dot z = \dot r\frac{\partial z}{\partial r} = kr\dot r,$$

the kinetic and potential energies are

$$T = \frac{1}{2}m\left(\dot r^2 + \dot z^2 + r^2\omega^2\right)$$

$$= \frac{1}{2}m\left(\dot r^2 + (kr\dot r)^2 + r^2\omega^2\right),$$

$$U = mgkr^2/2.$$

The equations of motion are then

$$\frac{d}{dt}\left\{m\dot r(1 + k^2r^2)\right\} = -mgkr + mk^2\dot r^2r + m\omega^2r,$$

$$\ddot r = \frac{-gkr + \omega^2r - k^2\dot r^2r}{1 + k^2r^2}$$

For a stable configuration, there needs to be a solution with $\dot r = 0$ and $\ddot r = 0$. This can only happen at $r = 0$, and then for the acceleration to be inward for small deviations of r one needs to have $gk > \omega^2$. If $\omega^2 > gk$ the bead will move outward indefinitely.

Small Vibrations and Normal Modes

Two examples are provided for solving for normal modes. These are solutions with multiple generalized coordinates, where the motion is that of simple harmonic motion. However, the motion is only simple for a particular set of coordinates q_1 and q_2,

$$q_1 = A\cos(\omega_1 t), \tag{19}$$
$$q_2 = B\cos(\omega_2 t),$$

while it is not necessarily simple in other coordinates. For example if $x = q_1 + q_2$, and $y = q_1 - q_2$, the x and y motions will contain mixtures of multiple frequencies. For many problems, or in the limit of small vibrations about a minimum, there is some coordinate system where the motion is simple. These are normal modes. Characterizing the normal modes involves finding the frequencies, ω_i, and the coordinate system where the motion is simple for each coordinate. This involves finding the direction, or the linear combination of x_i that form the coordinates q_i in which the motion is that of a single oscillator in each coordinate.

For a first example, we consider a system of springs, where we write the Lagrangian, then find the normal modes. For the second example, a double pendulum is considered. In this case, one must first make a small angle expansion before finding the modes. In principle, problems could have the same number of normal modes a degrees of freedom. For example, a system of 7 particles moving in three dimensions has 21 degrees of freedom. However, some of the degrees of freedom do not have oscillatory behavior. For example, for a rigid body in free space, the angles describing the orientation evolve, but do not oscillate. Also, the center-of-mass coordinates of a system of particles isolated from outside particles moves at constant velocity. One can also describe these as normal modes, but acknowledge that their characteristic frequency is zero, as there are no restoring forces.

Consider two springs, whose relaxed lengths are ℓ, connected to three masses as depicted in the figure here. Describe the two normal modes of the motion. We can write the Lagrangian as

$$\mathcal{L} = \frac{m}{2}\dot{x}_1^2 + m\dot{x}_2^2 + \frac{m}{2}\dot{x}_3^2 - \frac{k}{2}(x_2 - x_1 - \ell)^2 - \frac{k}{2}(x_3 - x_2 - \ell)^2.$$

There are three coordinates, thus there are three equations of motion,

$$m\ddot{x}_1 = -k(x_1 - x_2 + \ell)$$
$$2m\ddot{x}_2 = -k(x_2 - x_1 - \ell) - k(x_2 - x_3 + \ell)$$
$$= -k(2x_2 - x_1 - x_3)$$
$$m\ddot{x}_3 = -k(x_3 - x_2 + \ell).$$

This is a bit complicated because the center-of-mass motion does not easily separate from the three equations. Instead, choose the following coordinates,

$$X = \frac{x_1 + 2x_2 + x_3}{4},$$
$$q_1 = x_1 - x_2 + \ell,$$
$$q_3 = x_3 - x_2 - \ell.$$

In these coordinates the potential energy only involves two coordinates,

$$U = \frac{k}{2}(q_1^2 + q_3^2).$$

To express the kinetic energy express x_1, x_2 and x_3 in terms of X, q_1 and q_3,

$$x_1 = (3q_1 - q_3 - 4\ell + 4X)/4,$$
$$x_2 = (4X - q_1 - q_3)/4,$$
$$x_3 = (3q_3 - q_1 + 4\ell + 4X)/4.$$

The kinetic energy and Lagrangian are them

$$
\begin{aligned}
T &= \frac{m}{2}\frac{1}{16}(3\dot{q}_1 - \dot{q}_3 + 4\dot{X})^2 + m\frac{1}{16}(4\dot{X} - \dot{q}_1 - \dot{q}_3)^2 + \frac{m}{2}\frac{1}{16}(3\dot{q}_3 - \dot{q}_1 + 4\dot{X})^2 \\
&= \frac{3m}{8}(\dot{q}_1^2 + \dot{q}_3^2) - \frac{m}{4}\dot{q}_1\dot{q}_3 + 2m\dot{X}^2, \\
\mathcal{L} &= \frac{3m}{8}(\dot{q}_1^2 + \dot{q}_3^2) - \frac{m}{4}\dot{q}_1\dot{q}_3 + 2m\dot{X}^2 - \frac{k}{2}q_1^2 - \frac{k}{2}q_3^2.
\end{aligned}
$$

The three equations of motion are then,

$$\frac{3}{4}m\ddot{q}_1 - \frac{1}{4}m\ddot{q}_3 = -kq_1,$$
$$\frac{3}{4}m\ddot{q}_3 - \frac{1}{4}m\ddot{q}_1 = -kq_3,$$
$$4M\ddot{X} = 0.$$

The last equation simply states that the center-of-mass velocity is fixed. One could obtain the same result by summing the equations of motion for x_1, $2x_2$ and x_3 above. The second two equations are more complicated. To solve them, we assume a form

$$q_1 = Ae^{i\omega t},$$
$$q_3 = Be^{i\omega t},$$

Because this is a linear equation, we can multiply the solution by a constant and it will still be a solution. Thus, we can set $B = 1$, then solve for A, effectively solving for A/B. Putting this guess into the equations of motion,

$$-\frac{3}{4}\frac{A}{B}\omega^2 + \frac{1}{4}\omega^2 = -\omega_0^2\frac{A}{B},$$
$$-\frac{3}{4}\omega^2 + \frac{1}{4}\frac{A}{B}\omega^2 = -\omega_0^2.$$

This is two equations and two unknowns, ω^2 and A/B. Substituting for A/B gives a quadratic equation,

$$\omega^4 - 3\omega_0^2\omega^2 + 2\omega_0^4 = 0,$$
$$\omega_0^2 \equiv k/m.$$

The two solutions are

$$(1) \quad \omega = \omega_0, \quad A = -B,$$
$$(2) \quad \omega = \omega_0\sqrt{2}, \quad A = B.$$

The first solution corresponds to the two outer masses moving in opposite directions, in sync, with the middle mass fixed. The second solution has both outer masses moving in the same direction, but with the center mass moving opposite. These two solutions are referred to as normal modes, and are characterized by their frequency and by the linear combinations of coordinates that oscillate together. In general, the solution is a linear combination of normal modes, which usually results in a chaotic looking motion. However, once the solution is expressed in terms of the normal modes, each of which oscillates independently in a simple manner, one can better understand the motion. Further, the frequencies of these modes represent the natural resonant frequencies of the system. This is important in the construction of many structures, such as bridges or vehicles.

Consider a double pendulum confined to the $x - y$ plane, where y is vertical. A mass m is connected to the ceiling with a massless string of length ℓ. A second mass m hangs from the first mass with an identical massless string of the same length. Using θ_1 and θ_2 to describe the orientations of the strings relative to the vertical axis, find the Lagrangian and derive the equations of motion, both for arbitrary angles and in the small-angle approximation. Finally, express the equations of motion in the limit of small oscillations.

The kinetic and potential energies are:

$$
\begin{aligned}
T &= \frac{1}{2}m\ell^2\dot{\theta}_1^2 + \frac{1}{2}m\left\{(\ell\dot{\theta}_1\cos\theta_1 + \ell\dot{\theta}_2\cos\theta_2)^2 + (\ell\dot{\theta}_1\sin\theta_1 + \ell\dot{\theta}_2\sin\theta_2)^2\right\} \\
&= \frac{1}{2}m\ell^2\left\{2\dot{\theta}_1^2 + \dot{\theta}_2^2 + 2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1 - \theta_2)\right\}, \\
U &= mg\ell(1 - \cos\theta_1) + mg\left[\ell(1 - \cos\theta_1) + \ell(1 - \cos\theta_2)\right] \\
&= mg\ell(3 - 2\cos\theta_1 - \cos\theta_2)
\end{aligned}
$$

Lagrange's equations for θ_1 lead to

$$m\ell^2 \frac{d}{dt}\left\{2\dot{\theta}_1 + \dot{\theta}_2 \cos(\theta_1 - \theta_2)\right\} = -m\ell^2 \dot{\theta}_1 \dot{\theta}_2 \sin(\theta_1 - \theta_2) - 2mg\ell \sin\theta_1,$$

$$2\ddot{\theta}_1 + \ddot{\theta}_2 \cos(\theta_1 - \theta_2) + \dot{\theta}_2^2 \sin(\theta_1 - \theta_2) = -2\omega_0^2 \sin\theta_1,$$

$$\omega_0^2 \equiv g/\ell,$$

and the equations for θ_2 are

$$m\ell^2 \frac{d}{dt}\left\{\dot{\theta}_2 + \dot{\theta}_1 \cos(\theta_1 - \theta_2)\right\} = m\ell^2 \dot{\theta}_1 \dot{\theta}_2 \sin(\theta_1 - \theta_2) - mg\ell \sin\theta_2,$$

$$\ddot{\theta}_2 + \ddot{\theta}_1 \cos(\theta_1 - \theta_2) = -\omega_0^2 \sin\theta_2.$$

For small oscillations, one can only consider terms linear in θ_1 and θ_2 or their derivatives,

$$2\ddot{\theta}_1 + \ddot{\theta}_2 = -2\omega_0^2 \theta_1, \tag{20}$$
$$\ddot{\theta}_1 + \ddot{\theta}_2 = -\omega_0^2 \theta_2.$$

To find the solutions, assume they are of the form $\theta_1 = Ae^{i\omega t}, \theta_2 = Be^{i\omega t}$. Solve for ω and A/B, noting that B is arbitrary.

Plug in the desired form and find

$$e^{i\omega t}(-2\omega^2 A - \omega^2 B) = e^{i\omega t}(-2\omega_0^2 A),$$
$$e^{i\omega t}(-\omega^2 A - \omega^2 B) = e^{i\omega t}(-\omega_0^2 B).$$

We can treat B as arbitrary and set it to unity. When we find A, it is the same as A/B for arbitrary B. This gives the equations

$$2\omega^2 A + \omega^2 = 2\omega_0^2 A,$$
$$\omega^2 A + \omega^2 = \omega_0^2.$$

This is two equations and two unknowns. Solving them leads to a quadratic equation with solutions

$$A/B = \pm\frac{1}{\sqrt{2}},$$

$$\omega^2 = \frac{\omega_0^2}{1 \pm 1/\sqrt{2}}.$$

Again, these two solutions are the normal modes, and the general solution is a sum of the two solutions, with two arbitrary constants. For the angles θ_1 and θ_2 are:

$$\theta_1 = \frac{A_+}{\sqrt{2}} e^{i\omega_+ t}, \ \theta_2 = A_+ e^{i\omega_+ t},$$

$$\theta_1 = \frac{-A_-}{\sqrt{2}} e^{i\omega_- t}, \ \theta_2 = A_- e^{i\omega_- t},$$

$$\omega_\pm = \omega_0 \sqrt{\frac{1}{1 \pm 1/\sqrt{2}}}.$$

One can also express the solution in vector notation, with the vectors having arbitrary amplitudes A_+ and A_-,

$$\theta_+ = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 1 \end{pmatrix} A_+ e^{i\omega_+ t},$$

$$\theta_- = \begin{pmatrix} \frac{-1}{\sqrt{2}} \\ 1 \end{pmatrix} A_- e^{i\omega_- t}.$$

Here, the upper/lower components of the vector describe θ_1/θ_2 respectively.

These problems can be treated as linear algebra exercises. Linear algebra is not used in this course, but nonetheless we describe how this works for the curious student. In the limit of small vibrations, the equations of motion can be expressed in the form,

$$M\ddot{q} = -Kq,$$

a form that looks like the spring equation. However, q is an $n-$dimensional vector and M and k are $n \times n$ matrices. In the double pendulum example, the dimensionality is 2 and the q refers to the θ_1 and θ_2, and the matrices for M and K can be read off (add ref).

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \qquad K = \begin{pmatrix} 2\omega_0^2 & 0 \\ 0 & \omega_0^2 \end{pmatrix}.$$

Multiplying both sides of the equation by the inverse matrix M^{-1},

$$\ddot{q} = -\left(M^{-1}K\right) q.$$

Here,

$$M^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix},$$

$$M^{-1}K = \begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix} \omega_0^2.$$

One can find a transformation, basically a rotation, that transforms to a frame where $M^{-1}K$ is diagonal. In this coordinate system the diagonal components of $M^{-1}K$ represent the squared frequencies of the normal modes,

$$M^{-1}K \rightarrow - \left(\begin{array}{cc} \omega_+^2 & 0 \\ 0 & \omega_-^2 \end{array} \right) \ ,$$

and are known as "eigen" frequencies. The corresponding unit vectors,

$$\left(\begin{array}{c} 1 \\ 0 \end{array} \right) \text{ and } \left(\begin{array}{c} 0 \\ 1 \end{array} \right) \text{ in the new coordinate system,}$$

can be rotated back into the original frame, and become the solutions for the normal modes. These are then called "eigenvectors", which are the same as the normal modes. Finding the eigenfrequencies is performed by realizing that the determinant of a matrix is unchanged by the rotation between coordinate systems. Writing the equations of motion as an eigenvalue problem,

$$[A - \lambda_i \mathbb{1}] \, u_i \quad = \quad 0, \quad A \equiv M^{-1}K, \ \lambda_i \equiv \omega_i^2. \tag{21}$$

In the coordinate system where $M^{-1}K$ is diagonal, and the forms for u_i are simple this requires that in that system, the diagonal elements of $M^{-1}K$ are the eigenvalues, ω_i^2. For each ω_i^2, the determinant $|A - \lambda_i \mathbb{1}|$ must vanish. This is then true in any coordinate system,

$$\det [A - \lambda \mathbb{1}] \quad = \quad 0, \tag{22}$$

which for a 2×2 matrix becomes

$$\left| \begin{array}{cc} A_{11} - \lambda & A_{12} \\ A_{21} & A_{22} - \lambda \end{array} \right| \quad = \quad 0, \tag{23}$$

$$A_{11}A_{22} - \lambda A_{11} - \lambda A_{22} + \lambda^2 - A_{21}A_{12} \quad = \quad 0. \tag{24}$$

One can solve a quadratic equation for λ, which gives two eigenvalues corresponding to ω_+^2 and ω_-^2 found above. Choosing one of the eigenvalues, one can insert one of the eigenvalues λ_i into the eigenvalue problem and solve for u_i, then choose the other eigenvalue and solve for the other corresponding vector.

If this were a 3-dimensional set of equations, the determinant would include terms like λ^3 and would become a cubic equation with three eigenvalues. One would then solve for three eigenvectors. If one has a system with dimensionality $n > 2$, one usually resorts to solving the problem numerically due to the messiness of the algebra. The main programming languages all have packages which readily diagonalize matrices and find eigenvectors and eigenvalues.

Conservation Laws

Energy is conserved only when the Lagrangian has no explicit dependence on time, i.e. $L(q, \dot{q})$, not $L(q, \dot{q}, t)$. To show this, we first define the Hamiltonian,

$$H = \sum_i \left(\dot{q}_i \frac{\partial L}{\partial \dot{q}_i} \right) - L. \tag{25}$$

After showing that H is conserved, i.e. $(d/dt)H = 0$, we then show that H can be identified with the total energy, $H = T + V$.

One can see that H is conserved by applying first using the chain rule for $(d/dt)H$, then applying Lagrange's equations,

$$
\begin{aligned}
\frac{d}{dt}H &= \sum_i \left\{ \ddot{q}_i \frac{\partial L}{\partial \dot{q}_i} + \dot{q}_i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i} \right) - \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i - \frac{\partial L}{\partial q_i} \dot{q}_i \right\} \\
&= \sum_i \left\{ \ddot{q}_i \frac{\partial L}{\partial \dot{q}_i} + \dot{q}_i \frac{\partial L}{\partial q_i} - \frac{\partial L}{\partial \dot{q}_i} \ddot{q}_i - \frac{\partial L}{\partial q_i} \dot{q}_i \right\} \\
&= 0.
\end{aligned}
\tag{26}
$$

These steps assumed that L had no explicit time dependence, i.e. L is a function of q and $\dot{q}$, but not of t.

Next, we show that L can be identified with the energy. Because V does not depend on $\dot{q}$,

$$H = \sum_i \frac{\partial T}{\partial \dot{q}_i} \dot{q}_i - T + V. \tag{27}$$

If the kinetic energy has a purely quadratic form in terms of $\dot{q}$,

$$T = \sum_{ij} A_{ij}(q) \dot{q}_i \dot{q}_j, \tag{28}$$

the Hamiltonian becomes

$$
\begin{aligned}
H &= \sum_{ij} 2 A_{ij}(q) \dot{q}_i \dot{q}_j - \sum_{ij} A_{ij}(q) \dot{q}_i \dot{q}_j + V \\
&= T + V.
\end{aligned}
\tag{29}
$$

The proof that H equals the energy hinged on the fact that the kinetic energy was quadratic in $\dot{q}$. This can be attributed to time-reversal symmetry. Because the Cartesian coordinates x_i do not depend on $\dot{q}_i$ or on time, $\dot{x}_i = (\partial x_i / \partial q_j) \dot{q}_j$. Thus, the kinetic energy, $T = m\dot{x}_i^2/2$, should be proportional to two powers of $\dot{q}$, which validates the assumption above.

Here, energy conservation is predicated on the Lagrangian not having an explicit time dependence. Without an explicit time dependence the equations

of motion are unchanged if one translates a fixed amount in time because the physics does not depend on when the clock starts. In contrast, the absolute time becomes relevant if there is an explicit time dependence. In fact, conservation laws can usually be associated with symmetries. In this case the translation symmetry in time leads to energy conservation.

For another example of how symmetry leads to conservation laws, consider a Lagrangian for a particle of mass m moving in a two-dimensional plane where the generalized coordinates are the radius r and the angle θ. The kinetic energy would be

$$T = \frac{1}{2}m\left\{\dot{r}^2 + r^2\dot{\theta}^2\right\}, \tag{30}$$

and if the potential energy $V(r)$ depends only on the radius r and not on the angle, Lagrange's equations become

$$\frac{d}{dt}(m\dot{r}) = -\frac{\partial V}{\partial r} + m\dot{\theta}^2 r, \tag{31}$$
$$\frac{d}{dt}(mr^2\dot{\theta}) = 0.$$

The second equation implies that $mr^2\dot{\theta}$ is a constant. Indeed, it is the angular momentum which is conserved for a radial force. Here, the conservation of angular momentum is associated with the independence of the physics to changes in θ, or in other words, rotational invariance. Once one knows the fact that $L = mr^2\dot{\theta}$ is conserved, it can be inserted into the equations of motion for $\dot{r}$,

$$m\ddot{r} = -\frac{\partial V}{\partial r} + \frac{L^2}{mr^3}. \tag{32}$$

This is related to Emmy Noether's theorem

Simply stated, if the Lagrangian L is independent of q_i, one can see that the quantity $\partial L/\partial \dot{q}_i$ is conserved,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} = 0. \tag{33}$$

Another easy example is in Cartesian coordinates where the potential depends only on x and y but not on z. In that case, there is a translational symmetry. From the last equation, this translates into conservation of the momentum in the z direction.

Consider a pair of particles of mass m_1 and m_2 where the potential is of the form

$$U(\boldsymbol{r}_1, \boldsymbol{r}_2) = V_a(|m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2|/(m_1 + m_2)) + V_b(|\boldsymbol{r}_1 - \boldsymbol{r}_2|).$$

Using symmetry arguments alone, are there any conserved components of the momentum? or the angular momentum??

There is no translational invariance, hence there are no conserved components of the momentum. However, there is rotational invariance about any axis that goes through the origin. Hence, there is angular momentum conservation in all three directions. Symmetry arguments are great ways to recognize the existence of conserved quantities, but actually expressing them in terms of coordinates can be tricky. For instance, you may need to write the Lagrangian in terms of angles.