The Calculus of Variations

MECH 642 - *Advanced Dynamics*

Prof. James Richard Forbes ©

McGill University, Department of Mechanical Engineering

Motivation - The Brachistochrone Problem [1, Ch. 10.1] [2, Ch. 3.8]

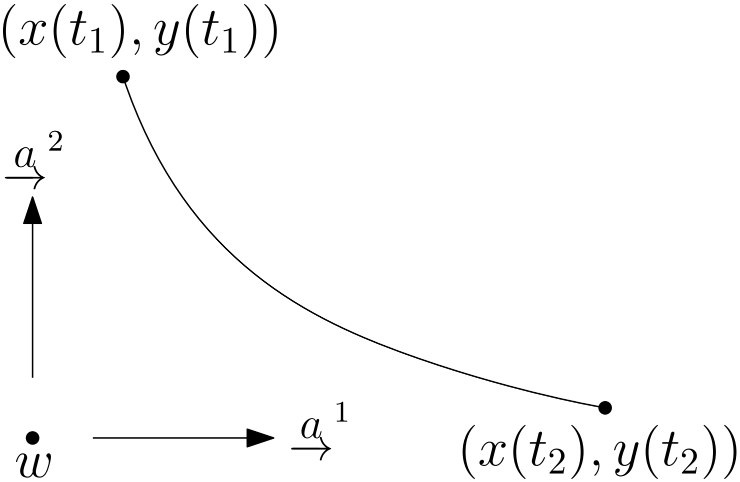


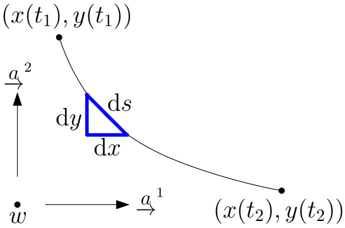
Figure 1: Particle sliding along a curve in a gravity field.

- The Brachistochrone problem is, "*find the shape of the curve from* $(x(t_1), y(t_1))$ *to* $(x(t_2), y(t_2))$ *such that a particle sliding on the curve goes from* $(x(t_1), y(t_1))$ *to* $(x(t_2), y(t_2))$ *in minimum time, without friction, in a gravity field.*"
- "Brachistos" = "shortest" and "chronos" = "time"
- Posed by Johan Bernoulli in 1696.
- Isaac Newton solved the Brachistochrone problem the day he heard of it.

An Optimization Problem

- The Brachistochrone problem is a type of optimization problem. Specifically, a “minimum time” problem.
- What is the cost function, or objective function, associated with the Brachistochrone problem?
- First recall

$$v(t) = \frac{\mathrm{d}s(t)}{\mathrm{d}t}, \qquad \mathrm{d}t = \frac{\mathrm{d}s(t)}{v(t)}. \tag{1}$$



- Via the pythagorean theorem,

$$(\mathrm{d}s(t))^2 = (\mathrm{d}x(t))^2 + (\mathrm{d}y(t))^2, \tag{2}$$

$$\left(\frac{\mathrm{d}s(t)}{\mathrm{d}x(t)}\right)^2 = 1 + \left(\frac{\mathrm{d}y(t)}{\mathrm{d}x(t)}\right)^2, \tag{3}$$

$$\mathrm{d}s(t) = \sqrt{1 + y'(t)}\mathrm{d}x(t), \tag{4}$$

where $y'(t) = \frac{\mathrm{d}y(t)}{\mathrm{d}x(t)}$.

- Going forward, drop the “(t)” for connivence.

- Next, consider energy conservation. Specifically, at any point on the curve

$$\tfrac{1}{2}mv^2 + mgy = \tfrac{1}{2}mv_1^2 + mgy_1 \tag{5}$$

  must hold, where $v_1 = v(t_1)$ and $y_1 = y(t_1)$.
- It follows that

$$v = \sqrt{2g(y_0 - y)}, \tag{6}$$

  where $y_0 = y_1 + \frac{v_1^2}{2g}$ is constant.
  - A positive velocity is assumed.
- Using (1), (4), and (6), it follows that

$$\mathrm{d}t = \frac{\mathrm{d}s(t)}{v(t)} = \frac{\sqrt{1+y'}\mathrm{d}x}{\sqrt{2g(y_0-y)}} = \sqrt{\frac{1+y'}{2g(y_0-y)}}\,\mathrm{d}x. \tag{7}$$

- Integrating both sides of (7) between t_1 and t_2 gives

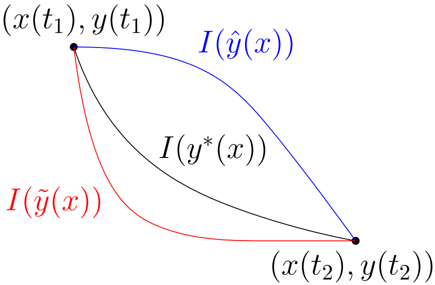
$$\int_{t_1}^{t_2} \mathrm{d}t = \int_{x_1}^{x_2} \sqrt{\frac{1+y'}{2g(y_0-y)}}\, \mathrm{d}x, \tag{8}$$

$$(t_2 - t_1) = \int_{x_1}^{x_2} \sqrt{\frac{1+y'(x)}{2g(y_0-y(x))}}\, \mathrm{d}x, \tag{9}$$

  where $x_1 = x(t_1)$ and $x_2 = x(t_2)$ and, as a reminder, $y = y(x) = y(x(t))$ and $y' = y'(x) = y(x(t))$.
  - That is, $y(x(t))$ is a function of $x(t)$, and $x(t)$ is a function of t.
- What is to be minimized? $t_2 - t_1$, right!
- Define the cost function to be minimized as

$$I(y(x)) = \int_{x_1}^{x_2} \sqrt{\frac{1+y'(x)}{2g(y_0-y(x))}}\, \mathrm{d}x. \tag{10}$$

- The cost function in (10) is a *functional* [3, Ch. 2.2.3].
  - A functional is a function of a function.
  - A functional maps a function to a real number.
  - Said another way, a functional assigns a number to a function.
- In this case, $I : C^\infty \to \mathbb{R}^+$
  - The domain of I is the set of continuously differentiable functions, C^∞.
  - The range of I is the positive real numbers, $\mathbb{R}^+$.
- Can think of guessing a candidate function $\tilde{y}(x)$, plugging it into (10), and getting a time $I(\tilde{y}(x)) = t_2 - t_1$ out.
- But guessing and checking isn't so smart . . .
- How can the function $y(x)$ that minimizes $I(y(x))$ be found?



- Optimal path, $y^*(x)$, and two nonoptimal paths, $\tilde{y}(x)$ and $\hat{y}(x)$.
- $I(y^*(x)) < I(\tilde{y}(x))$ and $I(y^*(x)) < I(\hat{y}(x))$.

Extremization - Finding a Minimum or Maximum

- The Taylor series expansion of a general function $f(\cdot) : \mathbb{R}^n \to \mathbb{R}$ is

$$f(\bar{\mathbf{x}} + \delta\mathbf{x}) = f(\bar{\mathbf{x}}) + \left[\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \bigg|_{\mathbf{x}=\bar{\mathbf{x}}} \right] \delta\mathbf{x} + \frac{1}{2}\delta\mathbf{x}^{\mathsf{T}} \left[\frac{\partial}{\partial \mathbf{x}} \left(\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} ^{\mathsf{T}} \right) \bigg|_{\mathbf{x}=\bar{\mathbf{x}}} \right] \delta\mathbf{x} + O(\delta x_i \delta x_j)$$

  where

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \bigg|_{\mathbf{x}=\bar{\mathbf{x}}} = \nabla f(\mathbf{x})^{\mathsf{T}}\big|_{\mathbf{x}=\bar{\mathbf{x}}}, \qquad \frac{\partial}{\partial \mathbf{x}} \left(\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} ^{\mathsf{T}} \right) \bigg|_{\mathbf{x}=\bar{\mathbf{x}}} = \nabla^2 f(\mathbf{x})\big|_{\mathbf{x}=\bar{\mathbf{x}}}$$

  are the Jacobain and Hessian of $f(\cdot)$ evaluated at $\mathbf{x} = \bar{\mathbf{x}}$, respectfully.

- A necessary condition for $\bar{\mathbf{x}}$ to be an extremum (either a maximum, a minimum, or a saddle point) is

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \bigg|_{\mathbf{x}=\bar{\mathbf{x}}} = \nabla f(\mathbf{x})^{\mathsf{T}}\big|_{\mathbf{x}=\bar{\mathbf{x}}} = \mathbf{0}.$$

- When $\nabla^2 f(\mathbf{x})\big|_{\mathbf{x}=\bar{\mathbf{x}}} > 0$ then $\bar{\mathbf{x}}$ corresponds to a minimum.
- When $\nabla^2 f(\mathbf{x})\big|_{\mathbf{x}=\bar{\mathbf{x}}} < 0$ then $\bar{\mathbf{x}}$ corresponds to a maximum.

Extremization of a Functional [1, Ch. 10.2] [2, Ch. 3.8.3] [4, Ch. 6.2]

- Consider a functional,

$$I(y(x)) = \int_{x_1}^{x_2} F(y(x), y'(x))\, \mathrm{d}x, \tag{11}$$

  subject to $y_1 = y(x_1) = y(x(t_1))$ and $y_2 = y(x_2) = y(x(t_2))$.
  - Could also have $F(y(x), y'(x), x)$. The forthcoming derivation does not change if $F(y(x), y'(x), x)$.
  - Can also constraints on $y_1' = y'(x_1)$ and $y_2' = y'(x_2)$, or even mixed boundary conditions such as $y_1 = y(x_1)$ and $y_2' = y'(x_2)$. For now the simple case of $y_1 = y(x_1)$ and $y_2 = y(x_2)$ is considered.
- The goal is to find $y(x)$ that extremizes (11).
- The tool of choice for extremization *anything* is a Taylor series expansion.
- However, in the present case, need to perform a Taylor series of *a functional*.

- Let $y^*(x)$ be the function that extremizes the functional $I(\cdot)$.
  - $I(y^*(x))$ is a max, a min, or a saddle (inflection) point.
- Consider a perturbed solution,

$$y(x) = y^*(x) + \epsilon\eta(x), \tag{12}$$

  where $\eta(x_1) = 0$ and $\eta(x_2) = 0$ so that the endpoint conditions $y^*(x_1) = y(x_1)$ and $y^*(x_2) = y(x_2)$ are satisfied.
- Other than the requirement that $\eta(x_1) = 0$ and $\eta(x_2) = 0$, $\eta(x)$ is arbitrary.
- $0 < \epsilon \ll 1$ is a small real number that scales $\eta(x)$.
- Together $\epsilon\eta(x)$ represent a perturbation to $y^*(x)$.
  - $\epsilon\eta(x)$ can be thought of as $\delta y(x)$.
- Note that $I(y^*(x)) < I(y(x)) = I(y^*(x) + \epsilon\eta(x))$.
- Also note that as $\epsilon \to 0$, $y(x) \to y^*(x)$.

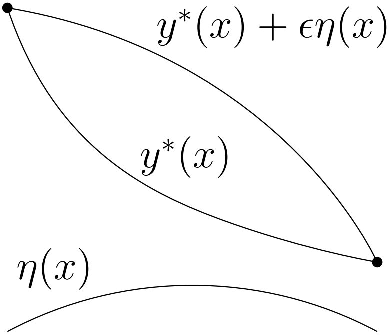


Figure 2: Optimal path $y^*(x)$. Perturbed path $y^*(x) + \epsilon\eta(x)$.

- With fixed $y^*(x)$ and fixed but arbitrary $\eta(x)$, via $y(x) = y^*(x) + \epsilon\eta(x)$ the functional is now a *function of* ϵ,

$$I(y(x)) = \int_{x_1}^{x_2} F(y(x), y'(x))\,\mathrm{d}x, \tag{11}$$

$$\Phi(\epsilon) = \int_{x_1}^{x_2} F(y^*(x) + \epsilon\eta(x), {y^*}'(x) + \epsilon\eta'(x))\,\mathrm{d}x. \tag{13}$$

- When $\epsilon = 0$, $\Phi(0) = I(y^*(x))$.
- The problem has been transformed from extremizing the functional $I(y(x))$ over all functions $y(x)$ to extremizing the function $\Phi(\epsilon)$ over all real numbers ϵ.

- To find the ϵ that extermizes $\Phi(\epsilon)$, execute a Taylor series expansion about $\epsilon = 0$.
  - Note that, in general, $\epsilon = \bar{\epsilon} + \delta\epsilon$.
  - However, when $\bar{\epsilon} = 0$, which is the present case, $\epsilon = 0 + \delta\epsilon$.
  - Rather than writing $\delta\epsilon$, just write ϵ.
- The Taylor expansion of $\Phi(\epsilon)$ is

$$\Phi(\epsilon) = \Phi(0) + \left.\frac{\mathrm{d}\Phi(\epsilon)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} \epsilon + \frac{1}{2} \left.\frac{\mathrm{d}^2\Phi(\epsilon)}{\mathrm{d}\epsilon^2}\right|_{\epsilon=0} \epsilon^2 + O(\epsilon^3). \tag{14}$$

- For an extremum, also called a *stationary value* of of $I(y(x))$, $\left.\frac{\mathrm{d}\Phi(\epsilon)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} = 0$.
- For the extremum to be a minimum, $\left.\frac{\mathrm{d}^2\Phi(\epsilon)}{\mathrm{d}\epsilon^2}\right|_{\epsilon=0} > 0$.
- For the extremum to be a maximum, $\left.\frac{\mathrm{d}^2\Phi(\epsilon)}{\mathrm{d}\epsilon^2}\right|_{\epsilon=0} < 0$.

- In detail, $\left.\frac{\mathrm{d}\Phi(\epsilon)}{\mathrm{d}\epsilon}\right|_{\epsilon=0}$ is

$$\left.\frac{\mathrm{d}\Phi(\epsilon)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} = \frac{\mathrm{d}}{\mathrm{d}\epsilon}\int_{x_1}^{x_2} F(y^*(x)+\epsilon\eta(x), {y^*}'(x)+\epsilon\eta'(x))\,\mathrm{d}x\Big|_{\epsilon=0} \tag{15}$$

$$= \int_{x_1}^{x_2} \frac{\mathrm{d}}{\mathrm{d}\epsilon} F(y^*(x)+\epsilon\eta(x), {y^*}'(x)+\epsilon\eta'(x))\,\mathrm{d}x\Big|_{\epsilon=0}, \tag{16}$$

which follows from the fact that the integral, and the integral bounds, do not depend on ϵ.

- Using the chain rule it follows that

$$\begin{aligned}\left.\frac{\mathrm{d}\Phi(\epsilon)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} &= \int_{x_1}^{x_2} \frac{\mathrm{d}}{\mathrm{d}\epsilon} F(y^*(x)+\epsilon\eta(x), {y^*}'(x)+\epsilon\eta'(x))\,\mathrm{d}x\Big|_{\epsilon=0} \\ &= \int_{x_1}^{x_2} \left(\frac{\partial F(y(x),y'(x))}{\partial y(x)}\frac{\mathrm{d}y(x)}{\mathrm{d}\epsilon} + \frac{\partial F(y(x),y'(x))}{\partial y'(x)}\frac{\mathrm{d}y'(x)}{\mathrm{d}\epsilon}\right)\,\mathrm{d}x\Big|_{\epsilon=0}. \end{aligned} \tag{17}$$

- Recalling that $y(x) = y^*(x) + \epsilon\eta(x)$, it follows that

$$\frac{\mathrm{d}y(x)}{\mathrm{d}\epsilon} = \eta(x), \qquad \frac{\mathrm{d}y'(x)}{\mathrm{d}\epsilon} = \eta'(x). \tag{18}$$

- Therefore,

$$\left.\frac{\mathrm{d}\Phi(\epsilon)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} = \int_{x_1}^{x_2} \left(\frac{\partial F(y(x), y'(x))}{\partial y(x)}\eta(x) + \frac{\partial F(y(x), y'(x))}{\partial y'(x)}\eta'(x) \right) \left.\mathrm{d}x\right|_{\epsilon=0}. \tag{19}$$

- The second term in (19) is

$$\int_{x_1}^{x_2} \frac{\partial F(y(x), y'(x))}{\partial y'(x)} \frac{\mathrm{d}\eta(x)}{\mathrm{d}x}\, \mathrm{d}x. \tag{20}$$

- Recall integration by parts, that is

$$\int_{x_1}^{x_2} u(x)\mathrm{d}v(x) = u(x)v(x)\Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} v(x)\mathrm{d}u(x). \tag{21}$$

- Define

$$u(x) = \frac{\partial F(y(x), y'(x))}{\partial y'(x)}, \qquad \frac{\mathrm{d}u(x)}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\right), \qquad \mathrm{d}u(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\right)\mathrm{d}x, \tag{22}$$

$$\mathrm{d}v(x) = \frac{\mathrm{d}\eta(x)}{\mathrm{d}x}\mathrm{d}x, \qquad \frac{\mathrm{d}v(x)}{\mathrm{d}x} = \frac{\mathrm{d}\eta(x)}{\mathrm{d}x}, \qquad v(x) = \eta(x) \tag{23}$$

- Using integration by parts, the second term in (19) can be written as

$$\int_{x_1}^{x_2} \frac{\partial F(y(x), y'(x))}{\partial y'(x)} \frac{\mathrm{d}\eta(x)}{\mathrm{d}x}\, \mathrm{d}x = F(y(x), y'(x))\eta(x)\Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\right)\eta(x)\, \mathrm{d}x \tag{24}$$

- Returning to (19), it follows that

$$\left.\frac{\mathrm{d}\Phi(\epsilon)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} = \left.\left(\int_{x_1}^{x_2}\left(\frac{\partial F(y(x),y'(x))}{\partial y(x)} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x),y'(x))}{\partial y'(x)}\right)\right)\eta(x)\,\mathrm{d}x + F(y(x),y'(x))\eta(x)\Big|_{x_1}^{x_2}\right)\right|_{\epsilon=0}. \tag{25}$$

- Setting $\left.\frac{\mathrm{d}\Phi(\epsilon)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} = 0$ and using the fact that $\eta(x_1) = 0$ and $\eta(x_2) = 0$ gives

$$0 = \left.\int_{x_1}^{x_2}\left(\frac{\partial F(y(x),y'(x))}{\partial y(x)} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x),y'(x))}{\partial y'(x)}\right)\right)\eta(x)\,\mathrm{d}x\right|_{\epsilon=0}. \tag{26}$$

- Furthermore, $y^*(x) = y(x)$ at $\epsilon = 0$, meaning

$$0 = \int_{x_1}^{x_2}\left(\frac{\partial F(y(x),y'(x))}{\partial y(x)} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x),y'(x))}{\partial y'(x)}\right)\right)\eta(x)\,\mathrm{d}x, \tag{27}$$

where $y^*(x)$ and $y(x)$ are interchangeable.

- The only way for the right hand side of (27) to be equal to zero for *any* $\eta(x)$ is that the integrand is exactly zero, meaning

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\right) - \frac{\partial F(y(x), y'(x))}{\partial y(x)} = 0, \qquad x_1 < x < x_2. \tag{28}$$

- Equation (28) is the *Euler-Lagrange* equation.
- In general, the Euler-Lagrange equation gives a differential equation.
- The solution to said differential equation extremizes the functional $I(y(x))$.
    - It is also common to say the solution to the Euler-Lagrange equation results in a stationary value of the functional $I(y(x))$.
    - The term "stationary value" is used because first-order variations in $y(x)$ do not change the value of $I(y(x))$ at the extremum.
    - Note, the stationary value could be a minimum, a maximum, or a saddle (inflection) point.

The Variational Operator [1, Ch. 10.2] [2, Ch. 3.8.4] [3, Ch. 2.2.4]

- Recall the variational operator.
- The variational operator is linear, meaning

$$\delta(\alpha g(x) + \beta h(x)) = \alpha \delta g(x) + \beta \delta h(x). \tag{29}$$

- Additionally, the variational operator commutes with both differentiation and integration, meaning

$$\delta \left(\frac{\mathrm{d}y(x)}{\mathrm{d}x} \right) = \frac{\mathrm{d}\delta y(x)}{\mathrm{d}x}, \tag{30}$$

$$\delta \int_{x_1}^{x_2} y(x) \, \mathrm{d}x = \int_{x_1}^{x_2} \delta y(x) \, \mathrm{d}x. \tag{31}$$

Using the Variational Operator

- Recall

$$y(x) = y^*(x) + \epsilon\eta(x). \tag{32}$$

  - $\eta(x)$ satisfies the BCs but is otherwise arbitrary.
  - Note that $\epsilon\frac{\mathrm{d}\eta(x)}{\mathrm{d}x} = \frac{\mathrm{d}(\epsilon\eta(x))}{\mathrm{d}x}$.
  - Similarly, note that $\epsilon\int_{x_1}^{x_2}\eta(x)\,\mathrm{d}x = \int_{x_1}^{x_2}\epsilon\eta(x)\,\mathrm{d}x$.
- Rather than working with ϵ and $\eta(x)$ work with the variational operator δ.
- In particular, replace $\epsilon\eta(x)$ with $\delta y(x)$ so that

$$y(x) = y^*(x) + \delta y(x) \tag{33}$$

- Just like $\epsilon\eta(x)$ is an arbitrary perturbation to $y^*(x)$, $\delta y(x)$ is an arbitrary variation to $y^*(x)$.
- Rather than writing * all the time, write

$$\tilde{y}(x) = y(x) + \delta y(x) \tag{34}$$

- Consider $F(y(x), y'(x))$.
- The variation in $F(y(x), y'(x))$ due to a variation in $y(x)$ is

$$
\begin{aligned}
& F(y(x) + \delta y(x), y'(x) + \delta y'(x)) && (35) \\
& = F(y(x), y'(x)) + \underbrace{\frac{\partial F(y(x), y'(x))}{\partial y(t)}\delta y(x) + \frac{\partial F(y(x), y'(x))}{\partial y'(t)}\delta y'(x) + O(\delta^2)}_{\delta F(y(x), y'(x))}\,. && (36)
\end{aligned}
$$

- For small variations $\delta y(x)$ and $\delta y'(x)$,

$$
\delta F(y(x), y'(x)) = \frac{\partial F(y(x), y'(x))}{\partial y(t)}\delta y(x) + \frac{\partial F(y(x), y'(x))}{\partial y'(t)}\delta y'(x), \qquad (37)
$$

which is the *first variation in* $F(y(x), y'(x))$.

Deriving the Euler-Lagrange Equation using the Variational Operator

- Consider the functional

$$I(y(x)) = \int_{x_1}^{x_2} F(y(x), y'(x))\,\mathrm{d}x. \tag{11}$$

- The first variation in $I(y(x))$ is

$$\begin{aligned}
\delta I(y(x)) &= \delta \int_{x_1}^{x_2} F(y(x), y'(x))\,\mathrm{d}x && (38)\\
&= \int_{x_1}^{x_2} \delta F(y(x), y'(x))\,\mathrm{d}x && (39)\\
&= \int_{x_1}^{x_2} \left(\frac{\partial F(y(x), y'(x))}{\partial y(t)} \delta y(x) + \frac{\partial F(y(x), y'(x))}{\partial y'(t)} \delta y'(x) \right)\,\mathrm{d}x, && (40)
\end{aligned}$$

where the fact δ commutes with the integral has been exploited.

- Again, recall integration by parts,

$$\int_{x_1}^{x_2} u(x)\mathrm{d}v(x) = u(x)v(x)\Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} v(x)\mathrm{d}u(x). \tag{21}$$

- Define

$$u(x) = \frac{\partial F(y(x), y'(x))}{\partial y'(x)}, \qquad \frac{\mathrm{d}u(x)}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\right), \qquad \mathrm{d}u(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\right)\mathrm{d}x, \tag{41}$$

$$\mathrm{d}v(x) = \delta\frac{\mathrm{d}y(x)}{\mathrm{d}x}\mathrm{d}x, \qquad \frac{\mathrm{d}v(x)}{\mathrm{d}x} = \frac{\mathrm{d}\delta y(x)}{\mathrm{d}x}, \qquad v(x) = \delta y(x). \tag{42}$$

- Using integration by parts it follows that

$$\delta I(y(x)) = \int_{x_1}^{x_2}\left(\frac{\partial F(y(x), y'(x))}{\partial y(t)}\delta y(x) + \frac{\partial F(y(x), y'(x))}{\partial y'(t)}\delta y'(x)\right)\,\mathrm{d}x, \tag{43}$$

$$= \int_{x_1}^{x_2}\left(\frac{\partial F(y(x), y'(x))}{\partial y(t)} - \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\right)\right)\delta y(x)\,\mathrm{d}x + \frac{\partial F(y(x), y'(x))}{\partial y'(x)}\delta y(x)\Big|_{x_1}^{x_2} \tag{44}$$

- The necessary condition for an extremum is $\delta I(y(x)) = 0$.
- For $\delta I(y(x)) = 0$,
  - the first term must be zero, meaning the integral must be zero, meaning coefficient in front of $\delta y(x)$ must be zero because $\delta y(x)$ is an arbitrary variation, and
  - the second term must be zero, meaning the BCs must result in the second term being zero.
- Therefore, for an extremum (a.k.a. a stationary value)

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\right) - \frac{\partial F(y(x), y'(x))}{\partial y(t)} = 0, \qquad x_1 < x < x_2, \tag{28}$$

and

$$\left.\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\delta y(x)\right|_{x_1} = 0, \qquad \left.\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\delta y(x)\right|_{x_2} = 0, \tag{45}$$

where

  - $\delta y(x_1) = 0$ and $\delta y(x_2) = 0$ are called *forced* or *geometric* boundary conditions, and
  - $\left.\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\right|_{x_1} = 0$ and $\left.\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\right|_{x_2} = 0$ are called *natural* boundary conditions.

- A forced boundary condition is one where $y(x)$ is prescribed.
  - For example, $y(x_1) = y(x)|_{x_1} = 0$ implies $\delta y(x_1) = \delta y(x)|_{x_1} = 0$ because $y(x)$ cannot vary at x_1.
  - In dynamics problems, a forced BC often takes the form of a geometric (e.g., position) constraint.
- If there is not a forced BC at x_1 or x_2, then from (45), the BC must be of the natural type.
  - For example, say $y(x)$ is not prescribed at x_1, but is prescribed at x_2.
  - Then the natural boundary condition $\left.\frac{\partial F(y(x),y'(x))}{\partial y'(x)}\right|_{x_1} = 0$ must be enforced.
  - In dynamics problems, a natural BC often takes the form of a kinematic (e.g., velocity or momentum) constraint.
- Note, the boundary conditions (BCs) can be mixed, such as $\delta y(x_1) = 0$ and $\left.\frac{\partial F(y(x),y'(x))}{\partial y'(x)}\right|_{x_2} = 0$.

Example [5, Ch. 2.4]

- Consider the Functional

$$I(y(x)) = \int_{x_1}^{x_2} \left(\tfrac{1}{2}m(y'(x))^2 - \tfrac{1}{2}ky(x)^2\right)\, \mathrm{d}x. \tag{46}$$

- The task at hand is to find $y(x)$, with appropriate BCs, resulting in a stationary value of the functional.

Methods 1 - Use Variational Calculus Directly

- Applying the variational operator gives

$$\delta I(y(x)) = \tfrac{1}{2}\int_{x_1}^{x_2} \left(2m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\delta\frac{\mathrm{d}y(x)}{\mathrm{d}x} - 2ky(x)\delta y(x)\right)\, \mathrm{d}x \tag{47}$$

$$= \int_{x_1}^{x_2} \left(m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\frac{\mathrm{d}\delta y(x)}{\mathrm{d}x} - ky(x)\delta y(x)\right)\, \mathrm{d}x. \tag{48}$$

- Recall, as usual, integration by parts,

$$\int_{x_1}^{x_2} u(x)\mathrm{d}v(x) = u(x)v(x)\Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} v(x)\mathrm{d}u(x). \tag{21}$$

- Define

$$u(x) = m\frac{\mathrm{d}y(x)}{\mathrm{d}x}, \qquad \frac{\mathrm{d}u(x)}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\left(m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right), \qquad \mathrm{d}u(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right)\mathrm{d}x, \tag{49}$$

$$\mathrm{d}v(x) = \delta\frac{\mathrm{d}y(x)}{\mathrm{d}x}\mathrm{d}x, \qquad \frac{\mathrm{d}v(x)}{\mathrm{d}x} = \frac{\mathrm{d}\delta y(x)}{\mathrm{d}x}, \qquad v(x) = \delta y(x). \tag{50}$$

- Using integration by parts it follows that

$$\delta I(y(x)) = \int_{x_1}^{x_2}\left(-\delta y(x)\frac{\mathrm{d}}{\mathrm{d}x}\left(m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right) - ky(x)\delta y(x)\right)\,\mathrm{d}x + \delta y(x)\left(m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right)\Bigg|_{x_1}^{x_2} \tag{51}$$

$$= \int_{x_1}^{x_2}\left(-\frac{\mathrm{d}}{\mathrm{d}x}\left(m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right) - ky(x)\right)\delta y(x)\,\mathrm{d}x + \delta y(x)\left(m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right)\Bigg|_{x_1}^{x_2}. \tag{52}$$

- The necessary condition for a stationary value is $\delta I(y(x)) = 0$ that in turn leads to

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right) + ky(x) = 0, \qquad x_1 < x < x_2, \tag{53}$$

$$\delta y(x)\left(m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right)\bigg|_{x_1} = 0, \qquad \delta y(x)\left(m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right)\bigg|_{x_2} = 0. \tag{54}$$

- The BCs can be satisfied in a variety of ways [2, Ch. 3.8.5].
- Prescribed "positions" $y_1 = y(x_1)$ and $y_2 = y(x_2)$ imply $\delta y(x_1) = 0$ and $\delta y(x_2) = 0$ and the BCs are satisfied.
  - In this case the "momentum" at the endpoints are not necessarily zero, that is $m\left.\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right|_{x_1}$ and $m\left.\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right|_{x_2}$ can be non-zero.
- A prescribed "position" $y_1 = y(x_1)$ at x_1 and zero "momentum" $m\left.\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right|_{x_2} = 0$ at x_2 satisfies the BCs.
  - And vise versa.
- Zero "momentum" at both endpoints, $m\left.\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right|_{x_1} = 0$ and $m\left.\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right|_{x_2} = 0$, satisfies the BCs.
  - In this case, δy_1 and δy_2 are not necessarily zero implying y_1 and y_2 are not prescribed.

Method 2 - Use the Euler-Lagrange Equation Directly

- From (46) it follows that

$$F(y(x), y'(x)) = \tfrac{1}{2}m(y'(x))^2 - \tfrac{1}{2}ky(x)^2. \tag{55}$$

- The Euler-Lagrange equation is

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\right) - \frac{\partial F(y(x), y'(x))}{\partial y(t)} = 0, \qquad x_1 < x < x_2. \tag{28}$$

- The first term in the Euler-Lagrange equation is

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\right) = \frac{\mathrm{d}}{\mathrm{d}x}\left(m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right). \tag{56}$$

- The second term in the Euler-Lagrange equation is

$$\frac{\partial F(y(x), y'(x))}{\partial y(t)} = -ky(x). \tag{57}$$

- It then follows that

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x),y'(x))}{\partial y'(x)}\right)-\frac{\partial F(y(x),y'(x))}{\partial y(t)}=0,\qquad x_1<x<x_2, \tag{28}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right)+ky(x)=0,\qquad x_1<x<x_2. \tag{58}$$

- The solution is not complete with out specifying the BCs. Using (45),

$$\left.\frac{\partial F(y(x),y'(x))}{\partial y'(x)}\delta y(x)\right|_{x_1}=0,\qquad \left.\frac{\partial F(y(x),y'(x))}{\partial y'(x)}\delta y(x)\right|_{x_2}=0, \tag{45}$$

$$\left.\left(m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right)\delta y(x)\right|_{x_1}=0,\qquad \left.\left(m\frac{\mathrm{d}y(x)}{\mathrm{d}x}\right)\delta y(x)\right|_{x_2}=0. \tag{59}$$

- Method 1 and Method 2 yield the same result, as they should.

Extremizing Subject to Constraints

- Consider the functional

$$I(y(x)) = \int_{x_1}^{x_2} F(y(x), y'(x))\,\mathrm{d}x \tag{11}$$

  to be extremized subject to the constraint

$$\varphi(y(x)) = 0. \tag{60}$$

- The constraint must be enforced over $x_1 \leq x \leq x_2$.
- Using the *method of Lagrange multipliers*, define

$$J(y(x), \lambda(x)) = \int_{x_1}^{x_2} F(y(x), y'(x))\,\mathrm{d}x + \int_{x_1}^{x_2} \lambda(x)\varphi(y(x))\,\mathrm{d}x \tag{61}$$

  - “Add zero”.

- Now continue on using variational calculus as per usual.
- The first variation in $J(y(x), \lambda(x))$ is

$$\begin{aligned}
\delta J(y(x), \lambda(x)) &= \delta \int_{x_1}^{x_2} F(y(x), y'(x))\, \mathrm{d}x + \delta \int_{x_1}^{x_2} \lambda(x)\varphi(y(x))\, \mathrm{d}x && (62)\\
&= \int_{x_1}^{x_2} \delta F(y(x), y'(x))\, \mathrm{d}x + \int_{x_1}^{x_2} \delta\lambda(x)\varphi(y(x))\, \mathrm{d}x + \int_{x_1}^{x_2} \lambda(x)\delta\varphi(y(x))\, \mathrm{d}x && (63)\\
&= \int_{x_1}^{x_2} \left(\frac{\partial F(y(x), y'(x))}{\partial y(t)}\delta y(x) + \frac{\partial F(y(x), y'(x))}{\partial y'(t)}\delta y'(x)\right)\, \mathrm{d}x && (64)\\
&+ \int_{x_1}^{x_2} \lambda(x)\frac{\mathrm{d}\varphi(y(x))}{\mathrm{d}y(x)}\delta y(x)\, \mathrm{d}x + \int_{x_1}^{x_2} \delta\lambda(x)\varphi(y(x))\, \mathrm{d}x && (65)
\end{aligned}$$

where the fact δ commutes with the integral has been exploited.

- Using integration by parts it follows that

$$J(y(x),\lambda(x)) = \int_{x_1}^{x_2} \left(\frac{\partial F(y(x),y'(x))}{\partial y(t)} - \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{\partial F(y(x),y'(x))}{\partial y'(x)} \right) \right) \delta y(x) \,\mathrm{d}x + \frac{\partial F(y(x),y'(x))}{\partial y'(x)} \delta y(x) \Big|_{x_1}^{x_2} \tag{66}$$

$$+ \int_{x_1}^{x_2} \lambda(x) \frac{\mathrm{d}\varphi(y(x))}{\mathrm{d}y(x)} \delta y(x) \,\mathrm{d}x + \int_{x_1}^{x_2} \delta\lambda(x) \varphi(y(x)) \,\mathrm{d}x \tag{67}$$

$$= \int_{x_1}^{x_2} \left(\frac{\partial F(y(x),y'(x))}{\partial y(t)} - \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{\partial F(y(x),y'(x))}{\partial y'(x)} \right) + \lambda(x) \frac{\mathrm{d}\varphi(y(x))}{\mathrm{d}y(x)} \right) \delta y(x) \,\mathrm{d}x \tag{68}$$

$$+ \frac{\partial F(y(x),y'(x))}{\partial y'(x)} \delta y(x) \Big|_{x_1}^{x_2} + \int_{x_1}^{x_2} \delta\lambda(x) \varphi(y(x)) \,\mathrm{d}x. \tag{69}$$

- The necessary condition for an extremum is $\delta J(y(x), \lambda(x)) = 0$.
- For $\delta J(y(x), \lambda(x)) = 0$,
  - the first term must be zero, meaning the integral must be zero, meaning coefficient in front of $\delta y(x)$ must be zero because $\delta y(x)$ is an arbitrary variation,
  - the second term must be zero, meaning the BCs must result in the second term being zero, and
  - the third term must be zero, meaning $\varphi(y(x)) = 0$ must hold because $\delta\lambda(x)$ is also an arbitrary variation.
- Therefore, for an extremum, must have

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\right) - \frac{\partial F(y(x), y'(x))}{\partial y(t)} = \lambda(x)\frac{\mathrm{d}\varphi(y(x))}{\mathrm{d}y(x)}, \qquad x_1 < x < x_2, \tag{70}$$

$$\varphi(y(x)) = 0, \tag{60}$$

and

$$\left.\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\delta y(x)\right|_{x_1} = 0, \qquad \left.\frac{\partial F(y(x), y'(x))}{\partial y'(x)}\delta y(x)\right|_{x_2} = 0. \tag{45}$$

- Equations (70) and (60) must be solved simultaneously subject to the BCs (45).

- In some situations the constraint is directly given in variational form, such as

$$g(y(x))\delta y(x) = 0. \tag{71}$$

- The previous derivation still holds, just with (71) replacing $\frac{\mathrm{d}\varphi(y(x))}{\mathrm{d}y(x)}\delta y(x)$.

The Multivariable Case [5, Ch. 2.6]

- Consider the functional

$$I(u(x), v(x)) = \int_{x_1}^{x_2} F(u(x), v(x), u'(x), v'(x))\, \mathrm{d}x. \tag{72}$$

  - The forthcoming derivation does not change if $F(u(x), v(x), u'(x), v'(x), x)$.
- $I(u(x), v(x))$ is to be extremized.

- The variation in $F(u(x), v(x), u'(x), v'(x))$ due to a variation in $u(x)$ and $v(x)$ is

$$
\begin{aligned}
& F(u(x)+\delta u(x), v(x)+\delta v(x), u'(x)+\delta u'(x), v'(x)+\delta v'(x)) && (73)\\
&= F(u(x), v(x), u'(x), v'(x)) && (74)\\
&+ \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u(x)}\delta u(x) + \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u'(x)}\delta u'(x) && (75)\\
&+ \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v(x)}\delta v(x) + \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v'(x)}\delta v'(x) + O(\delta^2). && (76)
\end{aligned}
$$

- For small variations $\delta y(x)$ and $\delta y'(x)$,

$$
\delta F(u(x), v(x), u'(x), v'(x)) = \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u(x)}\delta u(x) + \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u'(x)}\delta u'(x) \tag{77}
$$

$$
+ \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v(x)}\delta v(x) + \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v'(x)}\delta v'(x). \tag{78}
$$

- The first variation in $I(u(x), v(x))$ is

$$\delta I(u(x), v(x)) = \delta \int_{x_1}^{x_2} F(u(x), v(x), u'(x), v'(x)) \, \mathrm{d}x \tag{79}$$

$$= \int_{x_1}^{x_2} \delta F(u(x), v(x), u'(x), v'(x)) \, \mathrm{d}x \tag{80}$$

$$= \int_{x_1}^{x_2} \Bigg(\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u(x)} \delta u(x) + \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u'(x)} \delta u'(x) \tag{81}$$

$$+ \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v(x)} \delta v(x) + \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v'(x)} \delta v'(x) \Bigg) \, \mathrm{d}x, \tag{82}$$

  where the fact δ commutes with the integral has been exploited.

- Yet again, recall integration by parts,

$$\int_{x_1}^{x_2} a(x) \mathrm{d}b(x) = a(x)b(x) \Big|_{x_1}^{x_2} - \int_{x_1}^{x_2} b(x) \mathrm{d}a(x). \tag{21}$$

- Define

$$a(x) = \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u'(x)}, \tag{83}$$

$$\frac{\mathrm{d}a(x)}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u'(x)}\right), \tag{84}$$

$$\mathrm{d}a(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u'(x)}\right)\mathrm{d}x. \tag{85}$$

and

$$\mathrm{d}b(x) = \delta\frac{\mathrm{d}u(x)}{\mathrm{d}x}\mathrm{d}x, \tag{86}$$

$$\frac{\mathrm{d}b(x)}{\mathrm{d}x} = \frac{\mathrm{d}\delta u(x)}{\mathrm{d}x}, \tag{87}$$

$$b(x) = \delta u(x). \tag{88}$$

- Similarly, define

$$\alpha(x) = \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v'(x)}, \tag{89}$$

$$\frac{\mathrm{d}\alpha(x)}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v'(x)}\right), \tag{90}$$

$$\mathrm{d}\alpha(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v'(x)}\right)\mathrm{d}x. \tag{91}$$

and

$$\mathrm{d}\beta(x) = \delta\frac{\mathrm{d}v(x)}{\mathrm{d}x}\mathrm{d}x, \tag{92}$$

$$\frac{\mathrm{d}\beta(x)}{\mathrm{d}x} = \frac{\mathrm{d}\delta v(x)}{\mathrm{d}x}, \tag{93}$$

$$\beta(x) = \delta v(x). \tag{94}$$

- Using integration by parts it follows that

$$\delta I(u(x), v(x)) = \int_{x_1}^{x_2} \left(\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u(x)} - \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u'(x)} \right) \right) \delta u(x) \, \mathrm{d}x \tag{95}$$

$$+ \left. \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u'(x)} \delta u(x) \right|_{x_1}^{x_2} \tag{96}$$

$$+ \int_{x_1}^{x_2} \left(\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v(x)} - \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v'(x)} \right) \right) \delta v(x) \, \mathrm{d}x \tag{97}$$

$$+ \left. \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v'(x)} \delta v(x) \right|_{x_1}^{x_2} . \tag{98}$$

- Setting $\delta I(u(x), v(x)) = 0$, and recalling that $\delta u(x)$ and $\delta v(x)$ are arbitrary, it follows that

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u'(x)}\right) - \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u(x)} = 0, \tag{99}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v'(x)}\right) - \frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v(x)} = 0, \tag{100}$$

and

$$\left.\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u'(x)}\delta u(x)\right|_{x_1} = 0, \qquad \left.\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial u'(x)}\delta u(x)\right|_{x_2} = 0, \tag{101}$$

$$\left.\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v'(x)}\delta v(x)\right|_{x_1} = 0, \qquad \left.\frac{\partial F(u(x), v(x), u'(x), v'(x))}{\partial v'(x)}\delta v(x)\right|_{x_2} = 0, \tag{102}$$

must hold for an extremum (a.k.a., a stationary value).

References

Slides based on [1, Ch. 10], [2, Ch. 3], [4, Ch. 6], [3, Ch. 2], and [5, Ch. 2].

[1] G. M. T. D'Eleutario and G. R. Heppler, *Newton's Second Law And All That (Preprint)*. Cambridge University Press, 2011.

[2] J. Hansen, "AER 373 – Mechanics of Solids and Structures Course Notes," *University of Toronto Institute For Aerospace Studies*, 2007.

[3] J. N. Reddy, *Introduction to the Finite Element Method*, Second. New York, NY: McGraw Hill, 1993.

[4] H. Schaub and J. L. Junkins, *Analytical Mechanics of Space Systems*, 2nd. Reston, VA: American Institute of Aeronautics and Astronautics, Inc., 2009.

[5] F. B. Hildebrand, *Methods of Applied Mathematics*. Dover: Mineola, New York, 1992.