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.
**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 $x(t)$ and $u(t)$.

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

subject to

$$
\dot{x} = f(x, u, t) \\
x(t_0), t_0 \text{ given} \\
m(x(t_f), t_f) = 0
$$

- Call $J(x(t), u(t))$ 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 \( f(x) \) has a local minimum at \( x^* \) if
  
  \[
  f(x) \geq f(x^*)
  \]
  
  for all admissible \( x \) in \( \|x - x^*\| \leq \epsilon \)
  
  – 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 \( f(x) \), then statement equivalent to requiring that differential of \( f \) satisfy:

  \[
  df = \frac{\partial f}{\partial x} dx = 0
  \]
  
  for all small \( dx \), which gives the same necessary condition from Lecture 1
  
  \[
  \frac{\partial f}{\partial x} = 0
  \]
  
  • Note that this definition used norms to compare two vectors. Can do the same thing with functions ⇒ distance between two functions

  \[
  d = \| x_2(t) - x_1(t) \|
  \]

  where

  1. \( \| x(t) \| \geq 0 \) for all \( x(t) \), and \( \| x(t) \| = 0 \) only if \( x(t) = 0 \) for all \( t \) in the interval of definition.
  2. \( \| ax(t) \| = |a| \| x(t) \| \) for all real scalars \( a \).
  3. \( \| x_1(t) + x_2(t) \| \leq \| x_1(t) \| + \| x_2(t) \| \)

• **Common function norm:**

  \[
  \| x(t) \|_2 = \left( \int_{t_0}^{t_f} x(t)^T x(t) dt \right)^{1/2}
  \]
• **Maximum and Minimum of a Functional**
  
  – A functional $J(x(t))$ has a local minimum at $x^*(t)$ if

  $$J(x(t)) \geq J(x^*(t))$$

  for all admissible $x(t)$ in $\|x(t) - x^*(t)\| \leq \epsilon$

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

  – An **increment** of a functional

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

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

  $$\Delta J(x(t), \delta x(t)) = \delta J(x(t), \delta x(t)) + H.O.T.$$ i.e. $\delta J(x(t), \delta x(t))$ is linear in $\delta x(t)$.

---

**Figure 5.1:** Differential $df$ versus increment $\Delta f$ shown for a function, but the same difference holds for a functional.

---

June 18, 2008
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 \(x\) be a function of \(t\) in the class \(\Omega\), and \(J(x)\) be a differentiable functional of \(x\). Assume that the functions in \(\Omega\) are not constrained by any boundaries.
  - If \(x^*\) is an extremal function, then the variation of \(J\) must vanish on \(x^*\), i.e. for all admissible \(\delta x\),
    \[
    \delta J(x(t), \delta x(t)) = 0
    \]
  - Proof is in Kirk, page 121, but it is relatively straightforward.

- How compute the variation? If \(J(x(t)) = \int_{t_0}^{t_f} f(x(t))\,dt\) where \(f\) has cts first and second derivatives with respect to \(x\), then
  \[
  \delta J(x(t), \delta x) = \int_{t_0}^{t_f} \left\{ \frac{\partial f(x(t))}{\partial x(t)} \right\} \delta x \, dt + f(x(t_f)) \delta t_f - f(x(t_0)) \delta t_0 \\
  = \int_{t_0}^{t_f} f_x(x(t)) \delta x \, dt + f(x(t_f)) \delta t_f - f(x(t_0)) \delta t_0
  \]
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.

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

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

Note that \( \delta \dot{x} = \frac{d}{dt} \delta x \) so \( \delta x \) and \( \delta \dot{x} \) are not independent.

- Integrate by parts:

\[
\int uv \equiv uv - \int v du
\]

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

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

\[- \int_{t_0}^{t_f} \frac{d}{dt} g_{\dot{x}}(x(t), \dot{x}(t), t)\delta x dt\]

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

\[+ \left[ g_{\dot{x}}(x(t), \dot{x}(t), t)\delta x \right]_{t_0}^{t_f}\]

- 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:
  \[
  J = \int_{x_0}^{x_f} ds = \int_{x_0}^{x_f} \sqrt{(dx)^2 + (dy)^2} \\
  = \int_{x_0}^{x_f} \sqrt{1 + \left( \frac{dy}{dx} \right)^2} \, dx
  \]

- Take $y$ as dependent variable, and $x$ as independent one
  \[
  \frac{dy}{dx} \rightarrow \dot{y}
  \]

- New form of the cost:
  \[
  J = \int_{x_0}^{x_f} \sqrt{1 + \dot{y}^2} \, dx \rightarrow \int_{x_0}^{x_f} g(\dot{y}) \, dx
  \]

- Take partials: $\frac{\partial g}{\partial y} = 0$, and
  \[
  \frac{d}{dx} \left( \frac{\partial g}{\partial \dot{y}} \right) = \frac{d}{dy} \left( \frac{\partial g}{\partial \dot{y}} \right) \frac{d\dot{y}}{dx}
  = \frac{d}{dy} \left( \frac{\dot{y}}{(1 + \dot{y}^2)^{1/2}} \right) \dot{y} = \frac{\dot{y}}{(1 + \dot{y}^2)^{3/2}} = 0
  \]
  which implies that $\ddot{y} = 0$

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

June 18, 2008
Vector Functions

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

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

with \(t_0, t_f, x(t_0)\) fixed.

• Then (drop the arguments for brevity)

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

– Integrate by parts to get:

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

• The requirement then is that for \(t \in (t_0, t_f)\), \(x(t)\) must satisfy

\[
\frac{\partial g}{\partial x} - \frac{d}{dt} \frac{\partial g}{\partial \dot{x}} = 0
\]

where \(x(t_0) = x_0\) which are the given \(N\) boundary conditions, and

the remaining \(N\) more BC follow from:

– \(x(t_f) = x_f\) if \(x_f\) is given as fixed,

– If \(x(t_f)\) are free, then

\[
\frac{\partial g(x(t), \dot{x}(t), t)}{\partial \dot{x}(t_f)} = 0
\]

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

June 18, 2008
**Free Terminal Time**

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

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

with \( t_0, x(t_0) \) fixed, \( t_f \) free, and various constraints on \( x(t_f) \)

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

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

  - Integrate by parts to get:

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

\[
+ \ g_{\dot{x}}(x^*(t_f), \dot{x}^*(t_f), t_f) \delta x(t_f)
\]

\[
+ \ g(x^*(t_f), \dot{x}^*(t_f), t_f) \delta t_f
\]

- Looks standard so far, but we have to be careful how we handle the terminal conditions
Figure 5.3: Comparison of possible changes to function at end time when $t_f$ is free.

- By definition, $\delta x(t_f)$ is the difference between two admissible functions at time $t_f$ (in this case the optimal solution $x^*$ and another candidate $x$).
  - But in this case, must also account for possible changes to $\delta t_f$.
  - Define $\delta x_f$ as being the difference between the ends of the two possible functions – **total possible change** in the final state:
    \[
    \delta x_f \approx \delta x(t_f) + x^*(t_f) \delta t_f
    \]
    so $\delta x(t_f) \neq \delta x_f$ in general.

- Substitute to get
  \[
  \delta J(x^*(t), \delta x) = \int_{t_0}^{t_f} \left[ g_x - \frac{d}{dt} g_{\dot{x}} \right] \delta x(t) dt + g_{\dot{x}}(x^*(t_f), \dot{x}^*(t_f), t_f) \delta x_f
  \]
  \[
  + [g(x^*(t_f), \dot{x}^*(t_f), t_f) - g_x(x^*(t_f), \dot{x}^*(t_f), t_f) \dot{x}^*(t_f)] \delta t_f
  \]
Independent of the terminal constraint, the conditions on the solution $x^*(t)$ to be an extremal for this case are that it satisfy the Euler equations

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

Now consider the additional constraints on the individual elements of $x^*(t_f)$ and $t_f$ to find the other boundary conditions.

Type of terminal constraints determines how we treat $\delta x_f$ and $\delta t_f$

1. **Unrelated**
   - $x(t_f) = \Theta(t_f)$

2. **Related by a simple function**
   - $x(t_f) = \Theta(t_f)$

3. **Specified by a more complex constraint**
   - $m(x(t_f), t_f) = 0$

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

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

$$g(x^*(t_f), \dot{x}^*(t_f), t_f) - g_{\dot{x}}(x^*(t_f), \dot{x}^*(t_f), t_f)\dot{x}^*(t_f) = 0$$

$$g_{\dot{x}}(x^*(t_f), \dot{x}^*(t_f), t_f) = 0$$

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

---

June 18, 2008
• **Type 2:** If $t_f$ and $x(t_f)$ are free but related as $x(t_f) = \Theta(t_f)$, then

$$\delta x_f = \frac{d\Theta}{dt}(t_f)\delta t_f$$

– Substitute and collect terms gives

$$\delta J = \int_{t_0}^{t_f} \left[ g_x - \frac{d}{dt}g_x \right] \delta x dt + \left[ g_x(x^*(t_f), \dot{x}^*(t_f), t_f) \frac{d\Theta}{dt}(t_f) \right. \\
\left. + g(x^*(t_f), \dot{x}^*(t_f), t_f) - g_x(x^*(t_f), \dot{x}^*(t_f), t_f)\dot{x}^*(t_f) \right] \delta t_f$$

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

$$g_x(x^*(t), \dot{x}^*(t), t) - \frac{d}{dt}g_x(x^*(t), \dot{x}^*(t), t) = 0$$

$$g_x(x^*(t_f), \dot{x}^*(t_f), t_f) \left[ \frac{d\Theta}{dt}(t_f) - \dot{x}^*(t_f) \right] + g(x^*(t_f), \dot{x}^*(t_f), t_f) = 0$$

– Last equation called the **Transversality Condition**

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

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

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

  \[
  x^*(t) = c_1 t + c_0
  \]

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

- Transversality condition gives

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

  that simplifies to

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

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

  – 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
  \[
  \left[ \frac{\dot{x}^*(t_f)}{[1 + \dot{x}^*(t_f)^2]^{1/2}} \right] [t_f - 5 - \dot{x}^*(t_f)] + [1 + \dot{x}^*(t_f)^2]^{1/2} = 0 \\
  [\dot{x}^*(t_f)] [t_f - 5 - \dot{x}^*(t_f)] + [1 + \dot{x}^*(t_f)^2] = 0 \\
  c_1 [t_f - 5] + 1 = 0
  \]

- Now look at \( x^*(t) \) and \( \Theta(t) \) at \( t_f \)
  \[ x^*(t_f) = -\frac{t_f}{(t_f - 5)} = \frac{1}{2}([t_f - 5]^2 - 1) \]
  which gives \( t_f = 3 \), \( c_1 = 1/2 \) and \( x^*(t_f) = t/2 \)

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 $x(t)$ cts, but allow discontinuities in $\dot{x}(t)$, which occur at **corners**.
  – Naturally occur when intermediate state constraints imposed, or with jumps in the control signal.

• **Goal:** with $t_0$, $t_f$, $x(t_0)$, and $x(t_f)$ fixed, minimize cost functional

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

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

• Assume that $\dot{x}$ 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:

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

• Expand as before – note that $t_1$ is not fixed

$$\delta J = \int_{t_0}^{t_1} \left[ \frac{\partial g}{\partial x} \delta x + \frac{\partial g}{\partial \dot{x}} \delta \dot{x} \right] dt + g(t_1^-) \delta t_1$$

$$+ \int_{t_1}^{t_f} \left[ \frac{\partial g}{\partial x} \delta x + \frac{\partial g}{\partial \dot{x}} \delta \dot{x} \right] dt - g(t_1^+) \delta t_1$$
Now IBP

\[
\delta J = \int_{t_0}^{t_1} \left[ g_x - \frac{d}{dt} (g_x) \right] \delta x dt + g(t_1^-) \delta t_1 + g_x(t_1^-) \delta x(t_1^-) + \int_{t_1}^{t_f} \left[ g_x - \frac{d}{dt} (g_x) \right] \delta x dt - g(t_1^+) \delta t_1 - g_x(t_1^+) \delta x(t_1^+)
\]

As on 5–9, must constrain \( \delta x_1 \), which is the total variation in the solution at time \( t_1 \)

- Continuity requires that these two expressions for \( \delta x_1 \) be equal
- Already know that it is possible that \( \dot{x}(t_1^-) \neq \dot{x}(t_1^+) \), so possible that \( \delta x(t_1^-) \neq \delta x(t_1^+) \) as well.

Substitute:

\[
\delta J = \int_{t_0}^{t_1} \left[ g_x - \frac{d}{dt} (g_x) \right] \delta x dt + \left[ g(t_1^-) - g_x(t_1^-) \dot{x}(t_1^-) \right] \delta t_1 + g_x(t_1^-) \delta x(t_1^-) + \int_{t_1}^{t_f} \left[ g_x - \frac{d}{dt} (g_x) \right] \delta x dt - \left[ g(t_1^+) - g_x(t_1^+) \dot{x}(t_1^+) \right] \delta t_1 - g_x(t_1^+) \delta x(t_1^+)
\]

Necessary conditions are then:

\[
\begin{align*}
    g_x - \frac{d}{dt} (g_x) &= 0 \quad t \in (t_0, t_f) \\
    g_x(t_1^-) &= g_x(t_1^+) \\
    g(t_1^-) - g_x(t_1^-) \dot{x}(t_1^-) &= g(t_1^+) - g_x(t_1^+) \dot{x}(t_1^+)
\end{align*}
\]

- 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 $x(t_1) = \theta(t_1)$.

- 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 $\delta x_1$ and $\delta t$ since

$$\delta x_1 = \frac{d\theta}{dt} \delta t_1 = \dot{\theta} \delta t_1$$

- But can eliminate $\delta x_1$ in favor of $\delta t_1$ in the expression for $\delta J$ to get new corner condition:

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

- So now $g_x(t_1^-) = g_x(t_1^+)$ no longer needed, but have $x(t_1) = \theta(t_1)$
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)

\[
\text{segment 1} \quad x(t) = a + bt \\
\text{segment 2} \quad x(t) = c + dt
\]

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

- Apply corner condition:

\[
\sqrt{1 + \dot{x}(t_1^-)^2} + \frac{\dot{x}(t_1^-)}{\sqrt{1 + \dot{x}(t_1^-)^2}} \left[ 2t_1^- - \dot{x}(t_1^-) \right]
\]

\[
= \frac{1 + 2t_1^- \dot{x}(t_1^-)}{\sqrt{1 + \dot{x}(t_1^-)^2}} = \frac{1 + 2t_1^+ \dot{x}(t_1^+)}{\sqrt{1 + \dot{x}(t_1^+)^2}}
\]

which gives:

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

- Solve using \texttt{fsolve} to get:

\[
a = 3.0947, b = 1.5474, c = 2.8362, d = -2.8362, t_1 = -0.0590
\]

```matlab
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])
```

June 18, 2008
Constrained Solutions

Now consider variations of the basic problem that include constraints.

For example, if the goal is to find the extremal function \( x^* \) that minimizes

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

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

\[
f(x(t), \dot{x}(t), t) = 0
\]

where we assume that \( x \in \mathbb{R}^{n+m} \) (take \( t_f \) and \( x(t_f) \) 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.

\[
J_a(x(t), t) = \int_{t_0}^{t_f} \{ g(x, \dot{x}, t) + p(t)^T f(x, \dot{x}, t) \} \, dt
\]

- 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 \( x(t) \) and \( p(t) \)

\[
\delta J(x(t), \delta x(t), p(t), \delta p(t))
\]

\[
= \int_{t_0}^{t_f} \left\{ \left[ g_x + p^T f_x \right] \delta x(t) + \left[ g_x + p^T f_x \right] \delta \dot{x}(t) + f^T \delta p(t) \right\} \, dt
\]
As before, integrate by parts to get:

\[
\delta J(x(t), \dot{x}(t), p(t), \delta p(t))
\]

\[
= \int_{t_0}^{t_f} \left( \left\{ [g_x + p^T f_x] - \frac{d}{dt} [g_x + p^T f_x] \right\} \delta x(t) + f^T \delta p(t) \right) dt
\]

To simplify things a bit, define

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

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

\[
\frac{\partial g_a(x(t), \dot{x}(t), t)}{\partial x} - \frac{d}{dt} \left( \frac{\partial g_a(x(t), \dot{x}(t), t)}{\partial \dot{x}} \right) = 0
\]

\[
f(x(t), \dot{x}(t), t) = 0
\]

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 \( x(t_f) \) given by a condition:

\[
m(x(t_f), t_f) = 0
\]

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

\[
J(x(t), t) = h(x(t_f), t_f) + \int_{t_0}^{t_f} g(x(t), x'(t), t) dt
\]

with the constraint using Lagrange multipliers:

\[
J_a(x(t), \nu, t) = h(x(t_f), t_f) + \nu^T m(x(t_f), t_f) + \int_{t_0}^{t_f} g(x(t), x'(t), t) dt
\]

- Considering changes to \( x(t), t_f, x(t_f) \) and \( \nu \), the variation for \( J_a \) is

\[
\delta J_a = h_x(t_f) \delta x_f + h_{t_f} \delta t_f + m(t_f)^T \delta \nu + \nu^T \left( m_x(t_f) \delta x_f + m_{t_f}(t_f) \delta t_f \right) + \int_{t_0}^{t_f} \left[ g_x \delta x + g_x \delta \dot{x} \right] dt + g(t_f) \delta t_f
\]

\[
= \left[ h_x(t_f) + \nu^T m_x(t_f) \right] \delta x_f + \left[ h_{t_f} + \nu^T m_{t_f}(t_f) + g(t_f) \right] \delta t_f + m(t_f)^T \delta \nu + \int_{t_0}^{t_f} \left[ g_x - \frac{d}{dt} g_x \right] \delta x dt + g_x(t_f) \delta x(t_f)
\]

- Now use that \( \delta x_f = \delta x(t_f) + \dot{x}(t_f) \delta t_f \) as before to get

\[
\delta J_a = \left[ h_x(t_f) + \nu^T m_x(t_f) + g_x(t_f) \right] \delta x_f + \left[ h_{t_f} + \nu^T m_{t_f}(t_f) + g(t_f) - g_x(t_f) \dot{x}(t_f) \right] \delta t_f + m(t_f)^T \delta \nu + \int_{t_0}^{t_f} \left[ g_x - \frac{d}{dt} g_x \right] \delta x dt
\]
• Looks like a bit of a mess, but we can clean it up a bit using

\[ w(x(tf), \nu, tf) = h(x(tf), tf) + \nu^T m(x(tf), tf) \]

to get

\[ \delta J_a = \left[ w_x(tf) + g_x(tf) \right] \delta x_f \\
+ \left[ w_{tf} + g(tf) - g_x(tf) \dot{x}(tf) \right] \delta t_f + m^T(tf) \delta \nu \\
+ \int_{t_0}^{tf} \left[ g_x - \frac{d}{dt} g_x \right] \delta x dt \]

– 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

\[
\begin{align*}
g_x - \frac{d}{dt} g_x &= 0 \quad \text{(dim } n) \\
w_x=tf + g_x=tf &= 0 \quad \text{(dim } n) \\
w_{tf} + g(tf) - g_x(tf) \dot{x}(tf) &= 0 \quad \text{(dim } 1) 
\end{align*}
\]

– With \( x(t_0) = x_0 \) \((\text{dim } n)\) and \( m(x(tf), tf) = 0 \) \((\text{dim } m)\) combined with last 2 conditions ⇒ \( 2n + m + 1 \) constraints

– Solution of Eulers equation has \( 2n \) constants of integration for \( x(t) \), and must find \( \nu \) \((\text{dim } m)\) and \( t_f \) ⇒ \( 2n + m + 1 \) unknowns

• Some special cases:

– If \( t_f \) is fixed, \( h = h(x(tf)) \), \( m \rightarrow m(x(tf)) \) and we lose the last condition in box – others remain unchanged

– If \( t_f \) is fixed, \( x(tf) \) free, then there is no \( m \), no \( \nu \) and \( w \) reduces to \( h \).

• Kirk’s book also considers several other type of constraints.