PHY321: Variational calculus

Morten Hjorth-Jensen[1,2]

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

April 10-14

Aims and Overarching Motivation

Monday April 10.

- Lagrangian formalism, top down approach first and derivation of the Euler-Lagrange equations
- Principle of Least Action, watch Feynman Lecture.
- Video of lecture
- Handwritten notes

Reading suggestion this week: Taylor sections 6.1-6.4

Wednesday April 12.

- Euler-Lagrange equations and the Lagrangian with examples
- Principle of Least Action, watch Feynman Lecture.
- Video of lecture TBA
- Handwritten notes

Reading suggestion: Taylor sections 6.1-6.4
See also Variational Calculus

Friday April 14. Discussion and work on second midterm.

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

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

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.

Why is the Lagrangian defined as the difference between kinetic and potential energy?

To understand this let us develop some intuition before the math by looking at what we did in the second midterm. There we studied energy conservation.

```
# 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
DeltaT = 0.001
#set up arrays
tfinal = 10 # in years
n = ceil(tfinal/DeltaT)
# 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))
# Initial conditions as compact 2-dimensional arrays. Here: circular orbit conditions.
r0 = np.array([1.0,0.0])
v0 = np.array([0.0,5.0])
r[0] = r0
v[0] = v0
Fourpi2 = 4*pi*pi
# Start integrating using the Velocity-Verlet method
for i in range(n-1):
    # Set up the accelerationn
    rabs = sqrt(sum(r[i]*r[i]))
    a = -Fourpi2*r[i]/(rabs**3)
    # update velocity, time and position using the Velocity-Verlet  method
    r[i+1] = r[i] + DeltaT*v[i] + ((DeltaT**2)/2)*(a)
    rabs = sqrt(sum(r[i+1]*r[i+1]))
    anew = -4*(pi**2)*r[i+1]/(rabs**3)
    v[i+1] = v[i] + DeltaT*(0.5)*(a + anew)
    t[i+1] = t[i] + DeltaT
sns.set()
plt.plot(r[:,0], r[:,1])


# We check that the total energy is conserved. For a circular orbit, potential and kinetic energy

# Note that we have set the mass of the Earth = 1
def kinetic_energy(v):
    KE = []
    step = len(t)
    for i in range(step):
        KE.append(0)
        KE[i] += 0.5 *np.sum(v[i]*v[i])
    return np.array(KE)


# Note that G x Mass_sun = 4*pi*pi and the mass of the Earth = 1
# Note also that if you change the exponent in the force you need also to change the potential en
def pot():
    Pot = []
    step = len(t)
```

```
    for i in range(step):
        Pot.append(0)
        Pot[i] +=  - 4*pi*pi/ sqrt(np.sum(r[i]*r[i]))
    return np.array(Pot)

fig, ax = plt.subplots(1,1,figsize=(12,6))
ax.plot(t,kinetic_energy(v)+pot(),label='Total')
ax.plot(t,kinetic_energy(v),label='Kinetic')
ax.plot(t,pot(),label='Potential')
ax.set_title('Energy vs time')
ax.set_xlabel('t [yr]')
ax.legend()
ax.set_ylabel(r'E')
```

The energy is conserved and does not say much about the variations in position and velocity as functions of time.

What if we plot the difference between kinetic and potential energy instead?

```
fig, ax = plt.subplots(1,1,figsize=(12,6))
ax.plot(t,kinetic_energy(v)-pot(),label='L=K-V')
ax.plot(t,kinetic_energy(v),label='Kinetic')
ax.plot(t,pot(),label='Potential')
ax.set_title('Energy vs time')
ax.set_xlabel('t [yr]')
ax.legend()
ax.set_ylabel(r'E')
```