Hamilton's Principle

Ravitej Uppu

Introduction

In this talk, I will assume the *principle of virtual work* and related topics. The talk consists of introducing and later analyzing the *principle of least action* axiomatically and semi-rigorously.

Let's consider a holonomic[1] system of particles with 'n' degrees of freedom. It's evolution can be considered as a sequence of equilibrium states (à la d'Alembert) under the action of all the forces. On applying the principle of virual work, the equations of dynamics can be rewritten as

$$\frac{d}{dt}\frac{\partial \mathcal{T}}{\partial \dot{q_h}} - \frac{\partial \mathcal{T}}{\partial q_h} = \mathcal{Q}_h$$

where $\mathcal{Q}_h$ are the generalized forces. In our systems of interest, i.e. conservative systems, $\mathcal{Q}_h$ is expressed in term of the potential energy as

$$\mathcal{Q}_h = \frac{d}{dt}\frac{\partial \mathcal{U}}{\partial \dot{q_h}} - \frac{\partial \mathcal{U}}{\partial q_h} \tag{1}$$

This leads to the Euler-Lagrange equation of motion by the definition of the Lagrangian function $\mathcal{L} = \mathcal{T} - \mathcal{U}$ as

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q_h}} - \frac{\partial \mathcal{L}}{\partial q_h} = 0 \qquad h \in 1, \ldots, n$$

1 An axiomatic approach to the Hamilton's Principle

The state of a system is completely specified by specifying the *aptitude*[2] of the system.

The evolution is a sequence of such states. This sequence is completely defined by specifying the function $\mathcal{L}(q, \dot{q}, t)$ (which is defined on the set of states) and two different configurations q_A and q_B at t_A and t_B.

The Principle of Least Action says that among all the curves $q_h = q_h(t)$ joining the position A and B, the one for which the integral $S[q]$ is minimized[3] represents the path of evolution. The integral $S[q]$ is given as

$$S[q] = \int_{t_A}^{t_B} \mathcal{L}(q, \dot{q}, t)dt \tag{2}$$

[1]Refer Talk 1 : page 3

[2]Refer Talk 1: page 3

[3]The form of $S[q]$ holds on every small part of the trajectory. On the whole trajectory $S[q]$ may just have an extremum. This is sufficient to get to the equations of motion ($\because \delta S = 0$, for an extremum).

2 Introducing the concept of variation

- Let $u = u_0(x)$ extremize $I[u]$ among all the curves in the plane joining points A and B.
- Analyzing the properties of $I[u]$ implies we needs to compare the value of $I[u_0]$ taken by the integral on this curve with other curve v that has the value $I[v]$.
- The difference $v(x) - u_0(x)$ is the *variation* (for a fixed x) in passing from curve u_0 to v.
- Analogous to two close points being represented by dx, we represent the variation between two "close" curves as δu.
- Corresponding to δu, the variations that occur in $I[u]$ is $I[v] - I[u_0]$. This can be decomposed in terms of δI, $\delta^2 I, \ldots$
- The condition for $I[u]$ to take an extremum is $\delta I = 0$.

The next obvious extension that we would expect is identification of the type of extremum from the the sign of $\delta^2 I$. But, this is not correct. Because, unlike two points, two curves can be located arbitrarily close in space, yet at "close" points have tangents that are totally different. This effects even more when the integral is a function as u' which is precisely the case that we are dealing with here.

The clarification comes from the principles of the functional calculus which I will discuss in brief in the following section.

3 Building up functional calculus

3.1 Continuity

If we consider the obvious extension from calculus of one-variable to our case and say

> "Consider a functional $F[u]$. It is said to be continuous, if for a small change of $u(x)$, $F[u]$ undergoes a small change."
>
> But, this definition is very restricted as a generic functional $F[u]$ can also have an explicit dependance on u'. So the question is *How do we define "close curves"?*
>
> The definition should include the information of how close should our curves be. So, curves can be defined to be close to various orders as follows :
>
> - *Zero-order proximity close* : $u(x)$ and $v(x)$ are zero-order proximity close, if $|u(x)-v(x)|$ is small.
> - *First-order proximity close* : $u(x)$ and $v(x)$ are first-order proximity close, if $|u(x)-v(x)|$ and $|u'(x) - v'(x)|$ are small.
> - In general, $u(x)$ and $v(x)$ are *kth-order proximity close*, if $|u(x)-v(x)|, |u'(x)-v'(x)|, \ldots,$ $|u^{(k)}(x) - v^{(k)}(x)|$ are small.
>
> From this we can define a notion of distance (σ) between the two curves as follows
>
> Considering the existence of continuous derivatives of u and v upto order k, the distance of order k is given by
>
> $$\sigma_k(u, v) := \sum_{h=1}^{k} \max_{x_0 \le x \le x_1} |u^{(h)}(x) - v^{(h)}(x)|$$
>
> For close curves, σ_k is small.

3.2 Differentiability

Analogous to the total and partial derivatives that we know of, even here there exists two kinds of derivatives : Strong or the Frechet derivative and the Weak or the Gateaux derivative.

3.2.1 Strong derivative

Consider U and V (normed vector spaces). As a clarification to the question raised by Amaresh, U and V can be similar. Only possible distinguishing feature is the sense of distance that is considered on the space. For U, we define the distance as defined above. Whereas, the distance measure on V is the standard measure of separation between points.

A is an open subset in U. Consider a map F from A to U.

$$F : A \subseteq U \rightarrow V$$

F is said to be differentiable at $u \in U$, if $\exists\ F'$, such that

$$F[u+h] - F[u] = F'h + \sigma(u,h) \tag{3}$$

where, σ is infinitesimal with respect to the distance measure in V as given by the norm (i.e. $\lim_{||h|| \to 0} \frac{||\sigma(u,h)||}{||h||} = 0$).

$F'h$ is called the strong differential and F' is the strong derivative of F at u. Composition is straightforward. Uniqueness can be checked in the same way as total derivatives in calculus.

3.2.2 Weak derivative

Consider a map G that is defined as

$$G : A \subseteq U \rightarrow V$$

Consider the limit,

$$DG(u,h) = \frac{d}{dt} G[u+th]|_{t=0} = \lim_{t \to 0} \frac{G[u+th] - G[u]}{t} \tag{4}$$

The convergence is only with respect to the norm on V. This is the general definition for the weak derivative. In general DG can be non-linear with respect to h. In the case of DG being linear, we can write

$$DG(u,h) = G_u h$$

where G_u is the weak derivative of G at u.

Composition does not hold good for weak derivatives similar to the case of partial derivatives of functions. At this point few remarks will clarify the differences between the two kinds of derivatives.

- If strong derivative of F exists, weak derivative exists and they will coincide.

Proof : If F is strongly derivable, then

$$F[u+th]-F[u]=tF'[u]h+\sigma(u,th)$$

$$\Rightarrow \lim_{t\to 0}\frac{F[u+th]-F[u]}{t}=F'[u]h$$

where the RHS is nothing but F_u, from the definition of weak derivative. *Q.E.D.*

- Converse of the above statement does not hold. Proof was left as an excercise.
- Analogous to the situation where existence of partial derivates that are continuous, implies the existence of total derivate, we can say

  Given a map F, if weak derivative F_u exists in the neighbourhood of u_0 and represents a continuous operator at u_0, then F' exists at u_o and coincides with F_u.

3.3 Gradient of a functional

Consider U (has a scalar product defined on it), a space of numerical functions $u(x)$. Let F be a functional defined on U as

$$F : u \in U \to F[u] \in \mathbb{R}$$

Gradient of F *or* the functional derivative of F with respect to the scalar product is $G = \frac{\delta F}{\delta u}$, defined by the relation

$$\frac{d}{d\epsilon}F[u+\epsilon\phi]|_{\epsilon=0} =: \left(\frac{\delta F}{\delta u},\phi\right) \tag{5}$$

Second functional derivative is the weak derviative G_u of G i.e.

$$\frac{d^2}{d\epsilon d\tau}F[u+\epsilon\phi+\tau\psi]|_{\epsilon=\tau=0} =: (G_u\psi,\phi)$$

It is easy to check that G_u is symmteric with respect to scalar product.

4 A step away - an example

Let U be the space of all C^∞ functions[1] $u(x)$, defined on the interval $\mathcal{I}=[a,b)\in\mathbb{R}$. Let the functions and all their derivatives go to zero at a and b i.e.

$$u^{(k)}(a)=0 \;\;\&\;\; u^{(k)}(b)=0 \qquad \forall\;\; k\in[1,\infty)$$

Let $F[u]$ be a functional defined as

$$F[u]=\int_a^b f(u,u_x,u_{xx},\ldots,u_{nx},\ldots)dx$$

[1] These are also called smooth functions, because neither the functions nor their derivatives have any corners, as all of them are continuous

where $u_n x$ is the nth-dervative of u with respect to x. From the definition of gradient, we get

$$\begin{aligned} \delta F &= \frac{d}{d\epsilon} F[u + \epsilon h]|_{\epsilon=0} \\ &= \int_a^b \left[\frac{\partial f}{\partial u} h + \frac{\partial f}{\partial u_x} h_x + \frac{\partial f}{\partial u_{xx}} h_{xx} + \ldots + \frac{\partial f}{\partial u_{nx}} h_{nx} + \ldots \right] dx \\ &= \int_a^b \left[h \frac{\partial f}{\partial u} - h \frac{d}{dx} \frac{\partial f}{\partial u_x} + \ldots + h(-1)^n \frac{d^n}{dx^n} \frac{\partial f}{\partial u_{nx}} + \ldots \right] dx \end{aligned}$$

It can now be easily checked that

$$\frac{\delta F}{\delta u} = \frac{\partial f}{\partial u} - \frac{d}{dx} \frac{\partial f}{\partial u_x} + \ldots + (-1)^n \frac{d^n}{dx^n} \frac{\partial f}{\partial u_{nx}} + \ldots \tag{6}$$

5 Towards Hamilton's Principle

Consider the special case when f is only a function of u and u_x. Then the Eqn.(6) simplifies to

$$\frac{\delta F}{\delta u} = \frac{\partial f}{\partial u} - \frac{d}{dx} \frac{\partial f}{\partial u_x}$$

This can be generalized to n functions $u_1, \ldots, u_n$ such that

$$\frac{\delta F}{\delta u} \equiv \left(\frac{\delta F}{\delta u_1}, \ldots, \frac{\delta F}{\delta u_n} \right) \tag{7}$$

$$= \left(\frac{\partial f}{\partial u_1} - \frac{d}{dx} \frac{\partial f}{\partial u_{1_x}}, \ldots, \frac{\partial f}{\partial u_n} - \frac{d}{dx} \frac{\partial f}{\partial u_{n_x}} \right) \tag{8}$$

Now, if we recognize $x \mapsto t$ and $u \mapsto q$ (which implies $u_x \mapsto \dot{q}$, we can attribute $f \mapsto \mathcal{L}$ and $F \mapsto S$, which says that

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \frac{\partial \mathcal{L}}{\partial q_i} = \delta S \tag{9}$$

Extremal variation of S as defined in Eqn. (2) gives us the Euler-Lagrange equations of motion from the above result.

Additional remark

Actually, starting from Eqn. (6), and recognizing $x \mapsto t$ and $u \mapsto q$, we can also get the generalized expression Eqn. (1) for the Potential energy $\mathcal{U}$ in consevative systems by recognizing $f \mapsto \mathcal{U}(q, \dot{q})$. This is just a consequence of the definition of gradient of a functional and the fact that for conservative systems gradient vanishes.