Part I

Physics for mathematicians

Classical Mechanics

1 Galilean Spacetime

In classical mechanics the motion of (unconstrained) point particles are described by a special class of curves in **Galilean spacetime** M^4_{Gal} (we think of this as physical spacetime). A spacetime coordinate system, or **frame**, is a bijection $\psi : M^4_{\text{Gal}} \longrightarrow \mathbb{R} \times \mathbb{R}^3$:

$$\psi : \text{event} \mapsto (t, \mathbf{x}) = (t, x, y, z)$$

It is assumed that the motion of a particle is described in a frame by

$$t \mapsto (t, \mathbf{c}(t))$$

where $\mathbf{c} : \mathbb{R} \to \mathbb{R}^3$ will be taken to be a sufficiently differentiable curve.

It is assumed that once length and time units are chosen, then there is a distinguished class of frames called **inertial frames** such that a particle which experiences no influences from other bodies, or is not subject to any forces, will be described by linear motion

$$\mathbf{c}(t) = t\mathbf{v} + \mathbf{x}_0$$

Transformations from one inertial frame to another are effected by **Galilean transformations**:

$$g : (t, \mathbf{x}) \mapsto (t + \tau, R\mathbf{x} + t\mathbf{V} + \mathbf{a})$$

where $R \in \mathrm{O}(3)$ is a rotation, $\tau \in \mathbb{R}$,$\mathbf{V} \in \mathbb{R}^3$ and $\mathbf{a} \in \mathbb{R}^3$

$$\begin{array}{ccc} & & \mathbb{R} \times \mathbb{R}^3 \\ & \nearrow & \\ M^4_{\text{Gal}} & & \downarrow g \\ & \searrow & \\ & & \mathbb{R} \times \mathbb{R}^3 \end{array}$$

We apply such a transformation when we want to describe the system from the point of view of new orthogonal axes displaced from the original axes by $\mathbf{a}$, rotated by R and moving with velocity $\mathbf{V}$ with respect to the first set of axes.

We can represent the group of Galilean coordinate transformations as a matrix group by using a trick. Namely, represent $(t, \mathbf{x})$ by $\begin{bmatrix} \mathbf{x} \\ t \\ 1 \end{bmatrix}$. Then we have

$$\begin{bmatrix} \mathbf{x} \\ t \\ 1 \end{bmatrix} \mapsto \begin{bmatrix} R & \mathbf{V} & \mathbf{a} \\ 0 & 1 & \tau \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ t \\ 1 \end{bmatrix} = \begin{bmatrix} R\mathbf{x}+t\mathbf{V}+\mathbf{a} \\ t+\tau \\ 1 \end{bmatrix}$$

When expressed in inertial coordinates, Newtons famous law states that there is some vector valued function $\mathbf{F} : \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ such that

$$m\ddot{\mathbf{c}}(t) = \mathbf{F}(t, \mathbf{c}(t), \dot{\mathbf{c}}(t))$$

where m is the mass of the object. Actually more general force laws are possible but this is a good starting point. In any case we have $\mathbf{F} = m\ddot{\mathbf{c}}(t)$.

We often use $\mathbf{x} = \mathbf{c}(t)$ and even write $\mathbf{x} = \mathbf{x}(t)$.

Many particles are described in an inertial frame by maps $\mathbf{c}_a : I \to \mathbb{R}^3$ which describe the motion of the a-the particle. Suppose there are N particles. Then the map corresponding to the i-th particle then

$$\mathbf{c} := (\mathbf{c}_1, ..., \mathbf{c}_N)$$

gives a map (a motion of the system)

$$\mathbf{c} : I \to \mathbb{R}^3 \times \cdots \times \mathbb{R}^3 = \mathbb{R}^{3N} = \mathbb{R}^n$$

We assume these maps are twice continuously differentiable. We denote a typical element of $\mathbb{R}^{3N} = \mathbb{R}^n$ by $\mathbf{x}$.

In this case the motions is described by a systems of second order differential equations.

$$\ddot{\mathbf{c}} = \mathbf{G}(t, \mathbf{c}, \dot{\mathbf{c}})$$

If the system of particles is isolated from and uninfluenced by outside forces we say the system is **closed**. It turns out that for such systems the force law takes the simple form

$$\ddot{\mathbf{c}} = \mathbf{G}(\mathbf{c})$$

for some vector field $\mathbf{G}$ on $\mathbb{R}^{3N}$.

It is often assumed that the equations of motion may be derived as the **Euler-Lagrange** equations for a functional S defined in terms of a Lagrangian function $L : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$. Using L we define a functional as follows:

$$S_{t_1,t_2}[\mathbf{c}] = \int_{t_1}^{t_2} L(t, \mathbf{c}\,(t)\,, \dot{\mathbf{c}}\,(t)\,)\,dt.$$

Notice that for fixed end-times t_1, t_2, the functional $\mathbf{c} \mapsto S_{t_1,t_2}[\mathbf{c}]$ is a real valued map defined on some space of curves. If we fix times t_1 and t_2 and consider the variation of $S_{t_1,t_2}[\mathbf{c}]$ over all curves for which $\mathbf{c}\,(t_1) = p_1$ and $\mathbf{c}\,(t_2) = p_2$ then the functional will be stationary at $\mathbf{c}$ if $\mathbf{c}$ is a physical motion. This leads to the **Euler-Lagrange equations**.

We define the variation δS of S as follows. Fix a curve $\mathbf{c}$ as above and let $\mathbf{h} : [t_1, t_2] \to$ be any curve

$$\delta S_{\mathbf{c}}(\mathbf{h}) = \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} S[\mathbf{c} + \varepsilon \mathbf{h}]$$

Let the variables of $L : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be named $(t, \mathbf{x}, \mathbf{v})$.Then the chain rule gives

$$\begin{aligned} \delta S_{\mathbf{c}}(\mathbf{h}) &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial \mathbf{x}}(t, \mathbf{c}, \dot{\mathbf{c}}) \cdot \mathbf{h} + \frac{\partial L}{\partial \mathbf{v}}(t, \mathbf{c}, \dot{\mathbf{c}}) \cdot \dot{\mathbf{h}} \right] dt \\ &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial \mathbf{x}}(t, \mathbf{c}, \dot{\mathbf{c}}) - \left(\frac{d}{dt} \frac{\partial L}{\partial \mathbf{v}}(t, \mathbf{c}, \dot{\mathbf{c}}) \right) \right] \cdot \mathbf{h} \, dt \\ &\quad + \left[\frac{\partial L}{\partial \mathbf{v}}(t, \mathbf{c}, \dot{\mathbf{c}}) \cdot \mathbf{h} \right]_{t_1}^{t_2} \end{aligned}$$

where, for example,

$$\frac{\partial L}{\partial \mathbf{v}} \cdot \mathbf{h} = \sum_{i=1}^{3N} \frac{\partial L}{\partial v_i} h_i$$

Now if $\delta S_{\mathbf{c}}(\mathbf{h}) = 0$ for all $\mathbf{h}()$ which are zero at end-times then the last term vanishes for such $\mathbf{h}$ and

$$\int_{t_1}^{t_2} \left[\frac{\partial L}{\partial \mathbf{x}}(t, \mathbf{c}, \dot{\mathbf{c}}) - \left(\frac{d}{dt} \frac{\partial L}{\partial \mathbf{v}}(t, \mathbf{c}, \dot{\mathbf{c}}) \right) \right] \cdot \mathbf{h} \, dt = 0.$$

Since this is true for all such variations $\mathbf{h}$, it follows that

$$\frac{\partial L}{\partial \mathbf{x}}(t, \mathbf{c}, \dot{\mathbf{c}}) - \frac{d}{dt} \frac{\partial L}{\partial \mathbf{v}}(t, \mathbf{c}, \dot{\mathbf{c}}) = 0 \text{ for all } t \in (t_1, t_2)$$

It is common to use the same symbols to denote the elements of $\mathbb{R}^n \times \mathbb{R}^n$ as we use to denote the curve and its velocity. So we write $\mathbf{x}$ and $\mathbf{v}$ as $\mathbf{x}$ and $\dot{\mathbf{x}}$ and $\mathbf{c}(t)$ and $\dot{\mathbf{c}}(t)$ as $\mathbf{x}(t)$ and $\dot{\mathbf{x}}(t)$. Then, still using an inertial frame, we write the E-L equations as

$$\frac{\partial L}{\partial \mathbf{x}}(t, \mathbf{x}, \dot{\mathbf{x}}) - \frac{d}{dt} \frac{\partial L}{\partial \dot{\mathbf{x}}}(t, \mathbf{x}, \dot{\mathbf{x}}) = 0$$

where

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{x}} &= \left(\frac{\partial L}{\partial x_1}, ..., \frac{\partial L}{\partial x_{3N}} \right) \\ \frac{\partial L}{\partial \dot{\mathbf{x}}} &= \left(\frac{\partial L}{\partial \dot{x}_1}, ..., \frac{\partial L}{\partial \dot{x}_{3N}} \right) \end{aligned}$$

If we group the variables by threes $\mathbf{x} = (\mathbf{x}_1, ..., \mathbf{x}_N) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \cdots \times \mathbb{R}^3$, where each triple $\mathbf{x}_a$ gives the coordinates of possible positions of the a-th particle then the E-L equation gives N separate equations

$$\frac{\partial L}{\partial \mathbf{x}_a}(\mathbf{x}(t), \dot{\mathbf{x}}(t)) - \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{x}}_a}(\mathbf{x}(t), \dot{\mathbf{x}}(t)) = 0,$$

one vector equation for each particle. (still using an inertial frame with rectangular coordinates)

....

We will show below that for a certain Lagrangians the Euler-Lagrange equations produce Newton's $\mathbf{F} = m\mathbf{a}$ law.

....

We may even use curvilinear coordinates $\mathbf{q} = (q_1, ..., q_{3N})$ and derive Euler-Langrange equations

$$\frac{\partial L}{\partial \mathbf{q}}(t, \mathbf{q}, \dot{\mathbf{q}}) - \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}}(t, \mathbf{q}, \dot{\mathbf{q}}) = 0$$

If the motion is constrained to lie on a submanifold in the space $\mathbb{R}^{3N}$ then

there may only be k degrees of freedom (k = dimension of the submanifold). Then we take $\mathbf{q} = (q_1, ..., q_k)$ as local coordinates on this submanifold and we end up with

$$\begin{aligned}
\frac{\partial L}{\partial \mathbf{q}}(t, \mathbf{q}, \dot{\mathbf{q}}) - \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{x}}}(t, \mathbf{q}, \dot{\mathbf{q}}) &= 0 \\
&\quad \text{or} \\
\frac{\partial L}{\partial q_i}(t, \mathbf{q}, \dot{\mathbf{q}}) - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}(t, \mathbf{q}, \dot{\mathbf{q}}) &= 0 \text{ for all } 1 \leq i \leq k
\end{aligned}$$

We will give some justification later that such constrained motion can be treated using Lagrangians in this way.

If we assume that motion takes place along a line then we just obtain

$$\frac{\partial L}{\partial x}(t, x) - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = 0$$

Example 1 ***Spring***

Here $L(t, x, \dot{x}) = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2$
The Euler-Lagrange equation becomes

$$m\ddot{x} = -kx$$

Indeed,

$$\begin{aligned}
\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} &= \frac{\partial L}{\partial x} \\
\frac{d}{dt}m\dot{x} &= -kx
\end{aligned}$$

Note also that $\frac{d}{dt}\left(\frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2\right) = m\ddot{x} + kx = 0$ as long as $x(t)$ is a solution.

This later fact is a consequence of conservation of Energy which holds whenever L does not depend explicitly on t (see below).

2 Conservation of Energy

We may as well use curvilinear coordinates for maximum generality. Suppose that $L(\mathbf{q}, \dot{\mathbf{q}})$ does not depend explicitly on time. This must be the case for a closed system if the laws of physics remain "constant" with time.

Then we have for $L = L(\mathbf{q}(t), \dot{\mathbf{q}}(t))$ for a solution of the E-L equations,

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

Or

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

The quantity $\sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L$ will typically have units of energy and is called the **energy** of the system. In this case, the system is energy **conservative**. This is valid not only for closed systems but is also true anytime L does not explicitly depend on time.

If our system of point particles is a closed system in 3d space and therefore given by a curve in $\mathbb{R}^{3N} = \mathbb{R}^n$, then the proper Lagrangian (valid also for constant external fields) is of the form

$$L(\mathbf{q}, \dot{\mathbf{q}}) = T(\dot{\mathbf{q}}) - U(\mathbf{q})$$

where T is homogeneous of degree 2:

$$s^2 T(\dot{\mathbf{q}}_1, ..., \dot{\mathbf{q}}_n) = T(s\dot{\mathbf{q}}_1, ..., s\dot{\mathbf{q}}_n) \text{ for all } s$$

T is called the **kinetic energy** term and U is the **potential energy** term.

Now by **Euler's theorem on homogeneous functions** we have $\sum_i \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} = 2T$ and so

$$\begin{aligned} \sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L &= \sum_i \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} - L \\ &= 2T - L = T + U. \end{aligned}$$

Thus the total energy (conserved!) is given by

$$E_{\text{tot}} = T + U$$

3 Force on a particle

Let our indexing respect the possible coordinates of the particles as above, so that $\mathbf{q} = (\mathbf{q}_1, ..., \mathbf{q}_N)$ with $\mathbf{q}_a \in \mathbb{R}^3$.Then if $\mathbf{q}(t)$ satisfied the E-L equations then

$$0 = \frac{\partial L}{\partial \mathbf{q}_a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}_a} = -\frac{\partial U}{\partial \mathbf{q}_a} - \frac{d}{dt}\frac{\partial T}{\partial \dot{\mathbf{q}}_a}$$

$$\begin{aligned} \frac{d}{dt}\frac{\partial T}{\partial \dot{\mathbf{q}}_a} &= -\frac{\partial U}{\partial \mathbf{q}_a} \\ \frac{d}{dt}\mathbf{p}_a &= -\frac{\partial U}{\partial \mathbf{q}_a} \end{aligned}$$

$$0 = \frac{\partial L}{\partial \mathbf{q}_a} - \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}_a} = -\frac{\partial U}{\partial \mathbf{q}_a} - \frac{d}{dt}\frac{\partial T}{\partial \dot{\mathbf{q}}_a}$$

$$\begin{aligned} \frac{d}{dt}\frac{\partial T}{\partial \dot{\mathbf{q}}_a} &= -\frac{\partial U}{\partial \mathbf{q}_a} \\ \frac{d}{dt}\mathbf{p}_a &= -\frac{\partial U}{\partial \mathbf{q}_a} \end{aligned}$$

where $\mathbf{p}_a = \frac{\partial T}{\partial \dot{\mathbf{q}}_a}$ is a **generalized momentum**. If in *rectangular coordinates* we have $L(\mathbf{x}, \dot{\mathbf{x}}) = T(\dot{\mathbf{x}}) - U(\mathbf{x})$ where

$$T(\dot{\mathbf{x}}) = \sum \frac{1}{2} m_a \left\| \dot{\mathbf{x}}_a \right\|^2,$$

then we obtain

$$\frac{\partial T}{\partial \dot{\mathbf{x}}_a} = m_a \dot{\mathbf{x}}_a$$

and

$$\begin{aligned} \frac{d}{dt} m_a \dot{\mathbf{x}}_a &= -\frac{\partial U}{\partial \mathbf{x}_a} \\ m_a \ddot{\mathbf{x}}_a &= -\frac{\partial U}{\partial \mathbf{x}_a} \end{aligned}$$

The vector $-\frac{\partial U}{\partial \mathbf{x}_a} = \mathbf{F}_a$ is called the **force on the a-th particle**. Notice that $\mathbf{F}_a$ does not depend on particle velocities. For the a-th particle,

$$m_a \ddot{\mathbf{x}}_a = \mathbf{F}_a \quad \text{(Newton)}.$$

Another look at Mechanical Energy. Consider a single particle possibly in an external field.

$$W = \int_C \mathbf{F} \cdot \mathbf{dx} = \int_a^b \mathbf{F} \cdot \frac{\mathbf{dx}}{dt} dt = \int_a^b \mathbf{F} \cdot \frac{\mathbf{dx}}{dt} dt$$
$$= \int_a^b m\mathbf{a} \cdot \mathbf{v}\, dt = \int_a^b m \frac{d\mathbf{v}}{dt} \cdot \mathbf{v}\, dt = \int_a^b \frac{1}{2} m \left\| \mathbf{v} \right\|^2\, dt$$

$= \frac{1}{2}m \left\| \mathbf{v}(b) \right\|^2 - \frac{1}{2}m \left\| \mathbf{v}(a) \right\|^2 = \Delta\, KE = \Delta T$

On the other hand if $\mathbf{F} = \boldsymbol{\nabla} U$ then $\int_C \mathbf{F} \cdot \mathbf{dx} = \int_{\mathbf{C}} \boldsymbol{\nabla} U \cdot \mathbf{dx} = U(\mathbf{x}(b)) - U(\mathbf{x}(a))$ therefore $U(\mathbf{x}(b)) - U(\mathbf{x}(a)) = \frac{1}{2}m \left\| \mathbf{v}(b) \right\|^2 - \frac{1}{2}m \left\| \mathbf{v}(a) \right\|^2$ or

$$U(\mathbf{x}(a)) + \frac{1}{2}m \left\| \mathbf{v}(a) \right\|^2 = U(\mathbf{x}(b)) + \frac{1}{2}m \left\| \mathbf{v}(b) \right\|^2$$

so that $E_{\text{tot}} = U + T$ is constant along physical paths (solutions to the equations of motion)

4 Conservation of Total Linear Momentum.

Let us consider the consequences of an invariance of a Lagrangian under choice of center of coordinates. Call it **shift invariance**. That is, we suppose that when expressed in an inertial frame with rectangular coordinates, our Lagrangian satisfies

$$L(t, \mathbf{x} + \mathbf{s}, \mathbf{v}) = L(t, \mathbf{x}, \mathbf{v})$$

for all $\mathbf{s} \in \mathbb{R}^{3N}$ of the form $\mathbf{s} = (\mathbf{a}, ..., \mathbf{a})$. Then we have

$$\begin{aligned} 0 &= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} L(t, \mathbf{c}\,(t) + \varepsilon \mathbf{s}, \dot{\mathbf{c}}\,(t)) \\ &= \sum_a \frac{\partial L}{\partial \mathbf{x}_a}(t, \mathbf{c}\,(t)\,, \dot{\mathbf{c}}\,(t)) \cdot \mathbf{a} \\ &= \mathbf{a} \cdot \sum_a \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{x}}_a} = \mathbf{a} \cdot \left(\frac{d}{dt} \sum_a \frac{\partial L}{\partial \dot{\mathbf{x}}_a} \right) \end{aligned}$$

and since $\mathbf{a}$ is arbitrary we have $\frac{d}{dt}\sum_a \frac{\partial L}{\partial \dot{\mathbf{x}}_a} = 0$. So the vector quantity $\mathbf{P} = \sum_a \frac{\partial L}{\partial \dot{\mathbf{x}}_a}$ is conserved along solution curves shift invariant systems (such as closed systems):

$$\frac{d}{dt}\mathbf{P} = 0$$

In general, $\mathbf{P}$ is called the **total momentum** and $\mathbf{p}_a := \frac{\partial L}{\partial \dot{\mathbf{x}}_a}$ is the **momentum** of the the a-th particle.

$$\mathbf{P} = \sum_a \mathbf{p}_a$$

The Lagrangian $L = \sum \frac{1}{2}m_a \left\| \dot{\mathbf{x}}_a \right\|^2 - U(\mathbf{x})$ has this shift invariance and so since $\mathbf{p}_a = m_a \dot{\mathbf{x}}_a$ we have the conserved total momentum

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

Also, for a closed system we have

$$\sum_a \mathbf{F}_a = \frac{d}{dt}\mathbf{P} = \mathbf{0}$$

A common form for a shift invariant Lagrangian is the following:

$$\begin{aligned} L(\mathbf{x}, \dot{\mathbf{x}}) &= T(\dot{\mathbf{x}}_1, ...\dot{\mathbf{x}}_N) - \sum_{j<k} U_{jk}(\mathbf{x}_j, \mathbf{x}_k) \\ &= T(\dot{\mathbf{x}}_1, ...\dot{\mathbf{x}}_N) - \sum_{j<k} f_{jk}(\|\mathbf{x}_j - \mathbf{x}_k\|) \end{aligned}$$

where $f_{jk} = f_{kj}$. For example for gravitationally interacting bodies we have

$$U_{ab}(\mathbf{x}_a, \mathbf{x}_b) = -G\frac{m_a m_b}{\|\mathbf{x}_a - \mathbf{x}_b\|}$$

Then it can be checked that

$$\mathbf{F}_a = \sum_{\substack{b=1 \\ b\neq a}}^{N} \mathbf{F}_{ab}$$

where

$$\begin{aligned} \mathbf{F}_a(\mathbf{x}) &= \frac{\partial U}{\partial \mathbf{x}_a} = \frac{\partial}{\partial \mathbf{x}_a}\sum_{j<k} U_{jk}(\mathbf{x}_j, \mathbf{x}_k) \\ &= \frac{\partial}{\partial \mathbf{x}_a}\left[\sum_{\substack{j \\ j<a}} U_{ja}(\mathbf{x}_j, \mathbf{x}_a) + \sum_{\substack{j \\ a<j}} U_{aj}(\mathbf{x}_a, \mathbf{x}_j)\right] \\ &= \sum_{j\neq a} \frac{\partial}{\partial \mathbf{x}_a} U_{aj}(\mathbf{x}_a, \mathbf{x}_j) = \sum_{j\neq a} \mathbf{F}_{aj}(\mathbf{x}_a, \mathbf{x}_j) \end{aligned}$$

Note well that momentum is not a Galilean invariant in that is depends on the frame.

However, once a frame is chosen, then the total momentum vector for a closed system (or spatial shift invariant system) is constant with respect to time.

If the system is not closed then the appropriae Lagrangian often has the form

$$L(\mathbf{x}, \dot{\mathbf{x}}) = T(\dot{\mathbf{x}}_1, ...\dot{\mathbf{x}}_N) - \sum_a U_a(\mathbf{x}_a) - \sum_{j<k} U_{jk}(\mathbf{x}_j, \mathbf{x}_k)$$

in which case we have

$$\mathbf{F}_a(\mathbf{x}) = \mathbf{F}_a^{\text{ext}}(\mathbf{x}_a) + \sum_{j\neq a} \mathbf{F}_{aj}(\mathbf{x}_a, \mathbf{x}_j)$$

where $\mathbf{F}_a^{\rm ext}$ is the external force on the a-th particle. Note that if $\mathbf{F}_{jk} = -\mathbf{F}_{kj}$ as Newton supposed, then

$$\mathbf{F} = \sum \mathbf{F}_a = \sum \mathbf{F}_a^{\rm ext}(\mathbf{x}_a) = \mathbf{F}^{\rm ext}$$

where

$$\mathbf{F}_a^{\rm ext}(\mathbf{x}_a) := -\frac{\partial U_a}{\partial \mathbf{x}_a}(\mathbf{x}_a)$$

On the other hand,

$$\mathbf{F}^{\rm ext} = \mathbf{F} = \sum \mathbf{F}_a = \frac{d}{dt}\sum m_a \dot{\mathbf{x}}_a = \frac{d}{dt}\sum_a \mathbf{p}_a = \frac{d}{dt}\mathbf{P}$$

Thus for a particle system in an external force

$$\frac{d}{dt}\mathbf{P} = \mathbf{F}^{\rm ext} = \mathbf{F}$$

4.1 Center of Mass

If a frame K' moves with respect to a frame K with velocity $\mathbf{V}$ then the velocities of the particles are related in the two frames:

$$\mathbf{v}_a = \mathbf{v}'_a + \mathbf{V}.$$

The total momenta in the two frames are related by

$$\begin{aligned} \mathbf{P} &= \sum m_a \mathbf{v}_a = \sum m_a \mathbf{v}'_a + \mathbf{V}\sum m_a \\ \mathbf{P} &= \mathbf{P}' + \mathbf{V}\sum m_a \end{aligned}$$

Now we see it is possible to choose a frame so that $\mathbf{P}' = 0$. Indeed we see that $\mathbf{P}' = 0$ gives

$$\begin{aligned} \mathbf{P} &= \mathbf{V}\sum m_a \\ \mathbf{V} &= \mathbf{P}/M \\ \text{where } M &:= \sum m_a \text{ is the } \textbf{total mass} \end{aligned}$$

Now we have something interesting. First define the **center of mass** as

$$\mathbf{R} := \frac{1}{M}\sum m_a \mathbf{x}_a$$

then

$$\dot{\mathbf{R}} := \frac{1}{M}\sum m_a \dot{\mathbf{x}}_a = \frac{1}{M}\mathbf{P}$$

For a closed system $\mathbf{V} = \dot{\mathbf{R}}$ and $M\ddot{\mathbf{R}} = \mathbf{0}$. But if there *are* external forces then the system is not closed and we have

$$\begin{aligned} M\ddot{\mathbf{R}} &= \frac{d}{dt}\mathbf{P} = \mathbf{F}^{\text{ext}} = \mathbf{F} \\ \mathbf{F} &= M\ddot{\mathbf{R}} \text{ (Newton's law again)} \end{aligned}$$

It is time to notice that the total energy of a system, like total momentum, *depends on the frame.*

Next let us consider a closed system where $U(\mathbf{x})$ is shift invariant. Under a boost $\mathbf{x} = \mathbf{x}'+t\mathbf{V} + \mathbf{a}$ the potential for such a system satisfies $U(\mathbf{x}) = U'(\mathbf{x}', t) = U(\mathbf{x}')$. Then

$$\begin{aligned} E &= \frac{1}{2}\sum m_a \|\mathbf{v}_a\|^2 + U(\mathbf{x}) \\ &= \frac{1}{2}\sum m_a \|\mathbf{v}'_a + \mathbf{V}\|_a^2 + U(\mathbf{x}) \\ &= \frac{1}{2}M\|\mathbf{V}\|^2 + \mathbf{V}\cdot\sum m_a \mathbf{v}'_a + \frac{1}{2}\sum m_a \|\mathbf{v}'_a\|^2 + U \\ &= \frac{1}{2}\sum m_a \|\mathbf{v}'_a\|^2 + U'(\mathbf{x}') + \mathbf{V}\cdot\mathbf{P}' + \frac{1}{2}M\|\mathbf{V}\|^2 \\ &= E' + \mathbf{V}\cdot\mathbf{P}' + \frac{1}{2}M\|\mathbf{V}\|^2 \end{aligned}$$

Now notice that if the center of mass is at rest in K' then $\mathbf{P}' = 0$ and by definition we have the **internal energy** $E_{\text{int}} := E'$ Then we have

$$E = E_{\text{int}} + \frac{1}{2}M\|\mathbf{V}\|^2$$

5 Angular Momentum

Suppose that $\phi \mapsto R_{\phi\mathbf{u}}$ is a **one parameter group of rotations** about a vector in the direction of the **unit vector u**. Then

$$\left.\frac{d}{d\phi}\right|_{\phi=0} R_{\phi\mathbf{u}} = L_{\mathbf{u}}$$

where

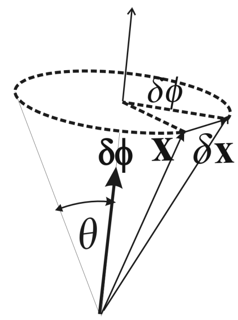
$$L_{\mathbf{u}}(\mathbf{v}) = \mathbf{u}\times\mathbf{v}$$

and we have

$$R_{\delta\phi\mathbf{u}} = \exp(\delta\phi L_{\mathbf{u}}) = 1 + \delta\phi L_{\mathbf{u}} + O(\,|\delta\phi|^2\,)$$

Lets use the notation $\boldsymbol{\delta\phi} = \delta\phi\mathbf{u}$. Then

$$R_{\boldsymbol{\delta\phi}}\mathbf{x} = \mathbf{x} + \boldsymbol{\delta\phi}\times\mathbf{x} + O(\,\|\boldsymbol{\delta\phi}\|^2\,)$$



$$\|\delta\mathbf{x}\| = \|\mathbf{x}\|\delta\phi\sin\theta$$

$$\delta\mathbf{x} = (\delta\phi\mathbf{u}) \times \mathbf{x}$$

$$\delta\mathbf{x} = \boldsymbol{\delta\phi} \times \mathbf{x}$$

Suppose that our Lagrangian is invariant under rotations. Then using the E-L equations of motion we have

$$\begin{aligned}
0 &= \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} L(R_{\epsilon\mathbf{u}}\mathbf{x}, R_{\boldsymbol{\epsilon}\mathbf{u}}\dot{\mathbf{x}}) \\
&= \sum_a \left(\frac{\partial L}{\partial \mathbf{x}_a} \cdot \mathbf{u} \times \mathbf{x}_a + \frac{\partial L}{\partial \dot{\mathbf{x}}_\mathbf{a}} \cdot \mathbf{u} \times \dot{\mathbf{x}}_\mathbf{a} \right) \\
&= \mathbf{u} \cdot \sum_a (\mathbf{x}_a \times \dot{\mathbf{p}}_a + \dot{\mathbf{x}}_\mathbf{a} \times \mathbf{p}_a) \\
&= \mathbf{u} \cdot \frac{d}{dt} \sum_a \mathbf{x}_a \times \mathbf{p}_a
\end{aligned}$$

Since this is true for all possible unit vectors $\mathbf{u}$, we obtain

$$\frac{d}{dt} \sum_a \mathbf{x}_a \times \mathbf{p}_a = 0$$

If we define the total angular momentum by

$$\mathbf{L} := \sum_a \mathbf{x}_a \times \mathbf{p}_a$$

then we have another conservation law for our system with rotational invariance (as for closed system for example).

$$\frac{d}{dt}\mathbf{L} = 0$$

While we are at it, let's let $\mathbf{L}_a := \mathbf{x}_a \times \mathbf{p}_a$ be the **angular momentum of the a-th particle.** Then we have

$$\mathbf{L} := \sum_a \mathbf{L}_a$$

But notice that for **any system** of particles in space

$$\begin{aligned} \frac{d}{dt}\mathbf{L}_a &= \frac{d}{dt}\mathbf{x}_a \times \mathbf{p}_a = \mathbf{x}_a \times \frac{d}{dt}\mathbf{p}_a + \frac{d}{dt}\mathbf{x}_a \times \mathbf{p}_a \\ &= \mathbf{x}_a \times \mathbf{F}_a := \boldsymbol{\tau}_a \text{ (\textbf{Torque} on } a\text{-th particle)} \end{aligned}$$

and then

$$\begin{aligned} \frac{d}{dt}\mathbf{L}_a &= \boldsymbol{\tau}_a \\ \frac{d}{dt}\mathbf{L} &= \boldsymbol{\tau} \\ \text{where } \boldsymbol{\tau} &:= \sum_a \boldsymbol{\tau}_a \text{ (\textbf{total torque})} \end{aligned}$$

Note that as defined here, $\mathbf{L}_a, \boldsymbol{\tau}_a$ and $\mathbf{L}, \boldsymbol{\tau}$ all depend on the choice of origin.

Let us now consider a *closed system* again and consider any two inertial frames K and K' with $\mathbf{x}_a = \mathbf{x}'_a + \mathbf{r}_0$. Then

$$\begin{aligned} \mathbf{L} &= \sum_a \mathbf{x}_a \times \mathbf{p}_a \\ &= \sum_a \mathbf{x}'_a \times \mathbf{p}_a + \mathbf{r}_0 \times \sum_a \mathbf{p}_a \\ &= \sum_a \mathbf{x}'_a \times \mathbf{p}_a + \mathbf{r}_0 \times \mathbf{P} \\ &= \mathbf{L}' + \mathbf{r}_0 \times \mathbf{P} \end{aligned}$$

However, if we are in a frame where the system as a whole is at rest, $\mathbf{P} = \mathbf{0}$, then we have a special situation that $\mathbf{L} = \mathbf{L}'$.

Now what if we have a system which is symmetric about a single axis (say the z axis) ?

The angular momentum about the z axis is given by

$$l_z = \sum_a \frac{\partial L}{\partial \dot{\phi}_a}$$

where $\mathbf{L} = (l_x, l_y, l_z)$. In fact, using cylindrical coordinates

$x_a = r_a \cos\phi_a$, $y_a = r_a \sin\phi_a$

$$\begin{aligned} l_z &= \sum_a m_a(x_a \dot{y}_a - y_a \dot{x}_a) \\ &= \sum_a m_a r_a^2 \dot{\phi}_a \end{aligned}$$

But

$$L=\frac{1}{2}\sum_a m_a\left(\dot{r}_a^2+r_a^2\dot{\phi}_a^2+\dot{z}_a^2\right)-U$$

and substitution gives the result.

A puzzle for you about boosts:

Free particle.

Use rectangular coordinates

$$L=\frac{1}{2}m\left\|\dot{\mathbf{x}}\right\|^2$$

The Euler-Lagrange equation become

$$m\ddot{\mathbf{x}}=\mathbf{0}$$

Note that under $\dot{\mathbf{x}}\to\dot{\mathbf{x}}+\mathbf{V}$ we can define a new Lagrangian

$$\begin{aligned}\widetilde{L}(\dot{\mathbf{x}}) &= L(\dot{\mathbf{x}}+\mathbf{V})=\frac{1}{2}m\left\|\dot{\mathbf{x}}+\mathbf{V}\right\|^2\\ &= \frac{1}{2}m\left\|\dot{\mathbf{x}}\right\|^2+m\dot{\mathbf{x}}\cdot\mathbf{V}+\left\|\mathbf{V}\right\|^2\end{aligned}$$

Then

$$\begin{aligned}\widetilde{L}(\dot{\mathbf{x}}(t)) &= L(\dot{\mathbf{x}}(t))+\frac{d}{dt}\left(m\mathbf{x}(t)\cdot\mathbf{V}+t\left\|\mathbf{V}\right\|^2\right)\\ &= L(\dot{\mathbf{x}}(t))+\frac{d}{dt}f(\mathbf{x}(t),t)\end{aligned}$$

So $\widetilde{L}$ and L give the Same Euler-Lagrange equations. How are we going to get a conservation law from considering boosts of closed systems?

6 Integration of the equations of motion

6.1 One dimension

$$L=\frac{1}{2}a(q)\dot{q}^2-U(q)$$

If motion is along a straight line and if x denotes the standard linear coordinate on $\mathbb{R}$ then

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

The equations of motion are the E-L equations. But, if we know the total energy E we can use energy conservation

$$\begin{aligned}
E &= \frac{1}{2}m\dot{x}^2 + U(x) \\
\dot{x} &= \sqrt{\frac{2}{m}}\left(E - U(x)\right)^{1/2} \\
t &= \int dt = \int \frac{m dx}{2\left(E - U(x)\right)^{1/2}} + C \\
\int_{t_0}^{t} dt &= \sqrt{\frac{m}{2}}\int_{x(t_0)}^{x(t)} \frac{d\xi}{\left(E - U(\xi)\right)^{1/2}} \\
t &= t_0 + \sqrt{\frac{m}{2}}\int_{x(t_0)}^{x(t)} d\xi \frac{1}{\left(E - U(\xi)\right)^{1/2}}
\end{aligned}$$

For example if $E = 1$ and $U(x) = x^2$ then

$$\begin{aligned}
t &= \sqrt{\frac{m}{2}}\int_0^x \left(\frac{1}{\left(1 - \xi^2\right)^{1/2}}\right) d\xi = \sqrt{\frac{m}{2}} \arcsin x \\
x &= \sin\left(\sqrt{\frac{2}{m}} t\right)
\end{aligned}$$

Now when $E = U(x)$ we have a turning point. Let's say that $x(0)$ is in a region between adjacent turning points x_1 and x_2. The the motion will oscillate between these points with period

$$T = \sqrt{2m}\int_{x(t_0)}^{x(t)} \frac{d\xi}{\left(E - U(\xi)\right)^{1/2}}$$

For our little example, $T = \sqrt{2m}\int_{-1}^{1} \frac{d\xi}{\left(1-\xi^2\right)^{1/2}} = \sqrt{2m}\pi$.

Exercise
$T = \sqrt{2m}\int_0^{\pi} \frac{1}{(1-\cos\xi)^{1/2}} d\xi =$???
(use physics!)

Preview: Here are some topics coming up:
Central force fields, The Kepler problem
Vibrations
Constrained motion (return of the Lagrangian)
Hamiltonians, Brackets, etc.
Lie groups and Momentum maps

Example of systems which can be described by a Lagrangian (vibrations etc.)

Rigid bodies and Euler's Equations

Hamiltonians, symplectic geometry and Poisson brackets etc.

Relativity and relativistic mechanics.

etc.

7 Appendix I

Galilean Spacetime

8 Galilean Spacetime

Roughly speaking an affine space is a vector space that has forgotten its zero element (no origin).

An **affine space** A with associated vector space V is a set on which V acts from the right (or left)

$$\begin{aligned} A \times V &\to A \\ (x, \mathbf{v}) &\mapsto x + \mathbf{v} \end{aligned}$$

such that

- $x + \mathbf{0} = p$
- $(x + \mathbf{v}_1) + \mathbf{v}_2 = x + (\mathbf{v}_1 + \mathbf{v}_2)$
- Given $x_1, x_2 \in A$ there is a unique element $\mathbf{v}$ of V denoted by $x_2 - x_1$ such that $x_1 + \mathbf{v} = x_2$.

Suppose that A_1 and A_2 are affine spaces modelled on V_1 and V_2 resp. Then a map $\alpha : A_1 \to A_2$ is called an affine map if there is a linear map $L : V_1 \to V_2$ such that

$$\begin{aligned} \alpha(x) - \alpha(x_0) &= \alpha_L(x - x_0) \\ \text{for all } p, p_0 &\in A_1 \end{aligned}$$

The set of affine isomorphisms $A \to A$ is the affine group $\mathrm{Aut}(A)$

If V has a positive definite inner product $\langle \cdot, \cdot \rangle$ then we can define the distance $d(x_1, x_2) = \|x_2 - x_1\|$. We then say that A is a **Euclidean affine space**. In this case an affine map α is called a Euclidean affine map if

$$\|\mathbf{v}\| = \|\alpha_L(\mathbf{v})\| \text{ for all } \mathbf{v} \in V_1$$

In particular we have the group of Euclidean isometries $\mathrm{Euc}(A) \subset \mathrm{Aut}(A)$.

A Galilean space time is a 4-dimensional affine space M^4_{Gal} together with the following

- A nonzero linear form $\tau : V^4 \to \mathbb{R}$
- An inner product defined on the kernel $S = \tau^{-1}(0)$

  $S =$"space"

  Point of M^4_{Gal} are called **events** (Galilean), τ is the **absolute time** function and $\tau(x_2 - x_1)$ is called the the time interval from x_1 to x_2. Two events are said to be simultaneous if $\tau(x_2 - x_1) = 0$.

For each event $x \in A^4$ the set $S_x = x + S$ is the set of all events simultaneous with x.

Now the group $\text{Aut}(M^4_{\text{Gal}})$ of affine automorphisms that preserve the Galilean structure is the Galilean group

$$G_{\text{Gal}} := \text{Aut}(M^4_{\text{Gal}})$$

By definition for $T \in G_{\text{Gal}}$ we have

- $\tau(x_2 - x_1) = \tau(Tx_2 - Tx_1)$
- $\tau(x_2 - x_1) = 0 \Longrightarrow \|Tx_2 - Tx_1\| = \|x_2 - x_1\|$

One may show that this is equivalent to

- $\tau = \tau \circ T^*$
- $T^*|_S$ is a linear isometry of S.

Now if $\mathbb{E}^3$ is Euclidean affine space modeled on an inner product space V^3 then $\mathbb{R} \times \mathbb{E}^3$ is a Galilean spacetime modeled on $\mathbb{R} \times V^3$ where

$$t(a, \mathbf{x}) = a$$

is the time function.

Now if M^4_{Gal} is a Galilean spacetime then a Galilean isomorphism $\phi : M^4_{\text{Gal}} \to \mathbb{R} \times S$ is called a an **inertial frame** and by choosing an orthonormal basis for S we obtain a map $u : \mathbb{R}^3 \to S$ and then Galilean isomorphism $\psi := u \circ \phi : M^4_{\text{Gal}} \to \mathbb{R} \times \mathbb{R}^3$ which we call an **inertial coordinate system**.

If we fix an ON basis once and for all then

Note that if $g \in G_{\mathbb{R}\times S}$ then $g \circ \phi$ is also an inertial frame. Similarly if $g \in G_{\mathbb{R}\times\mathbb{R}^3}$ and ψ an inertial coordinate system then $g \circ \phi$ is also.

Depending on the problem is may be convenient to choose a fixed frame or a fixed inertial coordinate system so that any of the following Galilean isomorphisms may be taken as an identification

$$\begin{aligned} M^4_{\text{Gal}} &\simeq \mathbb{E}^1 \times S \simeq \mathbb{E}^1 \times \mathbb{E}^3 \simeq \mathbb{R} \times \mathbb{R}^3 \\ S &\simeq \mathbb{E}^1 \simeq \mathbb{R}^3 (\text{Euclidean isometry}) \end{aligned}$$

A world line is a map of the form $c : I \to A^4$ such that

$$\begin{aligned} \tau(c(t) - c(t_0)) &= t - t_0 \\ t, t_0 &\in I \end{aligned}$$

In an inertial coordinate system this means that $c(t) := (t, \mathbf{x}(t))$ for some map $\mathbf{x} : I \to \mathbf{R}^3$.

9 Appendix II (Variational calculus)

Recall the basic definitions of the directional derivative of a map such as

$$f : \mathbb{R}^n \to \mathbb{R}^m.$$

From one point of view a n-tuple is just a function whose domain is the finite set $\{1, 2, ..., n\}$.
$h = (h^1, ..., h^n)$
is just the function $i \mapsto h^i$ which may as well have been written $i \mapsto h(i)$.
This suggests that we generalize to functions whose domain is an infinite set.
This brings us to the notion of a function space. An example of a function space is $C([0, 1])$, the space of continuous functions on the unit interval $[0, 1]$.

So one approach to generalizing the usual setting of calculus might be to consider replacing the space of n-tuples $\mathbb{R}^n$ by a space of functions then consider functions on that space. We call these **functionals**

$$f \mapsto F[f].$$

We shall follow the tradition of writing $F[f]$ instead of $F(f)$. Some books even write $F[f(x)]$. Notice that this is *not* a composition of functions. A simple example of a functional on $C([0, 1])$ is

$$F[f] = \int_0^1 f^2(x)dx.$$

We may then easily define a formal notion of **directional derivative**:

$$(D_h F)[f] = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \left(F[f + \epsilon h] - F[f]\right)$$

where h is some function which is the "direction vector".

This also allows us to define the **differential** δF which is a linear map on the functions space given at f by $\delta F|_f h = (D_h F)[f]$. This differential is traditionally called the **variation**. We use a δ instead of a d to avoid confusion between dx^i and δx^i (the latter is a function of t). Also, the test function h is sometimes denoted δf and also called **a variation**.

In general choosing the right function space for a particular problem nontrivial and in each case the function space must be given an appropriate topology.

In the following few paragraphs our discussion will be informal and we shall be rather cavalier with regard to the issues just mentioned.

The following is another typical example of a functional defined on the space $C^1([0,1])$ of continuously differentiable functions defined on the interval $[0,1]$:

$$S[c] := \int_0^1 \sqrt{1 + (dc/dt)^2} dt$$

The reader may recognize this example as the arc length functional.

The derivative *at* the function c in the *direction of* a function $h \in C^1([0,1])$ would be given by

$$\delta S|_c (h) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} (S[c + \varepsilon h] - S[c]) .$$

It is well known that if $\delta S|_c (h) = 0$ for every h then c is a linear function; $c(t) = at + b$. The condition $\delta S|_c (h) = 0$ (for all h) is often simply written as $\delta S = 0$.

For examples like this one, the analogy with multi-variable calculus is summarized as

$$\begin{aligned} i &\rightsquigarrow t \\ x^i &\rightsquigarrow c(t) \\ f(\vec{x}) &\rightsquigarrow S[c] \end{aligned}$$

Here we move from d-tuples (which are really functions with finite domain) to functions with a continuous domain. The function f of x becomes a functional S of functions c.

$L : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ which will be the basic ingredient in building an associated functional. A typical example is of the form $L(x, v) = \frac{1}{2}mv^2 - V(x)$. Define the action functional S by using L as follows: For a given function $t \longmapsto q(t)$ defined on $[a, b]$ let

$$S[q] := \int_a^b L(q(t), \dot{q}(t))dt.$$

We have used x and v to denote variables of L but since we are eventually to plug in $q(t), \dot{q}(t)$ we could also follow the common tradition of denoting these variables by q and $\dot{q}$

$$\delta S = \int \frac{\delta S}{\delta q(t)} \delta q(t) dt \tag{1}$$

Depending on one's training and temperament, the meaning of the notation may be a bit hard to pin down. First, what is the meaning of δq as opposed to, say, the differential dq? Second, what is the mysterious $\frac{\delta S}{\delta q(t)}$?

A good start might be to go back and settle on what we mean by the differential in ordinary multivariable calculus. For a differentiable function $f : \mathbb{R}^d \to \mathbb{R}$ we take df to just mean the map

$$df : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$$

given by $df(p,h) = f'(p)h$. We may also fix p and write $df\big|_p$ or $df(p)$ for the linear map $h \mapsto df(p,h)$. With this convention we note that $dx^i\big|_p(h) = h^i$ where $h = (h^1, ..., h^d)$. Thus applying both sides of the equation

$$df\big|_p = \sum \frac{\partial f}{\partial x^i}(p)\, dx^i\big|_p \tag{2}$$

to some vector h we get

$$f'(p)h = \sum \frac{\partial f}{\partial x^i}(p)h^i. \tag{3}$$

In other words, $df\big|_p h = D_h f(p) = \nabla f \cdot h = f'(p)(h)$.

Too many notations for the same concept.

The equation $df(h) = \sum \frac{\partial f}{\partial x^i} h^i$ is clearly very similar to $\delta S(\delta q) = \int \frac{\delta S}{\delta q(t)} \delta q(t) dt$ and so we expect that δS is a linear map and that $t \mapsto \frac{\delta S}{\delta q(t)}$ is to δS as $\frac{\partial f}{\partial x^i}$ is to df:

$$\begin{aligned} df &\rightsquigarrow \delta S \\ \frac{\partial f}{\partial x^i} &\rightsquigarrow \frac{\delta S}{\delta q(t)}. \end{aligned}$$

Roughly, $\frac{\delta S}{\delta q(t)}$ is taken to be whatever function (or distribution) makes the equation 1 true. We often see the following type of calculation

$$\begin{aligned} \delta S &= \delta \int L\, dt \\ &= \int \left(\frac{\partial L}{\partial q}\delta q + \frac{\partial L}{\partial \dot{q}}\delta \dot{q}\right) dt \\ &= \int \left\{\frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}\right\}\delta q\, dt \end{aligned} \tag{4}$$

from which we are to conclude that

$$\frac{\delta S}{\delta q(t)} = \frac{\partial L}{\partial q} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}$$

Actually, there is a subtle point here in that we must restrict δS to variations for which the integration by parts is justified.

....

We consider the space $A = C^2_{t_1,p_1;t_2,p_2}$ of all maps $\mathbf{x} : [t_1,t_2] \to \mathbf{R}^n$ such that $\mathbf{x}(t_i) = p_i$ as an affine space with associated vector space V

$$V = C^2_{t_1,0;t_2,0} := \{\mathbf{h} \in C^2([t_1,t_2], \mathbf{R}^n) : \mathbf{h}(t_1) = \mathbf{h}(t_2) = 0\}$$

$C^2_{t_1,p_1;t_2,p_2}$ is a subspace of the Banach space $C^2([t_1,t_2], \mathbf{R}^n)$. Here we take the norm on $C^k([t_1,t_2], \mathbf{R}^n)$ to be

$$\|\mathbf{f}\| = \max_{0 \le i \le k} \{\|D^i \mathbf{f}\|_\infty\}$$

Let us consider a Lagrangian

$$L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

and the functional on $C^2_{t_1,p_1;t_2,p_2}$ given by

$$S[\mathbf{c}] := \int_{t_1}^{t_2} L(\mathbf{c}(t), \dot{\mathbf{c}}(t))\, dt$$

One main example will be the case of a single particle of mass m in a potential V. We choose an inertial coordinate system. Then $n = 3$ and $L : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ will be of the form

$$\begin{aligned} L(\mathbf{x}, \mathbf{v}) &= \frac{1}{2} m \|\mathbf{v}\|^2 - U(\mathbf{x}) \\ L(\mathbf{c}(t), \dot{\mathbf{c}}(t)) &= \frac{1}{2} m \|\dot{\mathbf{c}}(t)\|^2 - U(\mathbf{c}(t)) \\ S[\mathbf{c}] &: = \int_{t_1}^{t_2} \left[\frac{1}{2} m \|\dot{\mathbf{c}}(t)\|^2 - U(\mathbf{c}(t))\right] dt \end{aligned}$$

In general for $\mathbf{h} \in C^2_{t_1,0;t_2,0}$ we have

$$\delta S_{\mathbf{c}(.)}(\mathbf{h}) = \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial \mathbf{x}}(\mathbf{c}(t), \dot{\mathbf{c}}(t)) \cdot \mathbf{h}(t) + \frac{\partial L}{\partial \mathbf{v}}(\mathbf{c}(t), \dot{\mathbf{c}}(t)) \cdot \dot{\mathbf{h}}(t)\right] dt$$

Integration by parts gives

$$\begin{aligned} \delta S_{\mathbf{c}(.)}(\mathbf{h}) &= \int_{t_1}^{t_2} \left[\frac{\partial L}{\partial \mathbf{x}}(\mathbf{c}(t), \dot{\mathbf{c}}(t)) - \frac{d}{dt}\frac{\partial L}{\partial \mathbf{v}}(\mathbf{c}(t), \dot{\mathbf{c}}(t))\right] \cdot \mathbf{h}(t)\, dt \\ &\quad \text{(boundary terms vanish)} \end{aligned}$$

Now since this is true for all $\mathbf{h}$ we conclude that

$$\frac{\partial L}{\partial \mathbf{x}}(\mathbf{c}(t), \dot{\mathbf{c}}(t)) - \frac{d}{dt}\frac{\partial L}{\partial \mathbf{v}}(\mathbf{c}(t), \dot{\mathbf{c}}(t)) = 0 \text{ for all } t$$

$$L(\mathbf{x},\mathbf{v}) = \frac{1}{2}m\sum v_i^2 - U(x_1,x_2,x_3)$$

so that $\frac{\partial L}{\partial v_i} = mv_i$ and $\frac{\partial L}{\partial x_i} = -\frac{\partial V}{\partial x_i}$ so

$$\begin{aligned}
-\frac{\partial V}{\partial x_i}(c_1(t),...,c_1(t)) - \frac{d}{dt}m\dot{c}_i(t) &= 0 \\
m\ddot{c}_i &= -\frac{\partial U}{\partial x_i}(c_1(t),...,c_1(t)) \\
m\ddot{\mathbf{c}} &= -\frac{\partial U}{\partial \mathbf{x}} \circ \mathbf{c} \\
m\mathbf{a} &= F
\end{aligned}$$

Let us abuse notation:
so that

$$\begin{aligned}
x_i &= x_i(t) \\
\dot{x}_i &= \dot{x}_i(t) = \frac{dx_i}{dt}(t)
\end{aligned}$$

Now suppose we rewrite our problem in general coordinates $(q_1,...,q_n)$ and let $q_i = q_i(t)$ so that q_i also describes the name of a functions. Now even more clever but also confusing, let $(\dot{q}_1,...,\dot{q}_n)$ be the names of n-more coordinates. But also let us agree that

$$\dot{q}_i = \dot{q}_i(t) = \frac{dq_i}{dt}(t)$$

in any context where $q_i = q_i(t)$. Thus

$$L(t,q,\dot{q}_i)$$

may simply refer to $L : \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$ evaluated at $(t,q,\dot{q}_i) \in \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^n$ or it may be an abbreviation for

$$L(t) = L(t,q(t),\dot{q}_i(t)).$$

In what follows we must be careful to understand

$$L, \quad \frac{\partial L}{\partial q_i}, \quad \frac{\partial L}{\partial \dot{q}_i}, \quad \frac{d}{dt}\frac{\partial L}{\partial q_i} \text{ etc.}$$

Lets do the derivation again in this more general setting and using the above abuses of notation:

$$\begin{aligned}
\delta S|_{q(.)}\, h &:= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}(S[q+\varepsilon h] - S[q]) \\
&= \int (\frac{\partial L}{\partial q_i}h_i + \frac{\partial L}{\partial \dot{q}_i}\dot{h}_i)\, dt
\end{aligned}$$

We restrict attention to "directions" h for which $h(a) = h(b) = 0$ and use integration by parts to obtain

$$\delta S|_{q(.)}\, h = \int \left(\frac{\partial L}{\partial q^i}(q(t), \dot{q}(t)) - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}(q(t), \dot{q}(t)) \right) h^i(t)\, dt.$$

since this is true for all h which vanish at the end times, we conclude

$$\frac{\partial L}{\partial q_i}(q(t), \dot{q}(t)) - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i}(q(t), \dot{q}(t)) \equiv 0$$

or

$$\begin{aligned} \frac{\partial L}{\partial q_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} &= 0 \text{ for } 1 \le i \le n \\ \nabla_q L - \frac{d}{dt}\nabla_{\dot{q}} L &= 0 \end{aligned}$$

Exercise 2 *Replace* $S[c] = \int L(c(t), \dot{c}(t))dt$ *by the similar function of several variables*

$$S(c_1, ...c_N) = \sum L(c_i, \triangle c_i).$$

Here $\triangle c_i := c_i - c_{i-1}$ *(taking* $c_0 = c_N$*) and* L *is a differentiable map* $\mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$. *What assumptions on* $c = (c_1, ...c_N)$ *and* $h = (h_1, ...h_N)$ *justify the following calculation?*

$$\begin{aligned} dS|_{(c_1,...c_N)}\, h &= \sum \frac{\partial L}{\partial c_i} h^i + \frac{\partial L}{\partial \triangle c_i} \triangle h^i \\ &= \sum \frac{\partial L}{\partial c_i} h^i + \sum \frac{\partial L}{\partial \triangle c_i} h^i - \sum \frac{\partial L}{\partial \triangle c_i} h^{i-1} \\ &= \sum \frac{\partial L}{\partial c_i} h^i + \sum \frac{\partial L}{\partial \triangle c_i} h^i - \sum \frac{\partial L}{\partial \triangle c_{i+1}} h^i \\ &= \sum \frac{\partial L}{\partial c_i} h^i - \sum (\frac{\partial L}{\partial \triangle c_{i+1}} - \frac{\partial L}{\partial \triangle c_i}) h^i \\ &= \sum \{ \frac{\partial L}{\partial c_i} h^i - (\triangle \frac{\partial L}{\partial \triangle c_i}) \} h^i \end{aligned}$$

The upshot of our discussion is that the δ notation is just an alternative notation to refer to the differential or derivative in the setting of **differentiable functionals** on **function spaces**.

Example 3 *Let* $F[c] := \int_{[0,1]} c^2(t)dt$ *and let* $c(t) = t^3$ *and* $h(t) = \sin(t^4)$*. Then*

$$\begin{aligned}
\delta F|_c (h) &= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\left(F[c+\varepsilon h] - F[c]\right) \\
&= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} F[c+\varepsilon h] \\
&= \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} \int_{[0,1]} (c(t)) + \varepsilon h(t))^2 dt \\
&= 2\int_{[0,1]} c(t)h(t)dt = 2\int_0^1 t^3 \sin(\pi t^4)dx \\
&= \frac{1}{\pi}
\end{aligned}$$

One final important remark (related to cohomology).

Suppose that L and $\widetilde{L}$ are two Lagrangians. In the case that there exists a function $f(t,q)$ of time and space coordinates only such that

$$\widetilde{L}(t, q(t), \dot{q}(t)) - L(t, q(t), \dot{q}(t)) = \frac{d}{dt} f(t, q(t))$$

then

$$\begin{aligned}
\widetilde{S} - S &= \int_{t_1}^{t_2} \widetilde{L}(t, q(t), \dot{q}(t))\, dt - \int_{t_1}^{t_2} L(t, q(t), \dot{q}(t))\, dt \\
&= \int_{t_1}^{t_2} \frac{d}{dt} f(t, q(t))\, dt = f(t_2, q(t_2)) - f(t_1, q(t_1))
\end{aligned}$$

from which we see that for $h(t)$ vanishing at end times, variations $\delta\widetilde{S}(h) = 0$ if and only if $\delta S(h) = 0$. Thus the both lead to the same Euler-Lagrange equations.

Let $\mathbf{q} = (\mathbf{q}_1, ..., \mathbf{q}_N) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \cdots \times \mathbb{R}^3$

$$\frac{\partial L}{\partial \mathbf{q}_a}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) - \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{q}}_a}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) = 0$$