16.323 Principles of Optimal Control
Spring 2008

16.323 Lecture 5

Calculus of Variations

- Calculus of Variations
- Most books cover this material well, but Kirk Chapter 4 does a particularly nice job.
- See here for online reference.

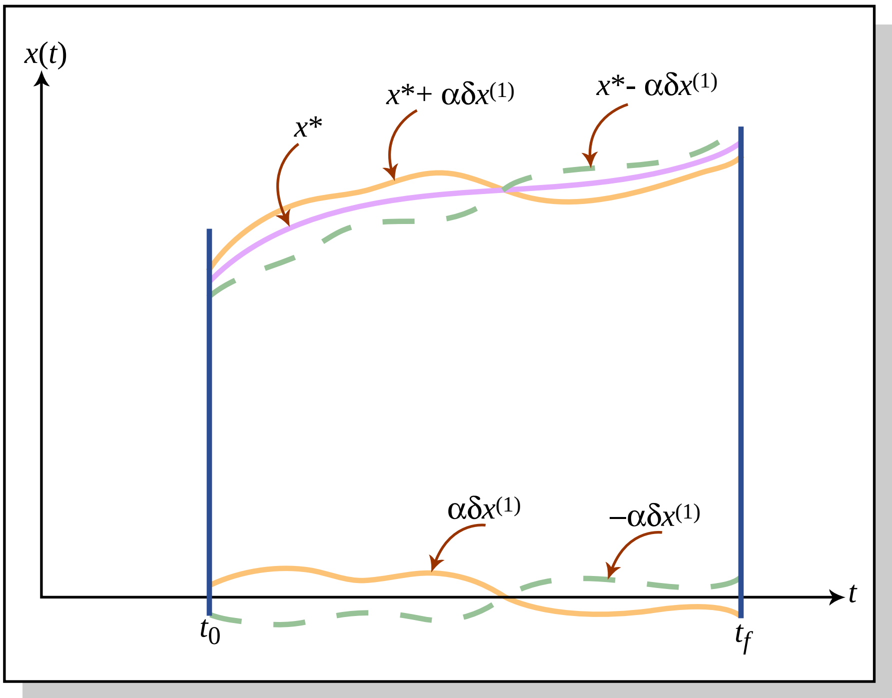


Figure by MIT OpenCourseWare.

Calculus of Variations

- **Goal:** Develop alternative approach to solve general optimization problems for continuous systems – **variational calculus**
  - Formal approach will provide new insights for constrained solutions, and a more direct path to the solution for other problems.

- **Main issue** – General control problem, the cost is a function of functions $\mathbf{x}(t)$ and $\mathbf{u}(t)$.

$$\min J = h(\mathbf{x}(t_f)) + \int_{t_0}^{t_f} g(\mathbf{x}(t), \mathbf{u}(t), t))\, dt$$

  subject to

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{u}, t) \\ \mathbf{x}(t_0), t_0 &\ \text{ given} \\ m(\mathbf{x}(t_f), t_f) &= 0 \end{aligned}$$

  - Call $J(\mathbf{x}(t), \mathbf{u}(t))$ a **functional**.

- Need to investigate how to find the optimal values of a functional.
  - For a function, we found the gradient, and set it to zero to find the stationary points, and then investigated the higher order derivatives to determine if it is a maximum or minimum.
  - Will investigate something similar for functionals.

- **Maximum and Minimum of a Function**
  - A function $f(\mathbf{x})$ has a local minimum at $\mathbf{x}^\star$ if

    $$f(\mathbf{x}) \geq f(\mathbf{x}^\star)$$

    for all admissible $\mathbf{x}$ in $\|\mathbf{x} - \mathbf{x}^\star\| \leq \epsilon$
  - Minimum can occur at (i) stationary point, (ii) at a boundary, or (iii) a point of discontinuous derivative.
  - If only consider stationary points of the differentiable function $f(\mathbf{x})$, then statement equivalent to requiring that differential of f satisfy:

    $$df = \frac{\partial f}{\partial \mathbf{x}} d\mathbf{x} = 0$$

    for all small $d\mathbf{x}$, which gives the same necessary condition from Lecture 1

    $$\frac{\partial f}{\partial \mathbf{x}} = 0$$

- Note that this definition used norms to compare two vectors. Can do the same thing with functions ⇒ distance between two functions

  $$d = \|\mathbf{x}_2(t) - \mathbf{x}_1(t)\|$$

  where

  1. $\|\mathbf{x}(t)\| \geq 0$ for all $\mathbf{x}(t)$, and $\|\mathbf{x}(t)\| = 0$ only if $\mathbf{x}(t) = 0$ for all t in the interval of definition.
  2. $\|a\mathbf{x}(t)\| = |a|\|\mathbf{x}(t)\|$ for all real scalars a.
  3. $\|\mathbf{x}_1(t) + \mathbf{x}_2(t)\| \leq \|\mathbf{x}_1(t)\| + \|\mathbf{x}_2(t)\|$

- Common function norm:

  $$\|\mathbf{x}(t)\|_2 = \left(\int_{t_0}^{t_f} \mathbf{x}(t)^T \mathbf{x}(t) dt \right)^{1/2}$$

- **Maximum and Minimum of a Functional**
  - A functional $J(\mathbf{x}(t))$ has a local minimum at $\mathbf{x}^\star(t)$ if

$$J(\mathbf{x}(t)) \geq J(\mathbf{x}^\star(t))$$

    for all admissible $\mathbf{x}(t)$ in $\|\mathbf{x}(t) - \mathbf{x}^\star(t)\| \leq \epsilon$

- Now define something equivalent to the differential of a function - called a **variation** of a functional.
  - An **increment** of a functional

$$\Delta J(\mathbf{x}(t), \delta\mathbf{x}(t)) = J(\mathbf{x}(t) + \delta\mathbf{x}(t)) - J(\mathbf{x}(t))$$

  - A **variation** of the functional is a linear approximation of this increment:

$$\Delta J(\mathbf{x}(t), \delta\mathbf{x}(t)) = \delta J(\mathbf{x}(t), \delta\mathbf{x}(t)) + H.O.T.$$

    i.e. $\delta J(\mathbf{x}(t), \delta\mathbf{x}(t))$ is linear in $\delta\mathbf{x}(t)$.

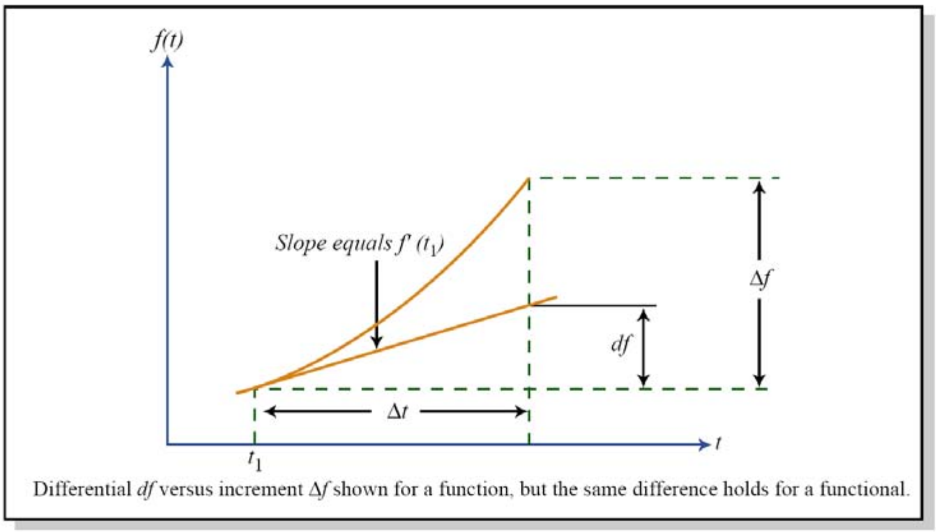


Figure by MIT OpenCourseWare.

Figure 5.1: Differential df versus increment Δf shown for a function, but the same difference holds for a functional.

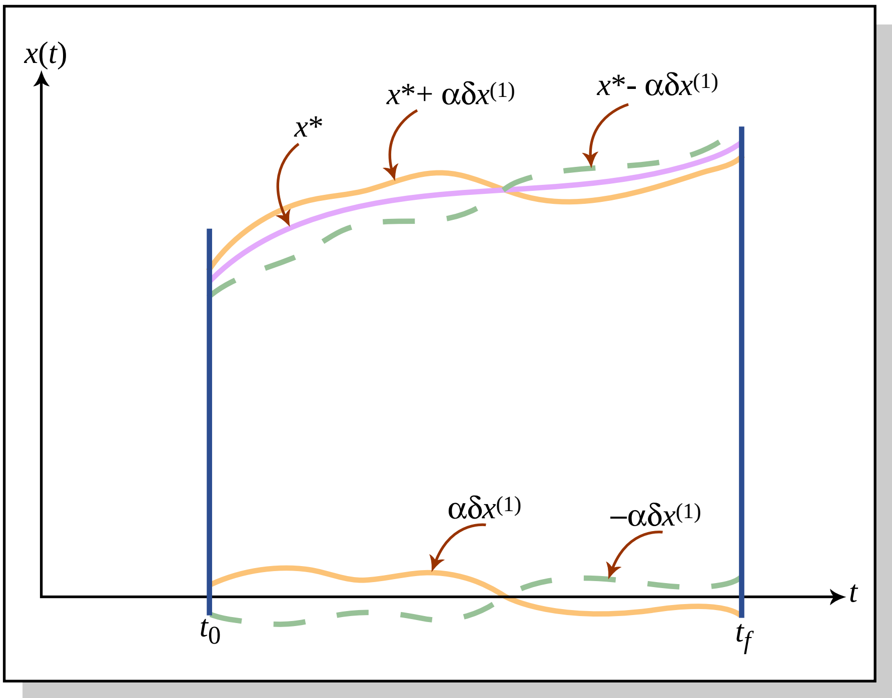


Figure by MIT OpenCourseWare.

Figure 5.2: Visualization of perturbations to function $x(t)$ by $\delta x(t)$ – it is a potential change in the value of x over the entire time period of interest. Typically require that if $x(t)$ is in some class (i.e., continuous), that $x(t) + \delta x(t)$ is also in that class.

- **Fundamental Theorem of the Calculus of Variations**
  - Let $\mathbf{x}$ be a function of t in the class Ω, and $J(\mathbf{x})$ be a differentiable functional of $\mathbf{x}$. Assume that the functions in Ω are not constrained by any boundaries.
  - If $\mathbf{x}^\star$ is an extremal function, then the variation of J must vanish on $\mathbf{x}^\star$, i.e. for all admissible $\delta\mathbf{x}$,

$$\delta J(\mathbf{x}(t), \delta\mathbf{x}(t)) = 0$$

  - Proof is in Kirk, page 121, but it is relatively straightforward.

- How compute the variation? If $J(\mathbf{x}(t)) = \int_{t_0}^{t_f} f(\mathbf{x}(t))dt$ where f has cts first and second derivatives with respect to $\mathbf{x}$, then

$$\begin{aligned}
\delta J(\mathbf{x}(t), \delta\mathbf{x}) &= \int_{t_0}^{t_f} \left\{ \frac{\partial f(\mathbf{x}(t))}{\partial \mathbf{x}(t)} \right\} \delta\mathbf{x} dt + f(\mathbf{x}(t_f))\delta t_f - f(\mathbf{x}(t_0))\delta t_0 \\
&= \int_{t_0}^{t_f} f_{\mathbf{x}}(\mathbf{x}(t))\delta\mathbf{x} dt + f(\mathbf{x}(t_f))\delta t_f - f(\mathbf{x}(t_0))\delta t_0
\end{aligned}$$

Variation Examples: Scalar

- For more general problems, first consider the cost evaluated on a scalar function $x(t)$ with t_0, t_f and the curve endpoints fixed.

$$J(x(t)) = \int_{t_0}^{t_f} g(x(t), \dot{x}(t), t)dt$$

$$\Rightarrow \quad \delta J(x(t), \delta x) = \int_{t_0}^{t_f} [\ g_x(x(t), \dot{x}(t), t)\delta x + g_{\dot{x}}(x(t), \dot{x}(t), t)\delta \dot{x}]\, dt$$

  – Note that

$$\delta \dot{x} = \frac{d}{dt}\delta x$$

  so δx and $\delta \dot{x}$ **are not independent**.

- Integrate by parts:

$$\int u dv \equiv uv - \int v du$$

  with $u = g_{\dot{x}}$ and $dv = \delta \dot{x} dt$ to get:

$$\begin{aligned}
\delta J(x(t), \delta x) &= \int_{t_0}^{t_f} g_x(x(t), \dot{x}(t), t)\delta x dt + [g_{\dot{x}}(x(t), \dot{x}(t), t)\delta x]_{t_0}^{t_f} \\
&\quad - \int_{t_0}^{t_f} \frac{d}{dt} g_{\dot{x}}(x(t), \dot{x}(t), t)\delta x dt \\
&= \int_{t_0}^{t_f} \left[g_x(x(t), \dot{x}(t), t) - \frac{d}{dt} g_{\dot{x}}(x(t), \dot{x}(t), t) \right] \delta x(t) dt \\
&\quad + [g_{\dot{x}}(x(t), \dot{x}(t), t)\delta x]_{t_0}^{t_f}
\end{aligned}$$

- Since $x(t_0),\ x(t_f)$ given, then $\delta x(t_0) = \delta x(t_f) = 0$, yielding

$$\delta J(x(t), \delta x) = \int_{t_0}^{t_f} \left[g_x(x(t), \dot{x}(t), t) - \frac{d}{dt} g_{\dot{x}}(x(t), \dot{x}(t), t) \right] \delta x(t) dt$$

- Recall need $\delta J = 0$ for all admissible $\delta x(t)$, which are arbitrary within (t_0, t_f) $\Rightarrow$ the (first order) necessary condition for a maximum or minimum is called **Euler Equation**:

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

- **Example**: Find the curve that gives the shortest distance between 2 points in a plane (x_0, y_0) and (x_f, y_f).
  - Cost function – sum of differential arc lengths:

$$\begin{aligned} J &= \int_{x_0}^{x_f} ds = \int_{x_0}^{x_f} \sqrt{(dx)^2 + (dy)^2} \\ &= \int_{x_0}^{x_f} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx \end{aligned}$$

  - Take y as dependent variable, and x as independent one

$$\frac{dy}{dx} \to \dot{y}$$

  - New form of the cost:

$$J = \int_{x_0}^{x_f} \sqrt{1 + \dot{y}^2}\, dx \to \int_{x_0}^{x_f} g(\dot{y}) dx$$

  - Take partials: $\partial g / \partial y = 0$, and

$$\begin{aligned} \frac{d}{dx}\left(\frac{\partial g}{\partial \dot{y}}\right) &= \frac{d}{d\dot{y}}\left(\frac{\partial g}{\partial \dot{y}}\right)\frac{d\dot{y}}{dx} \\ &= \frac{d}{d\dot{y}}\left(\frac{\dot{y}}{(1+\dot{y}^2)^{1/2}}\right)\ddot{y} = \frac{\ddot{y}}{(1+\dot{y}^2)^{3/2}} = 0 \end{aligned}$$

    which implies that $\ddot{y} = 0$
  - Most general curve with $\ddot{y} = 0$ is a line $y = c_1 x + c_2$

Vector Functions

- Can generalize the problem by including several (N) functions $x_i(t)$ and possibly free endpoints

$$J(\mathbf{x}(t)) = \int_{t_0}^{t_f} g(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)dt$$

with t_0, t_f, $\mathbf{x}(t_0)$ fixed.

- Then (drop the arguments for brevity)

$$\delta J(\mathbf{x}(t), \delta\mathbf{x}) = \int_{t_0}^{t_f} \left[\, g_{\mathbf{x}}\delta\mathbf{x}(t) + g_{\dot{\mathbf{x}}}\delta\dot{\mathbf{x}}(t)\right] dt$$

  – Integrate by parts to get:

$$\delta J(\mathbf{x}(t), \delta\mathbf{x}) \;=\; \int_{t_0}^{t_f} \left[g_{\mathbf{x}} - \frac{d}{dt} g_{\dot{\mathbf{x}}}\right] \delta\mathbf{x}(t)dt + g_{\dot{\mathbf{x}}}(\mathbf{x}(t_f), \dot{\mathbf{x}}(t_f), t_f)\delta\mathbf{x}(t_f)$$

- The requirement then is that for $t \in (t_0, t_f)$, $\mathbf{x}(t)$ must satisfy

$$\frac{\partial g}{\partial \mathbf{x}} - \frac{d}{dt}\frac{\partial g}{\partial \dot{\mathbf{x}}} = 0$$

where $\mathbf{x}(t_0) = \mathbf{x}_0$ which are the given N boundary conditions, and the remaining N more BC follow from:

  – $\mathbf{x}(t_f) = \mathbf{x}_f$ if $\mathbf{x}_f$ is given as fixed,

  – If $\mathbf{x}(t_f)$ are free, then

$$\frac{\partial g(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)}{\partial \dot{\mathbf{x}}(t_f)} = 0$$

- Note that we could also have a mixture, where parts of $\mathbf{x}(t_f)$ are given as fixed, and other parts are free – just use the rules above on each component of $x_i(t_f)$

Free Terminal Time

- Now consider a slight variation: the goal is to minimize

$$J(\mathbf{x}(t)) = \int_{t_0}^{t_f} g(\mathbf{x}(t), \dot{\mathbf{x}}(t), t) dt$$

with t_0, $\mathbf{x}(t_0)$ fixed, t_f free, and various constraints on $\mathbf{x}(t_f)$

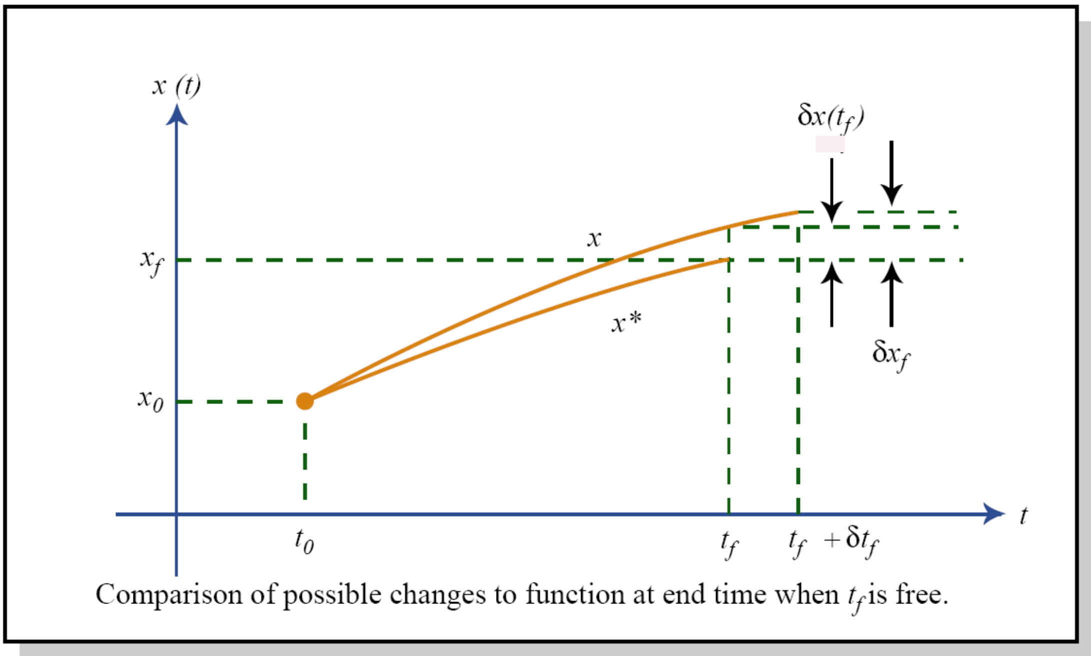
- Compute variation of the functional considering 2 candidate solutions:
  - $\mathbf{x}(t)$, which we consider to be a perturbation of the optimal $\mathbf{x}^\star(t)$ (that we need to find)

$$\delta J(\mathbf{x}^\star(t), \delta\mathbf{x}) \;=\; \int_{t_0}^{t_f} [\; g_{\mathbf{x}} \delta\mathbf{x}(t) + g_{\dot{\mathbf{x}}} \delta\dot{\mathbf{x}}(t)]\, dt + g(\mathbf{x}^\star(t_f), \dot{\mathbf{x}}^\star(t_f), t_f)\delta t_f$$

  - Integrate by parts to get:

$$\begin{aligned}\delta J(\mathbf{x}^\star(t), \delta\mathbf{x}) \;&=\; \int_{t_0}^{t_f} \left[\; g_{\mathbf{x}} - \frac{d}{dt} g_{\dot{\mathbf{x}}} \right] \delta\mathbf{x}(t) dt \\ &+\; g_{\dot{\mathbf{x}}}(\mathbf{x}^\star(t_f), \dot{\mathbf{x}}^\star(t_f), t_f)\delta\mathbf{x}(t_f) \\ &+\; g(\mathbf{x}^\star(t_f), \dot{\mathbf{x}}^\star(t_f), t_f)\delta t_f\end{aligned}$$

- Looks standard so far, but we have to be careful how we handle the terminal conditions



Comparison of possible changes to function at end time when t_f is free.

Figure by MIT OpenCourseWare.

Figure 5.3: Comparison of possible changes to function at end time when t_f is free.

- By definition, $\delta\mathbf{x}(t_f)$ is the difference between two admissible functions at time t_f (in this case the optimal solution $\mathbf{x}^\star$ and another candidate $\mathbf{x}$).
  - But in this case, must also account for possible changes to δt_f.
  - Define $\delta\mathbf{x}_f$ as being the difference between the ends of the two possible functions – **total possible change** in the final state:

$$\delta\mathbf{x}_f \approx \delta\mathbf{x}(t_f) + \dot{\mathbf{x}}^\star(t_f)\delta t_f$$

so $\delta\mathbf{x}(t_f) \neq \delta\mathbf{x}_f$ in general.

- Substitute to get

$$\begin{aligned}\delta J(\mathbf{x}^\star(t), \delta\mathbf{x}) &= \int_{t_0}^{t_f} \left[g_{\mathbf{x}} - \frac{d}{dt} g_{\dot{\mathbf{x}}} \right] \delta\mathbf{x}(t) dt + g_{\dot{\mathbf{x}}}(\mathbf{x}^\star(t_f), \dot{\mathbf{x}}^\star(t_f), t_f)\delta\mathbf{x}_f \\ &+ \left[g(\mathbf{x}^\star(t_f), \dot{\mathbf{x}}^\star(t_f), t_f) - g_{\dot{\mathbf{x}}}(\mathbf{x}^\star(t_f), \dot{\mathbf{x}}^\star(t_f), t_f)\dot{\mathbf{x}}^\star(t_f) \right] \delta t_f\end{aligned}$$

- Independent of the terminal constraint, the conditions on the solution $\mathbf{x}^\star(t)$ to be an extremal for this case are that it satisfy the Euler equations

$$g_{\mathbf{x}}(\mathbf{x}^\star(t), \dot{\mathbf{x}}^\star(t), t) - \frac{d}{dt} g_{\dot{\mathbf{x}}}(\mathbf{x}^\star(t), \dot{\mathbf{x}}^\star(t), t) = 0$$

  – Now consider the additional constraints on the individual elements of $\mathbf{x}^\star(t_f)$ and t_f to find the other boundary conditions

- Type of terminal constraints determines how we treat $\delta\mathbf{x}_f$ and δt_f
  1. Unrelated
  2. Related by a simple function $\mathbf{x}(t_f) = \Theta(t_f)$
  3. Specified by a more complex constraint $\mathbf{m}(\mathbf{x}(t_f), t_f) = 0$

- **Type 1:** If t_f and $\mathbf{x}(t_f)$ are free but unrelated, then $\delta\mathbf{x}_f$ and δt_f are independent and arbitrary $\Rightarrow$ their coefficients must both be zero.

$$\begin{aligned} g_{\mathbf{x}}(\mathbf{x}^\star(t), \dot{\mathbf{x}}^\star(t), t) - \frac{d}{dt} g_{\dot{\mathbf{x}}}(\mathbf{x}^\star(t), \dot{\mathbf{x}}^\star(t), t) &= 0 \\ g(\mathbf{x}^\star(t_f), \dot{\mathbf{x}}^\star(t_f), t_f) - g_{\dot{\mathbf{x}}}(\mathbf{x}^\star(t_f), \dot{\mathbf{x}}^\star(t_f), t_f)\dot{\mathbf{x}}^\star(t_f) &= 0 \\ g_{\dot{\mathbf{x}}}(\mathbf{x}^\star(t_f), \dot{\mathbf{x}}^\star(t_f), t_f) &= 0 \end{aligned}$$

  – Which makes it clear that this is a **two-point boundary value problem**, as we now have conditions at both t_0 and t_f

- **Type 2:** If t_f and $\mathbf{x}(t_f)$ are free but related as $\mathbf{x}(t_f) = \Theta(t_f)$, then

$$\delta \mathbf{x}_f = \frac{d\Theta}{dt}(t_f)\delta t_f$$

  – Substitute and collect terms gives

$$\begin{aligned}\delta J &= \int_{t_0}^{t_f} \left[g_{\mathbf{x}} - \frac{d}{dt} g_{\dot{\mathbf{x}}} \right] \delta \mathbf{x} dt + \left[g_{\dot{\mathbf{x}}}(\mathbf{x}^\star(t_f), \dot{\mathbf{x}}^\star(t_f), t_f) \frac{d\Theta}{dt}(t_f) \right. \\ &+ \left. g(\mathbf{x}^\star(t_f), \dot{\mathbf{x}}^\star(t_f), t_f) - g_{\dot{\mathbf{x}}}(\mathbf{x}^\star(t_f), \dot{\mathbf{x}}^\star(t_f), t_f)\dot{\mathbf{x}}^\star(t_f) \right] \delta t_f\end{aligned}$$

  – Set coefficient of δt_f to zero (it is arbitrary) $\Rightarrow$ full conditions

$$\begin{aligned} g_{\mathbf{x}}(\mathbf{x}^\star(t), \dot{\mathbf{x}}^\star(t), t) - \frac{d}{dt} g_{\dot{\mathbf{x}}}(\mathbf{x}^\star(t), \dot{\mathbf{x}}^\star(t), t) &= 0 \\ g_{\dot{\mathbf{x}}}(\mathbf{x}^\star(t_f), \dot{\mathbf{x}}^\star(t_f), t_f) \left[\frac{d\Theta}{dt}(t_f) - \dot{\mathbf{x}}^\star(t_f) \right] + g(\mathbf{x}^\star(t_f), \dot{\mathbf{x}}^\star(t_f), t_f) &= 0 \end{aligned}$$

  – Last equation called the **Transversality Condition**

- To handle third type of terminal condition, must address solution of constrained problems.

Image removed due to copyright restrictions.

Figure 5.4: Summary of possible terminal constraints (Kirk, page 151)

Example: 5–1

- Find the shortest curve from the origin to a specified line.

- **Goal:** minimize the cost functional (See page 5–6)

$$J = \int_{t_0}^{t_f} \sqrt{1 + \dot{x}^2(t)}\; dt$$

given that $t_0 = 0$, $x(0) = 0$, and t_f and $x(t_f)$ are free, but $x(t_f)$ must line on the line

$$\theta(t) = -5t + 15$$

- Since $g(x, \dot{x}, t)$ is only a function of $\dot{x}$, Euler equation reduces to

$$\frac{d}{dt}\left[\frac{\dot{x}^\star(t)}{[1 + \dot{x}^\star(t)^2]^{1/2}}\right] = 0$$

which after differentiating and simplifying, gives $\ddot{x}^\star(t) = 0 \Rightarrow$ answer is a straight line

$$x^\star(t) = c_1 t + c_0$$

but since $x(0) = 0$, then $c_0 = 0$

- Transversality condition gives

$$\left[\frac{\dot{x}^\star(t_f)}{[1 + \dot{x}^\star(t_f)^2]^{1/2}}\right][-5 - \dot{x}^\star(t_f)] + [1 + \dot{x}^\star(t_f)^2]^{1/2} = 0$$

that simplifies to

$$[\dot{x}^\star(t_f)]\,[-5 - \dot{x}^\star(t_f)] + [1 + \dot{x}^\star(t_f)^2] = -5\dot{x}^\star(t_f) + 1 = 0$$

so that $\dot{x}^\star(t_f) = c_1 = 1/5$

  – Not a surprise, as this gives the slope of a line orthogonal to the constraint line.

- To find final time: $x(t_f) = -5t_f + 15 = t_f/5$ which gives $t_f \approx 2.88$

Example: 5–2

- Had the terminal constraint been a bit more challenging, such as

$$\Theta(t) = \frac{1}{2}([t-5]^2 - 1) \Rightarrow \frac{d\Theta}{dt} = t - 5$$

- Then the transversality condition gives

$$\begin{aligned}
\left[\frac{\dot{x}^\star(t_f)}{[1+\dot{x}^\star(t_f)^2]^{1/2}}\right][t_f - 5 - \dot{x}^\star(t_f)] + [1+\dot{x}^\star(t_f)^2]^{1/2} &= 0 \\
[\dot{x}^\star(t_f)]\,[t_f - 5 - \dot{x}^\star(t_f)] + [1+\dot{x}^\star(t_f)^2] &= 0 \\
c_1\,[t_f - 5] + 1 &= 0
\end{aligned}$$

- Now look at $x^\star(t)$ and $\Theta(t)$ at t_f

$$x^\star(t_f) = -\frac{t_f}{(t_f - 5)} = \frac{1}{2}([t_f - 5]^2 - 1)$$

which gives $t_f = 3$, $c_1 = 1/2$ and $x^\star(t_f) = t/2$

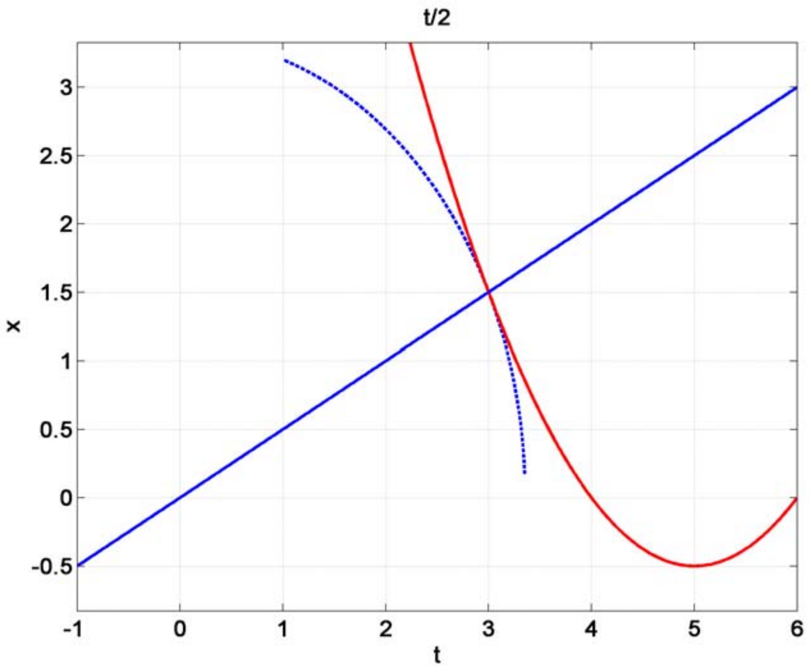


Figure 5.5: Quadratic terminal constraint.

Corner Conditions

- Key generalization of the preceding is to allow the possibility that the solutions not be as **smooth**
  - Assume that $\mathbf{x}(t)$ cts, but allow discontinuities in $\dot{\mathbf{x}}(t)$, which occur at **corners**.
  - Naturally occur when intermediate state constraints imposed, or with jumps in the control signal.

- **Goal:** with t_0, t_f, $\mathbf{x}(t_0)$, and $\mathbf{x}(t_f)$ fixed, minimize cost functional

$$J(\mathbf{x}(t),t) = \int_{t_0}^{t_f} g(\mathbf{x}(t),\dot{\mathbf{x}}(t),t)dt$$

  - Assume g has cts first/second derivatives wrt all arguments
  - Even so, $\dot{\mathbf{x}}$ discontinuity could lead to a discontinuity in g.

- Assume that $\dot{\mathbf{x}}$ has a discontinuity at some time $t_1 \in (t_0, t_f)$, which is not fixed (or typically known). Divide cost into 2 regions:

$$J(\mathbf{x}(t),t) = \int_{t_0}^{t_1} g(\mathbf{x}(t),\dot{\mathbf{x}}(t),t)dt + \int_{t_1}^{t_f} g(\mathbf{x}(t),\dot{\mathbf{x}}(t),t)dt$$

- Expand as before – note that t_1 is not fixed

$$\begin{aligned}\delta J \;&=\; \int_{t_0}^{t_1} \left[\frac{\partial g}{\partial \mathbf{x}}\delta\mathbf{x} + \frac{\partial g}{\partial \dot{\mathbf{x}}}\delta\dot{\mathbf{x}}\right] dt + g(t_1^-)\delta t_1 \\ &\quad + \int_{t_1}^{t_f} \left[\frac{\partial g}{\partial \mathbf{x}}\delta\mathbf{x} + \frac{\partial g}{\partial \dot{\mathbf{x}}}\delta\dot{\mathbf{x}}\right] dt - g(t_1^+)\delta t_1\end{aligned}$$

- Now IBP

$$\begin{aligned}\delta J &= \int_{t_0}^{t_1} \left[g_{\mathbf{x}} - \frac{d}{dt}(g_{\dot{\mathbf{x}}}) \right] \delta\mathbf{x} dt + g(t_1^-)\delta t_1 + g_{\dot{\mathbf{x}}}(t_1^-)\delta\mathbf{x}(t_1^-) \\ &+ \int_{t_1}^{t_f} \left[g_{\mathbf{x}} - \frac{d}{dt}(g_{\dot{\mathbf{x}}}) \right] \delta\mathbf{x} dt - g(t_1^+)\delta t_1 - g_{\dot{\mathbf{x}}}(t_1^+)\delta\mathbf{x}(t_1^+)\end{aligned}$$

- As on 5–9, must constrain $\delta\mathbf{x}_1$, which is the total variation in the solution at time t_1

$$\begin{aligned}\text{from lefthand side} \quad \delta\mathbf{x}_1 &= \delta\mathbf{x}(t_1^-) + \dot{\mathbf{x}}(t_1^-)\delta t_1 \\ \text{from righthand side} \quad \delta\mathbf{x}_1 &= \delta\mathbf{x}(t_1^+) + \dot{\mathbf{x}}(t_1^+)\delta t_1\end{aligned}$$

  - Continuity requires that these two expressions for $\delta\mathbf{x}_1$ be equal
  - Already know that it is possible that $\dot{\mathbf{x}}(t_1^-) \neq \dot{\mathbf{x}}(t_1^+)$, so possible that $\delta\mathbf{x}(t_1^-) \neq \delta\mathbf{x}(t_1^+)$ as well.

- Substitute:

$$\begin{aligned}\delta J &= \int_{t_0}^{t_1} \left[g_{\mathbf{x}} - \frac{d}{dt}(g_{\dot{\mathbf{x}}}) \right] \delta\mathbf{x} dt + \left[g(t_1^-) - g_{\dot{\mathbf{x}}}(t_1^-)\dot{\mathbf{x}}(t_1^-) \right] \delta t_1 + g_{\dot{\mathbf{x}}}(t_1^-)\delta\mathbf{x}_1 \\ &+ \int_{t_1}^{t_f} \left[g_{\mathbf{x}} - \frac{d}{dt}(g_{\dot{\mathbf{x}}}) \right] \delta\mathbf{x} dt - \left[g(t_1^+) - g_{\dot{\mathbf{x}}}(t_1^+)\dot{\mathbf{x}}(t_1^+) \right] \delta t_1 - g_{\dot{\mathbf{x}}}(t_1^+)\delta\mathbf{x}_1\end{aligned}$$

- Necessary conditions are then:

$$\boxed{\begin{aligned} g_{\mathbf{x}} - \frac{d}{dt}(g_{\dot{\mathbf{x}}}) &= 0 \quad t \in (t_0, t_f) \\ g_{\dot{\mathbf{x}}}(t_1^-) &= g_{\dot{\mathbf{x}}}(t_1^+) \\ g(t_1^-) - g_{\dot{\mathbf{x}}}(t_1^-)\dot{\mathbf{x}}(t_1^-) &= g(t_1^+) - g_{\dot{\mathbf{x}}}(t_1^+)\dot{\mathbf{x}}(t_1^+) \end{aligned}}$$

  - Last two are the **Weierstrass-Erdmann** conditions

- Necessary conditions given for a special set of the terminal conditions, but the form of the internal conditions unchanged by different terminal constraints
  - With several corners, there are a set of constraints for each
  - Can be used to demonstrate that there isn't a corner

- Typical instance that induces corners is intermediate time constraints of the form $\mathbf{x}(t_1) = \boldsymbol{\theta}(t_1)$.
  - i.e., the solution must touch a specified curve at some point in time during the solution.

- Slightly complicated in this case, because the constraint couples the allowable variations in $\delta\mathbf{x}_1$ and δt since

$$\delta\mathbf{x}_1 = \frac{d\boldsymbol{\theta}}{dt}\delta t_1 = \dot{\boldsymbol{\theta}}\delta t_1$$

  - But can eliminate $\delta\mathbf{x}_1$ in favor of δt_1 in the expression for δJ to get new corner condition:

$$g(t_1^-)+g_{\dot{\mathbf{x}}}(t_1^-)\left[\dot{\boldsymbol{\theta}}(t_1^-)-\dot{\mathbf{x}}(t_1^-)\right] = g(t_1^+)+g_{\dot{\mathbf{x}}}(t_1^+)\left[\dot{\boldsymbol{\theta}}(t_1^+)-\dot{\mathbf{x}}(t_1^+)\right]$$

  - So now $g_{\dot{\mathbf{x}}}(t_1^-) = g_{\dot{\mathbf{x}}}(t_1^+)$ no longer needed, but have $\mathbf{x}(t_1) = \boldsymbol{\theta}(t_1)$

Corner Example

- Find shortest length path joining the points $x = 0, t = -2$ and $x = 0, t = 1$ that touches the curve $x = t^2 + 3$ at some point

- In this case, $J = \int_{-2}^{1} \sqrt{1 + \dot{x}^2} dt$ with $x(1) = x(-2) = 0$ and $x(t_1) = t_1^2 + 3$

- Note that since g is only a function of $\dot{x}$, then solution $x(t)$ will only be linear in each segment (see 5–13)

$$\begin{aligned}\text{segment 1} \quad x(t) &= a + bt \\ \text{segment 2} \quad x(t) &= c + dt\end{aligned}$$

  – Terminal conditions: $x(-2) = a - 2b = 0$ and $x(1) = c + d = 0$

- Apply corner condition:

$$\begin{aligned}\sqrt{1 + \dot{x}(t_1^-)^2} \; + \; \frac{\dot{x}(t_1^-)}{\sqrt{1 + \dot{x}(t_1^-)^2}} \left[2t_1^- - \dot{x}(t_1^-)\right] \\ = \frac{1 + 2t_1^- \dot{x}(t_1^-)}{\sqrt{1 + \dot{x}(t_1^-)^2}} = \frac{1 + 2t_1^+ \dot{x}(t_1^+)}{\sqrt{1 + \dot{x}(t_1^+)^2}}\end{aligned}$$

  which gives:

$$\frac{1 + 2bt_1}{\sqrt{1 + b^2}} = \frac{1 + 2dt_1}{\sqrt{1 + d^2}}$$

- Solve using `fsolve` to get:

$$a = 3.0947, b = 1.5474, c = 2.8362, d = -2.8362, t_1 = -0.0590$$

```
function F=myfunc(x); %
% x=[a b c d t1]; %
F=[x(1)-2*x(2);
    x(3)+x(4);
    (1+2*x(2)*x(5))/(1+x(2)^2)^(1/2) - (1+2*x(4)*x(5))/(1+x(4)^2)^(1/2);
    x(1)+x(2)*x(5) - (x(5)^2+3);
    x(3)+x(4)*x(5) - (x(5)^2+3)];
return %
x = fsolve('myfunc',[2 1 2 -2 0]')
```

Constrained Solutions

- Now consider variations of the basic problem that include constraints.
- For example, if the goal is to find the extremal function $\mathbf{x}^\star$ that minimizes

$$J(\mathbf{x}(t), t) = \int_{t_0}^{t_f} g(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)dt$$

subject to the constraint that a given set of n differential equations be satisfied

$$\mathbf{f}(\mathbf{x}(t), \dot{\mathbf{x}}(t), t) = 0$$

where we assume that $\mathbf{x} \in \mathcal{R}^{n+m}$ (take t_f and $\mathbf{x}(t_f)$ to be fixed)

- As with the basic optimization problems in Lecture 2, proceed by augmenting cost with the constraints using Lagrange multipliers
  - Since the constraints must be satisfied at all time, these multipliers are also assumed to be functions of time.

$$J_a(\mathbf{x}(t), t) = \int_{t_0}^{t_f} \left\{ g(\mathbf{x}, \dot{\mathbf{x}}, t) + \mathbf{p}(t)^T \mathbf{f}(\mathbf{x}, \dot{\mathbf{x}}, t) \right\} dt$$

  - Does not change the cost if the constraints are satisfied.
  - Time varying Lagrange multipliers give more degrees of freedom in specifying how the constraints are added.

- Take variation of augmented functional considering perturbations to both $\mathbf{x}(t)$ and $\mathbf{p}(t)$

$$\delta J(\mathbf{x}(t), \delta\mathbf{x}(t), \mathbf{p}(t), \delta\mathbf{p}(t))$$
$$= \int_{t_0}^{t_f} \left\{ \left[g_{\mathbf{x}} + \mathbf{p}^T \mathbf{f}_{\mathbf{x}} \right] \delta\mathbf{x}(t) + \left[g_{\dot{\mathbf{x}}} + \mathbf{p}^T \mathbf{f}_{\dot{\mathbf{x}}} \right] \delta\dot{\mathbf{x}}(t) + \mathbf{f}^T \delta\mathbf{p}(t) \right\} dt$$

- As before, integrate by parts to get:

$$\delta J(\mathbf{x}(t), \delta\mathbf{x}(t), \mathbf{p}(t), \delta\mathbf{p}(t))$$
$$= \int_{t_0}^{t_f} \left(\left\{ [\, g_{\mathbf{x}} + \mathbf{p}^T \mathbf{f}_{\mathbf{x}}] - \frac{d}{dt} [g_{\dot{\mathbf{x}}} + \mathbf{p}^T \mathbf{f}_{\dot{\mathbf{x}}}] \right\} \delta\mathbf{x}(t) + \mathbf{f}^T \delta\mathbf{p}(t) \right) dt$$

- To simplify things a bit, define

$$g_a(\mathbf{x}(t), \dot{\mathbf{x}}(t), t) \equiv g(\mathbf{x}(t), \dot{\mathbf{x}}(t), t) + \mathbf{p}(t)^T \mathbf{f}(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)$$

- On the extremal, the variation must be zero, but since $\delta\mathbf{x}(t)$ and $\delta\mathbf{p}(t)$ can be arbitrary, can only occur if

$$\boxed{\begin{aligned} \frac{\partial g_a(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)}{\partial \mathbf{x}} - \frac{d}{dt}\left(\frac{\partial g_a(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)}{\partial \dot{\mathbf{x}}} \right) &= 0 \\ \mathbf{f}(\mathbf{x}(t), \dot{\mathbf{x}}(t), t) &= 0 \end{aligned}}$$

  - which are obviously a generalized version of the Euler equations obtained before.

- Note similarity of the definition of g_a here with the Hamiltonian on page 4–4.

- Will find that this generalization carries over to future optimizations as well.

General Terminal Conditions

- Now consider Type 3 constraints on 5–10, which are a very general form with t_f free and $\mathbf{x}(t_f)$ given by a condition:

$$\mathbf{m}(\mathbf{x}(t_f), t_f) = 0$$

- Constrained optimization, so as before, augment the cost functional

$$J(\mathbf{x}(t), t) = h(\mathbf{x}(t_f), t_f) + \int_{t_0}^{t_f} g(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)dt$$

with the constraint using Lagrange multipliers:

$$J_a(\mathbf{x}(t), \boldsymbol{\nu}, t) = h(\mathbf{x}(t_f), t_f) + \boldsymbol{\nu}^T \mathbf{m}(\mathbf{x}(t_f), t_f) + \int_{t_0}^{t_f} g(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)dt$$

- Considering changes to $\mathbf{x}(t)$, t_f, $\mathbf{x}(t_f)$ and $\boldsymbol{\nu}$, the variation for J_a is

$$\begin{aligned} \delta J_a &= h_{\mathbf{x}}(t_f)\delta\mathbf{x}_f + h_{t_f}\delta t_f + \mathbf{m}^T(t_f)\delta\boldsymbol{\nu} + \boldsymbol{\nu}^T \Big(\mathbf{m}_{\mathbf{x}}(t_f)\delta\mathbf{x}_f + \mathbf{m}_{t_f}(t_f)\delta t_f\Big) \\ &\quad + \int_{t_0}^{t_f} \left[g_{\mathbf{x}}\delta\mathbf{x} + g_{\dot{\mathbf{x}}}\delta\dot{\mathbf{x}}\right] dt + g(t_f)\delta t_f \\ &= \left[h_{\mathbf{x}}(t_f) + \boldsymbol{\nu}^T\mathbf{m}_{\mathbf{x}}(t_f)\right]\delta\mathbf{x}_f + \left[h_{t_f} + \boldsymbol{\nu}^T\mathbf{m}_{t_f}(t_f) + g(t_f)\right]\delta t_f \\ &\quad + \mathbf{m}^T(t_f)\delta\boldsymbol{\nu} + \int_{t_0}^{t_f}\left[g_{\mathbf{x}} - \frac{d}{dt}g_{\dot{\mathbf{x}}}\right]\delta\mathbf{x}dt + g_{\dot{\mathbf{x}}}(t_f)\delta\mathbf{x}(t_f) \end{aligned}$$

– Now use that $\delta\mathbf{x}_f = \delta\mathbf{x}(t_f) + \dot{\mathbf{x}}(t_f)\delta t_f$ as before to get

$$\begin{aligned} \delta J_a &= \left[h_{\mathbf{x}}(t_f) + \boldsymbol{\nu}^T\mathbf{m}_{\mathbf{x}}(t_f) + g_{\dot{\mathbf{x}}}(t_f)\right]\delta\mathbf{x}_f \\ &\quad + \left[h_{t_f} + \boldsymbol{\nu}^T\mathbf{m}_{t_f}(t_f) + g(t_f) - g_{\dot{\mathbf{x}}}(t_f)\dot{\mathbf{x}}(t_f)\right]\delta t_f + \mathbf{m}^T(t_f)\delta\boldsymbol{\nu} \\ &\quad + \int_{t_0}^{t_f}\left[g_{\mathbf{x}} - \frac{d}{dt}g_{\dot{\mathbf{x}}}\right]\delta\mathbf{x}dt \end{aligned}$$

- Looks like a bit of a mess, but we can clean it up a bit using

$$w(\mathbf{x}(t_f), \boldsymbol{\nu}, t_f) = h(\mathbf{x}(t_f), t_f) + \boldsymbol{\nu}^T \mathbf{m}(\mathbf{x}(t_f), t_f)$$

to get

$$\begin{aligned}
\delta J_a &= \left[w_{\mathbf{x}}(t_f) + g_{\dot{\mathbf{x}}}(t_f)\right] \delta \mathbf{x}_f \\
&+ \left[w_{t_f} + g(t_f) - g_{\dot{\mathbf{x}}}(t_f)\dot{\mathbf{x}}(t_f)\right] \delta t_f + \mathbf{m}^T(t_f)\delta\boldsymbol{\nu} \\
&+ \int_{t_0}^{t_f} \left[g_{\mathbf{x}} - \frac{d}{dt} g_{\dot{\mathbf{x}}}\right] \delta \mathbf{x} dt
\end{aligned}$$

  - Given the extra degrees of freedom in the multipliers, can treat all of the variations as independent ⇒ all coefficients must be zero to achieve $\delta J_a = 0$

- So the necessary conditions are

$$\begin{aligned}
g_{\mathbf{x}} - \frac{d}{dt} g_{\dot{\mathbf{x}}} &= 0 \quad (\texttt{dim}\ n) \\
w_{\mathbf{x}}(t_f) + g_{\dot{\mathbf{x}}}(t_f) &= 0 \quad (\texttt{dim}\ n) \\
w_{t_f} + g(t_f) - g_{\dot{\mathbf{x}}}(t_f)\dot{\mathbf{x}}(t_f) &= 0 \quad (\texttt{dim}\ 1)
\end{aligned}$$

  - With $\mathbf{x}(t_0) = \mathbf{x}_0$ ($\texttt{dim}\ n$) and $\mathbf{m}(\mathbf{x}(t_f), t_f) = 0$ ($\texttt{dim}\ m$) combined with last 2 conditions ⇒ $2n + m + 1$ constraints
  - Solution of Eulers equation has $2n$ constants of integration for $x(t)$, and must find $\boldsymbol{\nu}$ ($\texttt{dim}\ m$) and t_f ⇒ $2n + m + 1$ unknowns

- Some special cases:
  - If t_f is fixed, $h = h(\mathbf{x}(t_f))$, $\mathbf{m} \to \mathbf{m}(\mathbf{x}(t_f))$ and we lose the last condition in box – others remain unchanged
  - If t_f is fixed, $\mathbf{x}(t_f)$ free, then there is no $\mathbf{m}$, no $\boldsymbol{\nu}$ and w reduces to h.

- Kirk's book also considers several other type of constraints.