Dynamic Systems Optimization

5.1: Disrete-Time Optimization: Single-Stage Systems

- To begin our invesitigation of optimization of dynamic systems, we'll focus on the most elementary dynamic problem ⇒ *single-stage discrete-time systems*:

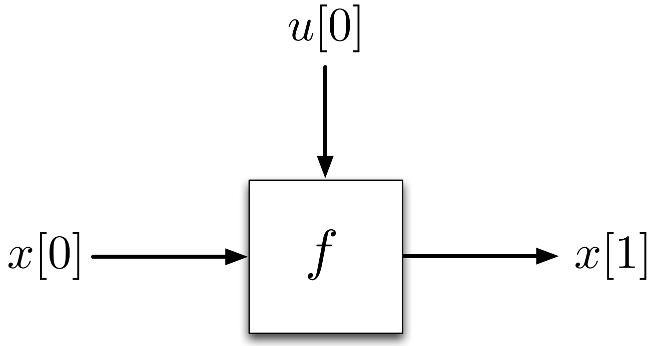


Figure 5.1 Single-stage system

$$x[1] = f\,(x[0],\,u[0])$$

where

$$J = \varphi\,\{x[1]\} + L\,\{x[0],\,u[0]\}$$

- Function $\varphi\,\{x[1]\}$ is a cost we may wish to place on the value of $x\,[1]$ and $L\,\{x[0],\,u[0]\}$ defines our performance cost.
- The basic idea can be stated as follows: *we wish to optimize some aspect of system performance across this single stage*.
- Dynamic process $f\,(x[0],\,u[0])$ establishes a *constraint* on the value of $x\,[1]$ for a given value of $u\,[0]$.
- So, what are the *unknowns* in this problem?

$x[1]$, $u[0]$ ($x[0]$ is a known initial condition)

- Exactly what *control* do we have?

  $u[0]$ ($x[1]$ will be fixed by the dynamic constraint once $u[0]$ is known)

> **Result:** This dynamic optimization problem is no more than a *parameter optimization problem* with equality constraints!

- So, here we'll use the *Lagrange multiplier* technique that we developed previously in order to attack this problem.
- We start by defining the augmented cost function $\bar{J}$, where here we'll adopt the notation J instead of L to denote these are dynamic optimizations.

$$\bar{J} = \varphi\left\{x[1]\right\} + J\left\{x[0],\, u[0]\right\} + \lambda[1]\left\{f\left(x[0],\, u[0]\right) - x[1]\right\}$$

- Next, define: $H = J + \boldsymbol{\lambda}^T \boldsymbol{f}$

$$\bar{J} = \varphi\left\{x[1]\right\} + H\left\{x[0],\, u[0],\, \lambda[1]\right\} - \lambda[1]x[1]$$

- Now, develop the first *variation* of $\bar{J}$:

$$\delta\bar{J} = \left\{\frac{d\varphi}{dx[1]} - \lambda[1]\right\} dx[1] + \frac{\partial H}{\partial u[0]} du[0]$$

$$+\frac{\partial H}{\partial x[0]} dx[0]$$

  – This expression tells us how small variations in $x[0]$, $x[1]$ and $u[0]$ affect the augmented cost $\bar{J}$.

- To find the minimum value of $\bar{J}$ (and hence J), we set $\delta\bar{J} = 0$ (stationarity condition).

- Now, by cleverly choosing $\lambda[1] = \dfrac{d\varphi}{dx[1]}$, we avoid determining $dx[1]$ in terms of $du[0]$.
- And, since $x[0]$ is given $\Rightarrow$ $dx[0] = 0$.
- Therefore, a stationary point of $\bar{J}$ will be obtained if,

$$\frac{\partial H}{\partial u[0]} = 0$$

- So the conditions for a stationary point in this single-stage problem can be stated as:

1. $\dfrac{d\varphi}{\partial x[1]}\lambda[1] = 0$
2. $\dfrac{\partial H}{\partial u[0]} = 0$
3. $x[1] = f\,\{x[0],\, u[0]\}$

   – These conditions provide $2n + m$ equations in $2n + m$ unknowns $\Rightarrow$ enough information to solve the problem!

5.2: Introduction to Multi-Stage Systems

- An obvious extension of the results above is to consider a system which changes dynamically over a series of stages:

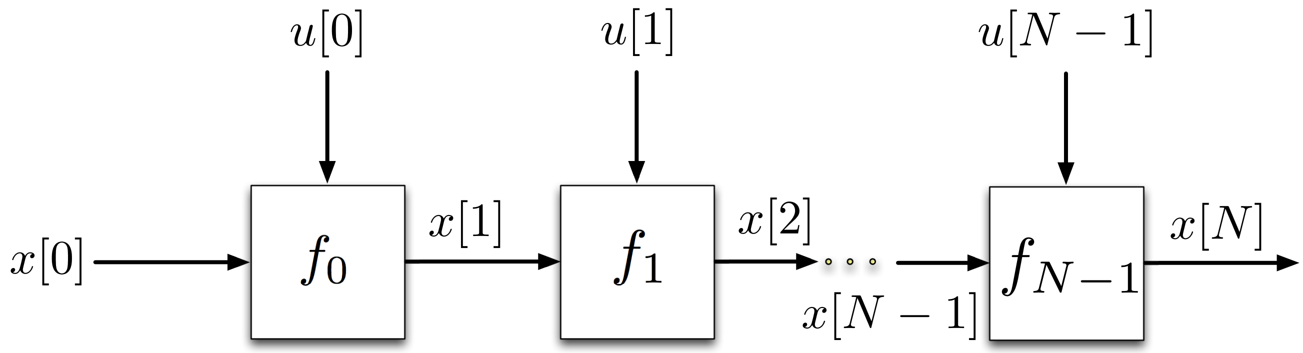


Figure 5.2 Multi-stage system

$$J = \varphi\left\{\boldsymbol{x}[N]\right\} + \sum_{k=0}^{N-1} L^{(k)}\left\{\boldsymbol{x}[k],\, \boldsymbol{u}[k]\right\}$$

where $\phi\left\{\boldsymbol{x}\,[N]\right\}$ is defined as the *terminal cost* and here $L^{(k)}$ is defined as the *running cost*.

  - In addition, for the multi-stage problem, our variable quantities are now vector-dimensioned for generality.
  - NOTE: Both $\boldsymbol{f}$ and $L^{(k)}$ could change at each stage; for simplicity however, we will assume here that this does not happen.

- OUR GOAL: Select the parameters $\{\boldsymbol{u}[k];\ k = 0,\, 1, \ldots,\, N-1\}$ and identify the corresponding parameters $\{\boldsymbol{x}[k];\ k = 1,\, 2, \ldots,\, N\}$ that minimize the cost.

- How should we attack this problem? Once again $\rightarrow$ LAGRANGE MULTIPLIERS!

  - Remember that we now have a set of N vector constraints (the dynamic equations) that must be adjoined to the cost:

$$J = J + \sum_{k=0}^{N-1} \boldsymbol{\lambda}^T[k+1]\{f(\boldsymbol{x}[k], \boldsymbol{u}[k]) - \boldsymbol{x}[k+1]\}$$

– And note, as before, that when constraints are met, $\bar{J} = J$.

- We'll now extend the definition of H:

$$H[k] = H\{\boldsymbol{x}[k], \boldsymbol{u}[k], \boldsymbol{\lambda}[k]\} = L\{\boldsymbol{x}[k], \boldsymbol{u}[k]\}$$

$$+\boldsymbol{\lambda}^T[k+1] f\{\boldsymbol{x}[k], \boldsymbol{u}[k]\}$$

$$\Rightarrow \bar{J} = \varphi\{\boldsymbol{x}[N]\} + H[0] + \sum_{k=1}^{N-1}\left\{H[k] - \boldsymbol{\lambda}^T[k]\boldsymbol{x}[k]\right\} - \boldsymbol{\lambda}^T[N]\boldsymbol{\lambda}[N]$$

- And now, since we've introduced Lagrange multipliers, we can take the first variation of $\bar{J}$ treating $\boldsymbol{u}[k]$ and $\boldsymbol{x}[k]$ as if they were independent:

$$\delta\bar{J} = \frac{\partial\varphi}{\partial\boldsymbol{x}[N]}d\boldsymbol{x}[N] + \sum_{k=1}^{N-1}\left\{\frac{\partial H[k]}{\partial\boldsymbol{x}[k]} - \boldsymbol{\lambda}^T[k]\right\}d\boldsymbol{x}[k]$$

$$+\sum_{k=1}^{N-1}\frac{\partial H[k]}{\partial\boldsymbol{u}[k]}d\boldsymbol{u}[k] + \frac{\partial H[0]}{\partial\boldsymbol{u}[0]}d\boldsymbol{u}[0]$$

$$+\frac{\partial H[0]}{\partial\boldsymbol{x}[0]}d\boldsymbol{x}[0] - \boldsymbol{\lambda}^T[N]d\boldsymbol{x}[N]$$

- Let's now examine this expression and determine the conditions for a stationary point ...

- We'll first simplify the expression,

– Remember, $\boldsymbol{\lambda}^T[k]$ is arbitrary, so we can choose it as we wish; a good choice is:

$$\boldsymbol{\lambda}^T[N] = \frac{\partial \varphi}{\partial \boldsymbol{x}[N]} \qquad \boldsymbol{\lambda}^T[k] = \frac{\partial H[k]}{\partial \boldsymbol{x}[k]},\ k = 1,\ 2, \ldots,\ N-1$$

which allows us to write,

$$\delta \bar{J} = \sum_{k=0}^{N-1} \frac{\partial H[k]}{\partial \boldsymbol{u}[k]} d\,\boldsymbol{u}[k] + \frac{\partial H[0]}{\partial \boldsymbol{x}[0]} d\,\boldsymbol{x}[0]$$

– If initial conditions are given, $d\,\boldsymbol{x}[0] = 0$ and $\delta \bar{J}$ simplifies to

$$\delta \bar{J} = \sum_{k=0}^{N-1} \frac{\partial H[k]}{\partial \boldsymbol{u}[k]} d\,\boldsymbol{u}[k]$$

– Thus, a necessary condition for a stationary point is:

$$\frac{\partial H[k]}{\partial \boldsymbol{u}[k]} = 0,\ k = 0,\ 1, \ldots,\ N-1$$

- Now that we've developed this stationarity condition, how do we solve the problem?

– Let's summarize: What are the unknowns involved?

$$\begin{aligned} &\boldsymbol{u}[k];\ k = 0,\ 1, \ldots,\ N-1 \quad (m * N) \\ &\boldsymbol{x}[k];\ k = 1,\ 2, \ldots,\ N \quad (n * N) \\ &\boldsymbol{\lambda}[k];\ k = 0,\ ,1, \ldots,\ N \quad (n * N) + n \end{aligned}$$

– And what are the equations available to solve for these unknowns?

$$\begin{aligned} (m * N) \qquad 0 \quad &= \frac{\partial H[k]}{\partial \boldsymbol{u}[k]} = \frac{\partial L}{\partial \boldsymbol{u}[k]} + \boldsymbol{\lambda}^T[k+1] \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}[k]},\ k = 0, \ldots, N-1 \\ (n * N) \quad \boldsymbol{\lambda}^T[k] \quad &= \frac{\partial H[k]}{\partial \boldsymbol{x}[k]} \quad k = 0, \ldots,\ N-1 \end{aligned}$$

$$n \quad \boldsymbol{\lambda}^T[N] = \frac{\partial \varphi}{\partial \boldsymbol{x}[N]}$$

$$(n * N) \;\; \boldsymbol{x}[k+1] = \boldsymbol{f}\,\{\boldsymbol{x}[k], \boldsymbol{u}[k]\} = \left\{ \frac{\partial H[k]}{\partial \boldsymbol{\lambda}^T[k+1]} \right\}, \; k = 0, \ldots, N-1$$

– Therefore, we have $N(2n+m)+n$ equations in $N(2n+m)+n$ unknowns $\Rightarrow$ solution exists if the equations are independent!

5.3: Example: Discrete-time Brachistochrone Problem

DESCRIPTION: A bead slides on a frictionless wire in a constant gravity field. The inclination angle, θ, may be changed at constant time intervals $\triangle t$.

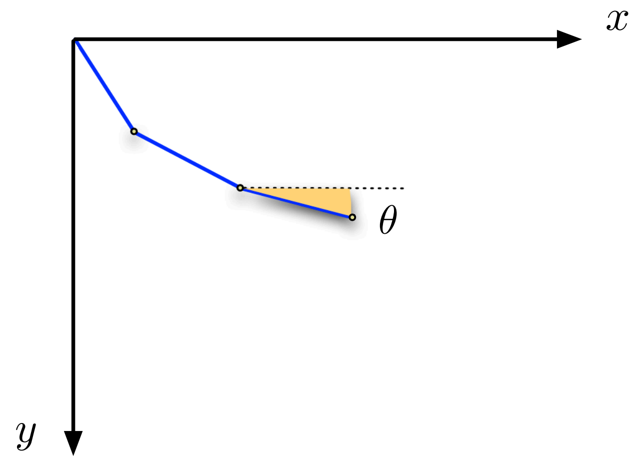


Figure 5.3 Brachistochrone Problem

FIND: $\theta[k]$ for $k = 0,\ 1, \ldots,\ N-1$ that maximizes the horizontal distance x at final time t_f with $\triangle t = {}^{t_f}/_{N}$.

SOLUTION:

1. The velocity and position along the wire at each corner point can be identified from:

$$v[k+1] = v[k] + g\sin\theta[k]\triangle t$$
$$\ell[k+1] = \ell[k] + v[k]\triangle t + \frac{1}{2}g\sin\theta[k]\triangle t^2$$

where $g = 9.81\ {}^{\mathrm{m}}/_{\mathrm{sec}^2}$ is gravitational acceleration.

(a) N can be introduced into the problem by normalizing the variables:

$$\tilde{v}[k] = \frac{v[k]}{gt_f} \qquad \tilde{\ell}[k] = \frac{\ell[k]}{gt_f^2}$$

$$\Rightarrow \tilde{v}[k+1] = \tilde{v}[k] + \frac{1}{N}\sin\theta[k]$$

$$\triangle\tilde{\ell}[k] = \tilde{\ell}[k+1] - \tilde{\ell}[k] = \frac{1}{N}\tilde{v}[k] + \frac{1}{2N^2}\sin\theta[k]$$

2. Using $\triangle\tilde{\ell}$, the x and y coordinates of the *corner points* can be identified:

$$\tilde{x}[k+1] = \tilde{x}[k] + \triangle\tilde{\ell}[k]\cos\theta[k]$$
$$\tilde{y}[k+1] = \tilde{y}[k] + \triangle\tilde{\ell}[k]\sin\theta[k]$$

3. For this problem, we define the *state vector* and *input* as:

$$\boldsymbol{x}[k] = \begin{bmatrix} \tilde{x}[k] \\ \tilde{y}[k] \\ \tilde{v}[k] \end{bmatrix}, \quad u[k] = \theta[k]$$

   (a) Thus we can express the *state equation* in the form:

$$\boldsymbol{x}[k+1] = f\left(\boldsymbol{x}[k], u[k]\right) = \begin{bmatrix} \tilde{x}[k] + \Delta\tilde{\ell}[k]\cos\left(\theta[k]\right) \\ \tilde{y}[k] + \Delta\tilde{\ell}[k]\sin\left(\theta[k]\right) \\ \tilde{v}[k] + \dfrac{1}{n}\sin\left(\theta[k]\right) \end{bmatrix}$$

4. What about the cost?

$$J = -\tilde{x}[N]$$

   (a) Note the minus sign is used because we wish to *maximize* the final horizontal position, $\tilde{x}$.

5. Build the augmented cost function using our previous definitions, $\phi = -\tilde{x}[N]$ and $L = 0$,

$$\bar{J} = -x[N] + \sum_{k=0}^{N-1}\left\{\lambda_v[k+1]\left\{\tilde{v}[k] + \frac{1}{N}\sin\theta[k]\right\}\right.$$

$$+\lambda_x[k+1]\left\{\tilde{x}[k] + \triangle\tilde{\ell}[k]\cos\theta[k]\right\}$$

$$+\lambda_y[k+1]\left\{\tilde{y}[k] + \triangle\tilde{\ell}[k]\sin\theta[k]\right\}\Big\}$$

$$-\sum_{k=1}^{N}\left\{\lambda_v[k]\tilde{v}[k] + \lambda_x[k]\tilde{x}[k] + \lambda_y[k]\tilde{y}[k]\right\}$$

(a) Why three λ's? Because we have three "states"; note also change of index in second summation.

6. So, the equations required to solve this problem are:

(a) Dynamic constraints ($3 * N$):

$$f(\boldsymbol{x}[k],\, u[k]) - \boldsymbol{x}\,[k+1] = 0$$

with intial conditions,

$$\tilde{x}[0]; = 0 \quad \tilde{y}[0] = 0 \quad \tilde{v}[0] = 0.$$

(b) Geometric constraint (N):

$$\triangle\tilde{\ell}[k] = \frac{1}{N}\tilde{v}[k] + \frac{1}{2N^2}\sin\theta[k]$$

(c) Costate equations:

$$\boldsymbol{\lambda}^T[k] = \frac{\partial H[k]}{\partial \boldsymbol{x}[k]}$$

where

$$\boldsymbol{\lambda}[k] = \begin{bmatrix} \lambda_x[k] \\ \lambda_y[k] \\ \lambda_v[k] \end{bmatrix}$$

and

$$\lambda_x[k] = \lambda_x[k+1]$$

$$\lambda_y[k] = \lambda_y[k+1]$$

$$\lambda_v[[k] = \lambda_v[k+1] + \lambda_x[k+1]\left(\frac{1}{N}\right)\cos\theta[k] + \lambda_y[k+1]\left(\frac{1}{N}\right)\sin\theta[k]$$

with terminal conditions,

$$\lambda_x[N] = -1 \quad \lambda_y[N] = 0 \quad \lambda_v[k] = 0.$$

(d) We can now express the optimality condition:

$$\frac{\partial H[k]}{\partial \boldsymbol{u}[k]} = 0$$

i. Noting that

$$H\,[k] = \lambda_x[k+1]\left(\tilde{x}[k] + \Delta\tilde{\ell}[k]\cos\left(\theta[k]\right)\right)$$

$$+\,\lambda_y[k]\left(\tilde{y}[k] + \Delta\tilde{\ell}[k]\sin\left(\theta[k]\right)\right)$$

$$+\,\lambda_v\,[k+1]\left(\tilde{v}_k + \frac{1}{N}\sin\left(\theta[k]\right)\right)$$

and from the terminal conditions for $\boldsymbol{\lambda}[N]$,

$$\lambda_x[k] = -1\; k = 1, \ldots, N$$
$$\lambda_y[k] = 0\; k = 1, \ldots, N$$

ii. We can write

$$\frac{\partial H[k]}{\partial \boldsymbol{u}[k]} = -\frac{\partial}{\partial \theta[k]}\left(\Delta\tilde{\ell}[k]\cos\left(\theta[k]\right)\right) + \frac{\partial}{\partial \theta[k]}\left(\tilde{v}_k + \frac{1}{N}\sin\left(\theta[k]\right)\right)$$

$$= \Delta\tilde{\ell}[k]\sin\left(\theta[k]\right) - \frac{1}{2N^2}\cos^2\left(\theta\,[k]\right) +$$

$$+\,\lambda_v[k+1]\left(\frac{1}{N}\right)\cos\left(\theta[k]\right) = 0$$

NOTE: $\theta[k]$ is a function of $\lambda_v[k+1]$ and $\tilde{v}[k]$.

- Even in this relatively simple example, the number of equations suggests the solution process could be extremely complicated.
- The existence of non-linear relationships makes the process even worse! So what can we do? DEVELOP NUMERICAL METHODS.
- What do we have available for use?
  1. A set of difference equations that develop *forward* in time; i.e., the *state equations*:
  $$\boldsymbol{x}[k+1] = \boldsymbol{f}\,\{\boldsymbol{x}[k], \boldsymbol{u}[k]\} \quad \boldsymbol{x}[0] = \boldsymbol{x}_0$$
  2. A set of difference equations that develop *backward* in time; i.e., the *co-state equations*:
  $$\boldsymbol{\lambda}[k] = \left\{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}[k]}\right\}^T \boldsymbol{\lambda}[k+1] + \left\{\frac{\partial L}{\partial \boldsymbol{x}[k]}\right\}^T$$
  $$\boldsymbol{\lambda}[N] = \left\{\frac{\partial \varphi}{\partial \boldsymbol{x}[N]}\right\}^T$$
- These two difference equations are coupled and define a TWO-POINT BOUNDARY VALUE PROBLEM.
  - The boundary conditions are split between the end points.
  - Once $\boldsymbol{\lambda}[k+1]$ and $\boldsymbol{x}[k]$ are known, $\boldsymbol{u}[k]$ can be computed using the algebraic equations defined by the optimality condition:
  $$\frac{\partial H[k]}{\partial \boldsymbol{u}[k]} = 0$$
- So, solving a two-point boundary value problem provides a means of identifying the solution of our dynamic optimization problem.
  - Curiosity: Even though we don't really care about $\boldsymbol{\lambda}[k]$, we must calculate it to identify $\boldsymbol{u}[k]$!

5.4: Solution methods for two-point boundary value problems

- What do we have available for use to solve the multi-stage problem?
  1. One set of difference equations that develop *forward* in time; i.e., the *state equations*:
  $$\boldsymbol{x}[k+1] = \boldsymbol{f}\,\{\boldsymbol{x}[k],\, \boldsymbol{u}[k]\} \qquad \boldsymbol{x}[0] = \boldsymbol{x}_0$$
  2. And one set of difference equations that develop *backward* in time; i.e., the *co-state equations*:
  $$\boldsymbol{\lambda}[k] = \left\{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}[k]}\right\}^T \boldsymbol{\lambda}[k+1] + \left\{\frac{\partial L}{\partial \boldsymbol{x}[k]}\right\}^T$$
  $$\boldsymbol{\lambda}[N] = \left\{\frac{\partial \varphi}{\partial \boldsymbol{x}[N]}\right\}^T$$
- These two difference equations are coupled and define a TWO-POINT BOUNDARY VALUE PROBLEM.
  - Note the boundary conditions are split between the end points – half at the start and half and the end.
  - Once $\boldsymbol{\lambda}[k+1]$ and $\boldsymbol{x}[k]$ are known, $\boldsymbol{u}[k]$ can be computed using the algebraic equations defined by the optimality condition:
  $$\frac{\partial H[k]}{\partial \boldsymbol{u}[k]} = 0$$
- So, solving a two-point boundary value problem provides a means of identifying the solution of our dynamic optimization problem.
  - Curiosity: Even though we don't really care about $\boldsymbol{\lambda}[k]$, we must calculate it to identify $\boldsymbol{u}[k]$!
- A number of methods exist to solve two-point boundary value problems

Shooting Method

1. Guess $\boldsymbol{\lambda}[0]$
2. Compute $\boldsymbol{\lambda}[1]$ using the co-state difference equations
3. Compute $\boldsymbol{u}[0]$ using $\boldsymbol{x}[0]$, $\boldsymbol{\lambda}[1]$, and the stationarity conditions
4. Compute $\boldsymbol{x}[1]$ using the state difference equations
5. Continue steps (2) through (4) up to time N
6. If $\boldsymbol{\lambda}^T[N] = \dfrac{\partial\varphi}{\partial\boldsymbol{x}[N]}$, the solution is correct; but if $\boldsymbol{\lambda}^T[N] \neq \dfrac{\partial\varphi}{\partial\boldsymbol{x}[N]}$, a new $\boldsymbol{\lambda}[0]$ must be chosen and steps (2) through (5) repeated
   (a) how do you choose $\boldsymbol{\lambda}[0]$? It's an art.
   (b) you might try varying each element of $\boldsymbol{\lambda}[0]$ individually to observe the sensitivity of the results to these changes, and then use this information to select the new $\boldsymbol{\lambda}[0]$

- Problem $\Rightarrow$ the process is very sensitive to the initial guess; the solution may not converge unless the first guess is pretty accurate

Gradient Method

1. Guess all of the control variables $\{\boldsymbol{u}[k];\ k = 0,\ 1,\ \ldots,\ N-1\}$
2. Compute $\boldsymbol{x}[k]$ using the state difference equations
3. Compute $\boldsymbol{\lambda}[k]$ backwards using the co-state difference equation
4. Stop when all $\partial H[k]/\partial\boldsymbol{u}[k]$ are sufficiently close to zero
   (a) Why? We want to set $\delta\bar{J} = \sum_{k=0}^{N-1} \partial H[k]/\partial\boldsymbol{u}[k]\, d\,\boldsymbol{u}[k] = 0$ which can only happen when $\partial H[k]/\partial\boldsymbol{u}[k] = 0$

(b) A useful criterion is the following RMS-type measurement:

$$\left[\left(\frac{1}{N}\right)\sum_{k=0}^{N-1}\left\{\frac{\partial H[k]}{\partial \boldsymbol{u}[k]}\right\}\left\{\frac{\partial H[k]}{\partial \boldsymbol{u}[k]}\right\}^T\right]^{\frac{1}{2}} < \epsilon$$

5. If the stopping criterion is not satisfied, another guess at the control variables must be made.

(a) This can be done during step (3) by setting

$$\boldsymbol{u}_{NEW}[k] = \boldsymbol{u}[k] - K\frac{\partial H[k]}{\partial \boldsymbol{u}[k]} \quad \text{for some } K > 0$$

(b) Why does this work?

i. if $\partial H/\partial \boldsymbol{u} > 0$, then $d\,\boldsymbol{u} < 0$ will produce $\delta \bar{J} < 0$ and hence $\bar{J}$ will decrease.

ii. if $\partial H/\partial \boldsymbol{u} < 0$, then $d\,\boldsymbol{u} > 0$ will produce $\delta J \bar{<} 0$ and hence $\bar{J}$ will decrease.

- Problem $\Rightarrow$ just like the parameter optimization problem, the selection of K here is tricky.

Solution Example: Gradient Method

- We can apply the gradient method outlined above to the brachistochrone problem developed previously.
- For this example, we chose the following parameters:

$$t_f = 10 \text{ sec}$$
$$N = 50$$

- The optimization was carried out using a gain $K = 2.0$ and a tolerance $\epsilon = 10^{-8}$.

- Doing so, we arrive arrive at the following result which gives the cyloid shape characteristic of the brachistichrone problem.

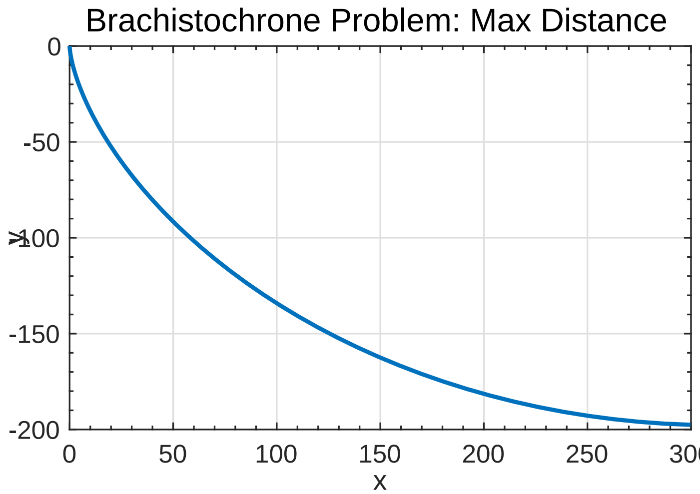


Figure 5.4 Brachistochrone Example

5.5: Continuous-Time: Fixed-Time, No Terminal Constraints

- Dynamic optimization problems for continuous-time systems are problems of the *Calculus of Variations.*
- Such problems can be often be considered as limiting cases of discrete-time systems where the time interval becomes infinitesimally small.
- Consider the system described by the non-linear vector differential equation,

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), t); \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0; \quad t_0 \le t \le t_f$$

  and cost function

$$J = \varphi\left\{\boldsymbol{x}(t_f), t_f\right\} + \int_{t_0}^{t_f} L\left\{\boldsymbol{x}(t), \boldsymbol{u}(t), t\right\} dt.$$

  Here, $\varphi\left\{\boldsymbol{x}(t_f), t_f\right\}$ represents the terminal cost and $L\left\{\boldsymbol{x}(t), \boldsymbol{u}(t), t\right\}$ is the running cost as with the discrete-time formulation.

- GOAL: Minimize J by selecting input $\boldsymbol{u}(t)$ and determine resulting state vector $\boldsymbol{x}(t)$.
- SOLUTION:
- First, define the augmented cost function $\bar{J}$ by adjoining the system dynamic (state) equation to J with Lagrange multipliers,

$$\bar{J} = \varphi\left\{\boldsymbol{x}(t_f), t_f\right\} + \int_{t_0}^{t_f} \left\{L(\boldsymbol{x}(t), \boldsymbol{u}(t), t) + \boldsymbol{\lambda}^T (f(\boldsymbol{x}(t), \boldsymbol{u}(t), t) - \dot{\boldsymbol{x}}t())\right\}$$

- Next, define the continuous-time *Hamiltonian* function,

$$H(\boldsymbol{x}(t), \boldsymbol{u}(t), t) = L(\boldsymbol{x}(t), \boldsymbol{u}(t)) + \boldsymbol{\lambda}(t)^T \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), t)$$

- In the subsequent development, we shall occasionally drop the functional dependence on t to simplify notation.

- Next, we make use of the following relationship:

$$\int_{t_0}^{t_f} -\boldsymbol{\lambda}^T(t)\,\dot{\boldsymbol{x}}(t)\,dt = \boldsymbol{\lambda}^T(t)\,\boldsymbol{x}(t)\Big|_{t_0}^{t_f} + \int_{t_0}^{t_f} \dot{\boldsymbol{\lambda}}^T(t)\,\boldsymbol{x}(t)\,dt$$

where we have used integration by parts $\left(\int u dv = uv - \int v du\right)$ where,

$$dv = \dot{\boldsymbol{x}}(t)\,dt \quad v = \boldsymbol{x}(t)$$

$$du = -\dot{\boldsymbol{\lambda}}^T(t)\,dt \quad u = -\boldsymbol{\lambda}^T(t)$$

- This gives $\Rightarrow$

$$\bar{J} = \varphi\left\{\boldsymbol{x}(t_f),\, t_f\right\} + \boldsymbol{\lambda}^T(t_0)\boldsymbol{x}(t_0) - \boldsymbol{\lambda}^T(t_f)\boldsymbol{x}(t_f) + \int_{t_0}^{t_f}\left(H + \dot{\boldsymbol{\lambda}}^T(t)\,\boldsymbol{x}(t)\right)dt$$

- Now, vary parameters,

$$\delta\bar{J} = \left.\frac{\partial\varphi}{\partial\boldsymbol{x}(t_f)}\right|\delta\boldsymbol{x}(t_f) + \boldsymbol{\lambda}^T(t_0)\delta\boldsymbol{x}(t_0) - \boldsymbol{\lambda}^T(t_f)\delta\boldsymbol{x}(t_f)$$

$$+ \int_{t_0}^{t_f}\left\{\left(\frac{\partial H}{\partial\boldsymbol{x}(t)} + \dot{\boldsymbol{\lambda}}^T(t)\right)\delta\boldsymbol{x}(t) + \frac{\partial H}{\partial\boldsymbol{u}(t)}\delta\boldsymbol{u}(t)\right\}dt$$

- Here, as before, we choose Lagrange multipliers such that the coefficients of $\delta\boldsymbol{x}$ vanish, giving:

$$\frac{\partial H}{\partial\boldsymbol{x}(t)} + \dot{\boldsymbol{\lambda}}^T(t) = 0$$

and by the boundary conditions at t_f we have

$$\frac{\partial\varphi}{\partial\boldsymbol{x}(t_f)} - \boldsymbol{\lambda}^T(t_f) = 0$$

- And finally, for an extremum, δJ must be zero for arbitrary variation $\delta\boldsymbol{u}(t)$, giving:

$$\frac{\partial H}{\partial \boldsymbol{u}(t)} = 0$$

- The three equations above are called the EULER-LAGRANGE EQUATIONS of the calculus of variations.
- In summary, to solve for the optimal control input $\boldsymbol{u}(t)$ that minimizes the performance function J, we need to solve the following set of differential equations:

$$\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t), \boldsymbol{u}(t), t)$$

$$\dot{\boldsymbol{\lambda}}(t) = -\left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}(t)}\right)^T \boldsymbol{\lambda} - \left(\frac{\partial L}{\partial \boldsymbol{x}(t)}\right)^T$$

where $\boldsymbol{u}(t)$ is determined by

$$\frac{\partial H}{\partial \boldsymbol{u}(t)} = 0 \quad \text{or} \quad \left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}(t)}\right)^T \boldsymbol{\lambda} + \left(\frac{\partial L}{\partial \boldsymbol{u}(t)}\right)^T = 0$$

- The boundary conditions are again split – some given for $t = t_0$ and some given for $t = t_f$:

$$\boldsymbol{x}(t_0) = \boldsymbol{x}_0$$

$$\boldsymbol{\lambda}(t_f) = \left(\frac{\partial \varphi}{\partial \boldsymbol{x}(t)}\right)^T$$

- So, again, as in the discrete-time multi-stage problem, we're faced with a *two-point boundary value problem.*

5.6: Example: Minimum Energy Room Temperature Control

- PROBLEM STATEMENT: We desire to heat a room using the least possible energy. Let T be the temperature in the room, T_a the ambient temperature outside (assumed constant) and $h(t)$ the rate of heat supplied to the room.

- Simplified dynamic equations may be written,

$$\dot{T}(t) = -a\left(T(t) - T_a\right) + bh(t)$$

  for some constants a and b which depend on room design and construction.

- We define the state as

$$x(t) \equiv T(t) - T_a$$

  and the control input as

$$u(t) \equiv h(t)$$

  - Thus, we may express the scalar state equation as:

$$\dot{x}(t) = -ax(t) + bu(t)$$

- In order to control the room temperature over the time interval $\left[t_0,\ t_f\right]$ using least energy, we define the cost function

$$J = \frac{1}{2}k\left(x(t_f) - x_d\right)^2 + \frac{1}{2}\int_{t_0}^{t_f} u(t)^2 dt$$

  for some weighting factor, k.

- SOLUTION:

  - Define the augmented cost function, $\bar{J}$

$$\bar{J} = J + \int_{t_0}^{t_f} \lambda(t)\left(f\left(x(t),\, u(t)\right) - \dot{x}(t)\right) dt$$

– Satisfy Euler-Lagrange equations:

$$\frac{\partial \varphi}{\partial t_f} - \lambda(t_f) = 0$$

$$\frac{\partial H}{\partial x(t)} + \dot{\lambda}(t) = 0$$

$$\frac{\partial H}{\partial u(t)} = 0$$

and

$$\dot{x}(t) = -ax(t) + bu(t)$$

– Recall,

$$\begin{aligned} H &= L(x(t), u(t)) + \lambda(t) f(x(t), u(t)) \\ &= \frac{1}{2}u(t)^2 + \lambda(t)[-ax(t) + bu(t)] \end{aligned}$$

so we get →

$$\begin{aligned} \dot{\lambda}(t) &= a\lambda(t) \qquad & \lambda(t_f) &= k\left[x(t_f) - x_d\right] \\ \dot{x}(t) &= -ax(t) + bu(t) \qquad & x(t_0) &= 0 \end{aligned}$$

$$u(t) + \lambda(t)b = 0 \qquad \Rightarrow \qquad u(t) = -b\lambda(t)$$

– Substituting for $u(t)$,

$$\begin{aligned} \dot{x} &= -ax - b^2\lambda \\ \dot{\lambda} &= a\lambda \end{aligned}$$

or, writing in matrix form,

$$\begin{bmatrix} \dot{x}(t) \\ \dot{\lambda}(t) \end{bmatrix} = \begin{bmatrix} -a & -b^2 \\ 0 & a \end{bmatrix} \begin{bmatrix} x(t) \\ \lambda(t) \end{bmatrix}$$

$$\dot{\boldsymbol{z}}(t) = A\boldsymbol{z}(t)$$

- Solving, we obtain

$$z(t) = e^{At}z(0)$$

$\Rightarrow$

$$\begin{aligned} x(t) &= e^{-at}x(0) - \frac{b^2}{a} \cdot \frac{1}{2}\left(e^{at} - e^{-at}\right)\lambda(0) \\ &= e^{-at}x(0) - \frac{b^2}{a}\sinh(at)\lambda(0) \\ \lambda(t) &= e^{at}\lambda(0) \end{aligned}$$

- But... we don't know $\lambda(0)$
  - However,

$$\begin{aligned} \lambda(0) &= \lambda(t_f)e^{-at_f} \\ &= k\left[x(t_f) - x_d\right]e^{-at_f} \end{aligned}$$

  - But where do we get e^{-at_f}?

$$\begin{aligned} x(t_f) &= -\frac{b^2}{2a}\left(e^{at_f} - e^{-at_f}\right)e^{-at_f}\lambda(t_f) \\ &= -\frac{b^2}{2a}\left(1 - e^{-2at_f}\right)\lambda(t_f) \end{aligned}$$

  and

$$\lambda(t_f) = k\left[x(t_f) - x_d\right]$$

- Solving,

$$x(t_f) = \frac{x_d}{1 + \frac{ae^{at_f}}{b^2k\sinh(at_f)}}$$

$$\lambda(t_f) = \frac{-2x_d ak}{2a - b^2k\left(1 - e^{-2at_f}\right)}$$

$\Rightarrow$

$$u(t) = \frac{x_d abk e^{at}}{ae^{at_f} + b^2 k \sinh(at_f)}$$

$$x(t) = \frac{x_d b^2 k \sinh(at)}{ae^{at_f} + b^2 k \sinh(at_f)}$$

- Putting some numbers to the example, let's let

$$a = 0.4$$
$$b = 0.8$$

and assume $t_0 = 0$ and $t_f = 20$.

- Further assume $T_a = 5$, $T_0 = T_a$, and define a target temperature of $T_d = 20\,(x_d = 15)$.

- The following plots show resulting temperature and control effort plots for weightings $k = 1,\ 2,\ 4,\ 16$.

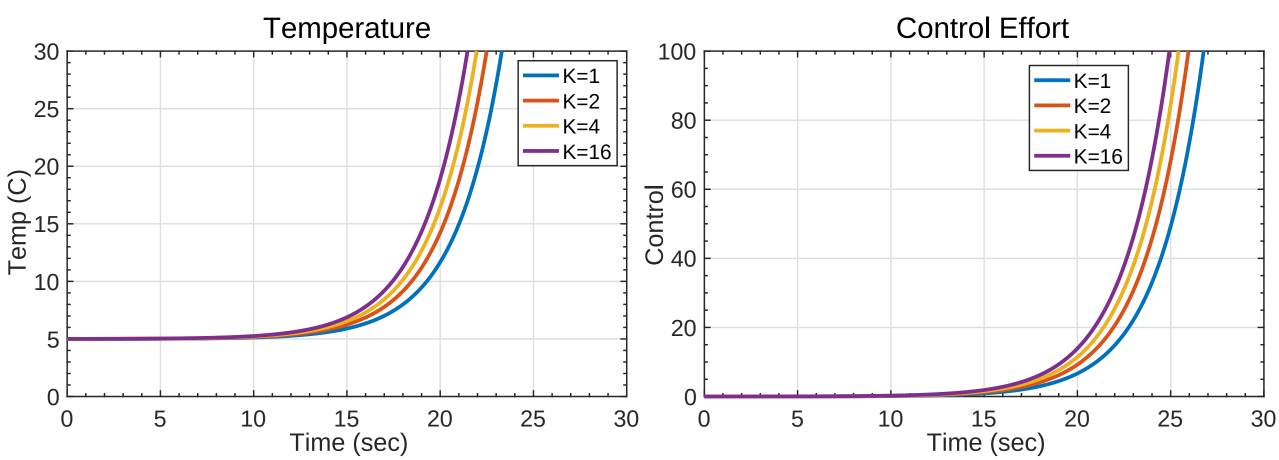


Figure 5.5 Room Heating Example

5.7: Continuous-Time: Fixed-Time, Terminal Constraints

- We'll now look at a Calculus of Variations approach to solving a slightly more difficult optimization problem: one with constraints imposed at the *terminal time*.

PROBLEM:

Dynamic System	$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), t)$	$\boldsymbol{x}(t_0) = \boldsymbol{x}_0$
Cost	$J = \varphi\left\{\boldsymbol{x}\left(t_f\right)\right\} + \int_{t_0}^{t_f} L\{\boldsymbol{x}(t), \boldsymbol{u}(t), t\}\, dt$	
Terminal Constraints	$\boldsymbol{\psi}\left\{\boldsymbol{x}\left(t_f\right)\right\} = \boldsymbol{c}$	$\boldsymbol{\psi}$ is $q \times 1$, $q \le n$
Goal	Select $\boldsymbol{u}(t)$ to minimize J subject to terminal constraints.	

SOLUTION:

- As before, we'll adjoin the constraints to the cost function using Lagrange multipliers.
- The difference is that we now have *two types* of constraints:
  - Intermediate dynamic constraints $\Rightarrow$ $\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), t)$
  - Terminal constraints $\Rightarrow$ $\boldsymbol{\psi}\left\{\boldsymbol{x}(t_f)\right\} = \boldsymbol{c}$
- Building the augmented cost function,

$$\bar{J} = \varphi\left\{\boldsymbol{x}(t_f)\right\} + \boldsymbol{\nu}^T\left\{\boldsymbol{\psi}\left[\boldsymbol{x}(t_f)\right] - \boldsymbol{c}\right\}$$

$$+ \int_{t_0}^{t_f} \left\{L\left[\boldsymbol{x}(t), \boldsymbol{u}(t)\right], t + \boldsymbol{\lambda}^T(t)\left[\boldsymbol{f} - \dot{\boldsymbol{x}}(t)\right]\right\} dt$$

- Note:
  - $\varphi\left\{\boldsymbol{x}\left(t_f\right)\right\}$ is an aggregate function of the *final states* that we want to make “small” in some sense;
  - $\boldsymbol{\psi}\left\{\boldsymbol{x}\left(t_f\right)\right\}$ is a vector function of the final states that we want “fixed” at constraint $\boldsymbol{c}$.
- Furthermore, if we define $\Phi\left\{\boldsymbol{x}(t_f)\right\} = \varphi + \boldsymbol{\nu}^T\left\{\boldsymbol{\psi} - \boldsymbol{c}\right\}$, the problem looks identical to the previous one and can be solved in the same way.
- There are some mathematical distinctions, however:
  - System must be *controllable* so that it is possible to reach the specified terminal constraints;
  - Variation $\delta\boldsymbol{u}$ is no longer arbitrary since the only admissible values are ones which ensure terminal constraints are satisfied.
- Let’s now step through the solution process:

$$\bar{J} = \Phi\left\{\boldsymbol{x}(t_f)\right\} + \int_{t_0}^{t_f}\left\{H(\boldsymbol{x}(t), \boldsymbol{u}(t), t) - \boldsymbol{\lambda}^T(t)\,\dot{\boldsymbol{x}}(t)\right\}dt$$

- Integrating by parts yields:

$$\bar{J} = \Phi\left\{\boldsymbol{x}(t_f)\right\} + \boldsymbol{\lambda}^T(t_0)\,\boldsymbol{x}_0 - \boldsymbol{\lambda}^T\left(t_f\right)\boldsymbol{x}\left(t_f\right) + \int_{t_0}^{t_f}\left\{H(\boldsymbol{x}(t), \boldsymbol{u}(t), t) + \dot{\boldsymbol{\lambda}}^T(t)\,\boldsymbol{x}\right.$$

- Taking the first variation of $\bar{J}$ yields:

$$\delta\bar{J} = \left\{\frac{\partial\Phi}{\partial\boldsymbol{x}(t_f)} - \boldsymbol{\lambda}^T\left(t_f\right)\right\}\delta\boldsymbol{x}(t_f) + \boldsymbol{\lambda}^T(t_0)\,\delta\boldsymbol{x}_0 + \frac{\partial\Phi}{\partial\boldsymbol{c}}\delta\boldsymbol{c}$$

$$+ \int_{t_0}^{t_f}\left\{\left[\frac{\partial H}{\partial\boldsymbol{x}(t)} + \dot{\boldsymbol{\lambda}}^T(t)\right]\delta\boldsymbol{x}(t) + \frac{\partial H}{\partial\boldsymbol{u}(t)}\delta\boldsymbol{u}(t)\right\}dt$$

- Eliminate the $\delta \boldsymbol{x}$ terms by selecting appropriate Lagrange multipliers,

$$\dot{\boldsymbol{\lambda}}(t) = -\frac{\partial H}{\partial \boldsymbol{x}(t)}^T = -\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}(t)}^T \boldsymbol{\lambda} - \frac{\partial L}{\partial \boldsymbol{x}(t)}^T$$

$$\boldsymbol{\lambda}^T(t_f) = \frac{\partial \Phi}{\partial \boldsymbol{x}(t_f)} = \frac{\partial \varphi}{\partial \boldsymbol{x}(t_f)} + \boldsymbol{\nu}^T \frac{\partial \boldsymbol{\psi}}{\partial \boldsymbol{x}(t_f)}$$

- Can we ignore $\delta \boldsymbol{x}_0$ and $\delta \boldsymbol{c}$? Normally *yes* because $\boldsymbol{x}_0$ and $\boldsymbol{c}$ are fixed

$$\Rightarrow \ \delta \boldsymbol{x}_0 = \delta \boldsymbol{c} = 0$$

- We've included these terms in $\delta \bar{J}$ to make a point about $\boldsymbol{\lambda}(t_0)$ and $\boldsymbol{\nu} = -{}^{\partial \Phi}/_{\partial \boldsymbol{c}}{}^T$:
  - $\boldsymbol{\lambda}_0$ and $\boldsymbol{\nu}$ represent the sensitivity of $\bar{J}$ to changes in initial conditions and terminal constraints respectively.
  - So, if we know $\boldsymbol{\lambda}_0$ and $\boldsymbol{\nu}$, we can estimate (to first order) changes in J that would be caused by changing $\boldsymbol{x}_0$ and $\boldsymbol{c}$.

- Using these results, we can simplify $\delta \bar{J}$ to $\Rightarrow$

$$\delta \bar{J} = \int_{t_0}^{t_f} \frac{\partial H}{\partial \boldsymbol{u}(t)} \delta \boldsymbol{u}(t)\, dt$$

  and so δJ can only be non-negative for admissible $\delta \boldsymbol{u}(t)$ if

$$\frac{\partial H}{\partial \boldsymbol{u}(t)}^T = \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}(t)}^T \boldsymbol{\lambda}(t) + \frac{\partial L}{\partial \boldsymbol{u}(t)}^T = 0 \quad \text{[stationarity conditions]}$$

- The stationarity conditions are a set of algebraic equations that will be used to define $\boldsymbol{u}(t)$.

- Finally, we must remember that all constraints must be satisfied:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), t), \quad \boldsymbol{x}(t_0) = \boldsymbol{x}_0$$

$$\boldsymbol{\psi}\left[\boldsymbol{x}(t_f)\right] = \boldsymbol{c}$$

- Summarizing:

Unknowns	Equations
$\boldsymbol{x}(t) \rightarrow n \times 1$	$\dot{\boldsymbol{x}}(t) = \boldsymbol{f} \rightarrow n$
$\boldsymbol{u}(t) \rightarrow m \times 1$	$\partial H / \partial \boldsymbol{u}(t) = 0 \rightarrow m$
$\boldsymbol{\lambda}(t) \rightarrow n \times 1$	$\dot{\boldsymbol{\lambda}}(t) = -\partial H / \partial \boldsymbol{x}(t)^T \rightarrow n$
$\boldsymbol{\nu} \rightarrow q \times 1$	$\boldsymbol{\psi} = \boldsymbol{c} \rightarrow q$

- This is still a difficult two-point boundary value problem – now with extra parameters $\boldsymbol{\nu}$ to be identified!

5.8: Example: Fixed-Time with Terminal Constraints

- This example is taken from Kirk, pg. 313, Problem 5-12.

GIVEN:

- State equations:

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = -x_2(t) + u(t)$$

$$x_1(0) = x_2(0) = 0$$

with cost function,

$$J = \frac{1}{2}\{x_1(2) - 5\}^2 + \frac{1}{2}\{x_2(2) - 2\}^2 + \frac{1}{2}\int_0^2 u(t)^2\,dt$$

and terminal constraint,

$$x_1(2) + 5x_2(2) = 15$$

FIND:

- Optimal control input, $u(t)$.

SOLUTION:

- Form the augmented cost function:

$$\bar{J} = \frac{1}{2}[x_1(2) - 5]^2 + \frac{1}{2}[x_2(2) - 2]^2 + \nu[x_1(2) + 5x_2(2) - 15]$$
$$+ \int_0^2 \left\{ \frac{1}{2}u(t)^2 + \lambda_1(t)\,x_2(t) + \lambda_2(t)\,[-x_2(t) + u(t)] \right.$$
$$\left. -\lambda_1(t)\,\dot{x}_1(t) - \lambda_2(t)\,\dot{x}_2(t) \right\} dt$$

– After integration by parts, we have

$$\bar{J} = \frac{1}{2}[x_1(2) - 5]^2 + \frac{1}{2}[x_2(2) - 2]^2 + \nu[x_1(2) + 5x_2(2) - 15]$$

$$+\lambda_1(0)x_1(0) + \lambda_2(0)x_2(0) - \lambda_1(2)x_1(2) - \lambda_2(2)x_2(2)$$

$$+\int_0^2 \left\{H - \dot{\lambda}_1(t)\, x_1(t) - \dot{\lambda}_2(t)\, x_2(t)\right\} dt$$

$$\left[\text{where } H = \frac{1}{2}u(t)^2 + \lambda_1(t)\, x_2(t) + \lambda_2(t)\left(-x_2(t) + u(t)\right)\right]$$

- The costate equations are found from:

$$\dot{\boldsymbol{\lambda}}(t) = \frac{-\partial H}{\partial \boldsymbol{x}(t)} \;\Rightarrow\; \begin{aligned} \dot{\lambda}_1(t) &= 0 \\ \dot{\lambda}_2(t) &= -\lambda_1(t) + \lambda_2(t) \end{aligned}$$

- Solving for $\lambda_1(t)$,

$$\lambda_1(t) = \lambda_1(0)$$

- Solving for $\lambda_2(t)$ we'll use Laplace transforms to write,

$$\begin{aligned} s\Lambda_2(s) - \lambda_2(0) &= \Lambda_2(s) - \frac{\lambda_1(0)}{s} \\ s\Lambda_2(s) - \Lambda_2(s) &= \frac{\lambda_2(0) - \lambda_1(0)}{s} \\ \Lambda_2(s) &= \frac{\lambda_2(0)\, s - \lambda_1(0)}{s(s-1)} \end{aligned}$$

and using partial fraction expansion, we obtain,

$$\lambda_2(t) = \lambda_1(0) + \left(\lambda_2(0) - \lambda_1(0)\right) e^t$$

- From the boundary conditions, we have

$$\frac{\partial \Phi}{\partial x_1(2)} = x_1(t) - 5 + \nu - \lambda_1(2) = 0$$

$$\frac{\partial \Phi}{\partial x_2(2)} = x_2(t) - 2 + 5\nu - \lambda_2(2) = 0$$

$$x_1(2) - 5 + \nu = k_1$$
$$x_2(2) - 2 + 5\nu = k_1\left(1 - e^2\right) + k_2 e^2$$

- We may express the stationarity condition as
$$\frac{\partial H}{\partial u(t)} = u(t) + \lambda_2(t) = 0 \;\Rightarrow\; u(t) = -\lambda_2(t)$$
- Solving the state equations we write,

1. $\dot{x}_2 = -x_2 - \lambda_2 \quad \Rightarrow$
$$(s+1)\,X_2(s) - x_2(0) = -\frac{[\lambda_2(0)\,s - \lambda_1(0)]}{s(s-1)}$$
$$(s+1)\,X_2(s) = -\frac{[\lambda_2(0)\,s - \lambda_1(0)]}{s(s-1)} + x_2(0)$$
$$X_2(s) = \frac{-\lambda_2(0)\,s + \lambda_1(s) + x_2(0)\,s(s-1)}{s(s-1)(s+1)}$$
   - Using the method of partial fraction expansion (and the fact that $x_2(0) = 0$) gives:
$$x_2(t) = -\lambda_1(0) - \frac{1}{2}(\lambda_1(0) - \lambda_2(0))\,e^t + \frac{1}{2}(\lambda_1(0) + \lambda_2(0))\,e^{-t}$$
   or equivalently,
$$x_2(t) = -\lambda_1(0) + \lambda_1(0)\cosh(t) - \lambda_2(0)\sinh(t)\,.$$

$\dot{x}_1 = x_2 \quad \Rightarrow$
$$x_1(t) = \int_0^t x_2(\tau)\,d\tau$$
$$x_1(t) = -\lambda_1(0)\,t + \lambda_2(0) + \lambda_1(0)\sinh(t) - \lambda_2(0)\cosh(t)$$

- To finish the problem, we must identify $\lambda_1(0)$, $\lambda_2(0)$, and ν

– Collect boundary conditions ⇒

$$
\begin{aligned}
x_1(t_f) - 5 + \nu &= \lambda_1(0) \\
x_2(t_f) - t_f + 5\nu &= \lambda_1(0)\left(1 - e^{t_f}\right) + \lambda_2(0)\,e^{t_f} \\
x_1(t_f) &= \lambda_1(0)\left(-t_f - {}^1\!/_2 e^{-t_f} + {}^1\!/_2 e^{t_f}\right) \\
&\quad + \lambda_2(0)\left(1 - {}^1\!/_2 e^{-t_f} - {}^1\!/_2 e^{t_f}\right) \\
x_2(t_f) &= \lambda_1(0)\left(-1 + {}^1\!/_2 e^{-t_f} + {}^1\!/_2 e^{t_f}\right) \\
&\quad + \lambda_2(0)\left({}^1\!/_2 e^{-t_f} - {}^1\!/_2 e^{t_f}\right) \\
x_1(t_f) + 5x_2(t_f) &= 15
\end{aligned}
$$

– Substituting values and expressing in matrix form

$$
\begin{bmatrix} 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & 5 & 6.3891 & -7.3891 \\ 1 & 0 & 0 & -1.6269 & 2.7622 \\ 0 & 1 & 0 & -2.7622 & 3.6269 \\ 1 & 5 & 0 & 0 & 0 \end{bmatrix}
\begin{bmatrix} x_1(2) \\ x_2(2) \\ \nu \\ \lambda_1(0) \\ \lambda_2(0) \end{bmatrix}
=
\begin{bmatrix} 5 \\ 2 \\ 0 \\ 0 \\ 15 \end{bmatrix}
$$

– Solving gives,

$$
\begin{bmatrix} x_1(2) \\ x_2(2) \\ \nu \\ \lambda_1(0) \\ \lambda_2(0) \end{bmatrix}
=
\begin{bmatrix} 3.0576 \\ 2.3885 \\ -0.6553 \\ -2.5976 \\ -2.6369 \end{bmatrix}
$$

– Since $u(t) = -\lambda_2(t) = -\lambda_1(0)\left(1 - e^t\right) - \lambda_2(0)\,e^t$,

$$
u^*(t) = 2.5976 + 0.03927e^t
$$

– The optimal cost is computed as:

$$
J^* = 1.8864 + 0.07546 + \frac{1}{2}\int_0^2 \left(2.5976 + 0.03927e^t\right)^2 dt = 9.3818
$$

- The accompanying figure shows the state trajectory for three different values of t_f – note how all arrive exactly at the constraint.

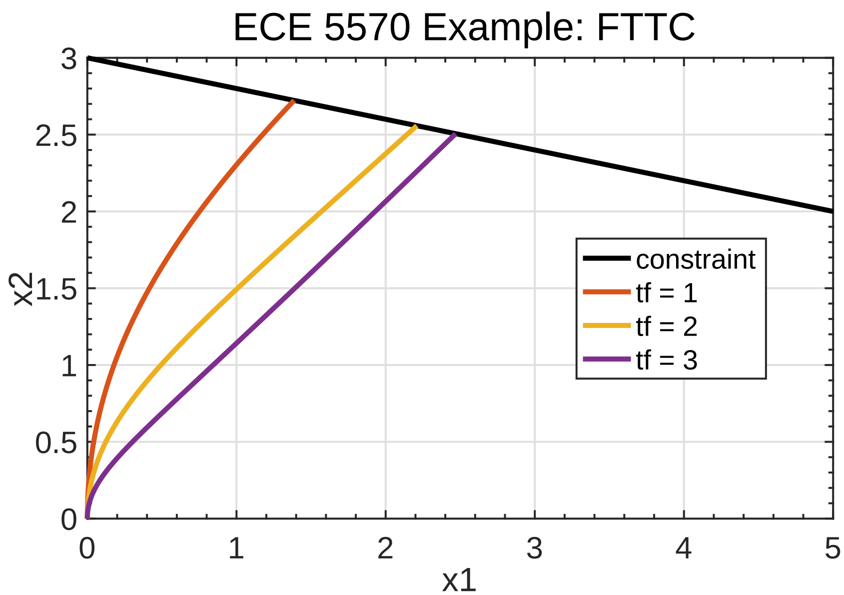


Figure 5.6 Example State Trajectory, $t_f = 1.0,\ 2.0,\ 3.0$

- Question: What would the cost be if I changed the terminal constraint to $x_1(2) + 5x_2(2) = 15.3$?
  - First order approximation:
    $\delta J = -\nu\delta c = -(-.6553)(.3) = 0.1966 \;\Rightarrow\; J = 9.5784$
  - Actual: $J = 9.5829$

- In many instances, the terminal constraints placed on the problem examined above are not functions of the final states, but rather constraints on the final states themselves:

$$\psi_i\left\{\boldsymbol{x}(t_f)\right\} = x_i(t_f) = c_i, \quad i = 1,\, 2,\, \ldots,\, q;\; q \le n$$

  - Does this change the general solution process? NO!

– What does change?

$$\boldsymbol{\lambda}(t_f) = \frac{\partial \Phi}{\partial \boldsymbol{x}(t_f)} \quad \Rightarrow \quad \lambda_i(t_f) = \begin{cases} \nu_i & i = 1, 2, \ldots, q \\ \dfrac{\partial \varphi}{\partial x_i(t_f)} & i = q+1, \ldots, n \end{cases}$$

∘ NOTE: φ will not be a function of $(x_i;\ i = 1, \ldots, q)$ because these states are already constrained

$$\boldsymbol{\psi}\left\{\boldsymbol{x}(t_f)\right\} = \boldsymbol{c} \quad \Rightarrow \quad x_i(t_f) = c_i$$

- Let's now examine the problem with fixed constraints placed on the terminal states:

$$J = \frac{1}{2}\int_0^2 u^2(t)\,dt$$
$$x_1(2) = 5$$
$$x_2(2) = 2$$

- H is the same $\Rightarrow$ $\lambda_1(t),\ \lambda_2(t),\ x_1(t),\ x_2(t)$ same form as before
- The two constraints give rise to 4 unknowns to identify: $(\nu_1,\ \nu_2,\ k_1,\ k_2)$
- ... And 4 equations,

$$\nu_1 = \lambda_1(0)$$
$$\nu_2 = \lambda_2(0)$$
$$x_1(t_f) = \left(-t_f + \sinh(t_f)\right)\lambda_1(0) + \left(1 - \cosh(t_f)\right)\lambda_2(0)$$
$$x_2(t_f) = \left(-1 + \cosh(t_f)\right)\lambda_1(0) - \sinh(t_f)\,\lambda_2(0)$$

- Substituting values,

$$\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 19.0855 & -20.0855 \\ 0 & 0 & 7.0179 & -9.0677 \\ 0 & 0 & 9.0677 & -10.0179 \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \lambda_1(0) \\ \lambda_2(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 5 \\ 2 \end{bmatrix}$$

- Solving

$$\begin{aligned} \nu_1 &= -2.6811 \\ \nu_2 &= -1.5832 \\ \lambda_1(0) &= -2.6811 \\ \lambda_2(0) &= -2.6264 \end{aligned}$$

with optimal control input

$$u^*(t) = 7.292 - 1.187e^t$$

and optimal cost

$$J^* = 16.75.$$

- As before, the accompanying figure shows the resulting state trajectory for three different values of t_f – here all arrive exactly at the target point.

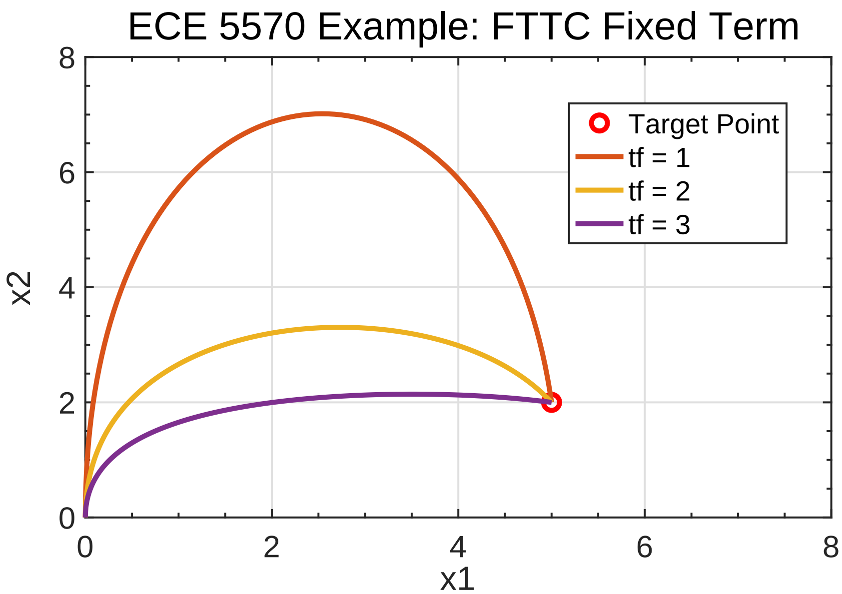


Figure 5.7 Example State Trajectory, $t_f = 1.0,\ 2.0,\ 3.0$

Changing Hard Terminal Constraints to Soft Constraints

- In many situations, the fixed-time terminal constraint problem may be too difficult to solve analytically so we must resort to numerical solutions.
  - Can the software developed for the no terminal constraint problem be used here without modification? *YES!*
  - How? Change the *hard* constraints to *soft* constraints with weighting and apply one of the algorithms presented previously, e.g.,

$$\varphi(t_f) \rightarrow \varphi(t_f) + \sum_i w_i \left\{\psi_i - c_i\right\}^2$$

  - What does this approach imply? That we'll be satisfied with small deviations from the terminal constraints.

5.9: Continuous-Time Optimization: Free Time Problems

- So far, we've examined the continuous-time optimization problem assuming the final time was *fixed*; but in many cases, it may be *free* and will thus be another parameter to be selected in the optimization.
  - Note however that any changes in $\boldsymbol{x}(t_f)$ are not independent of changes in t_f !

$$d\boldsymbol{x}(t_f) = \delta\boldsymbol{x}(t_f) + \dot{\boldsymbol{x}}(t_f)dt_f$$

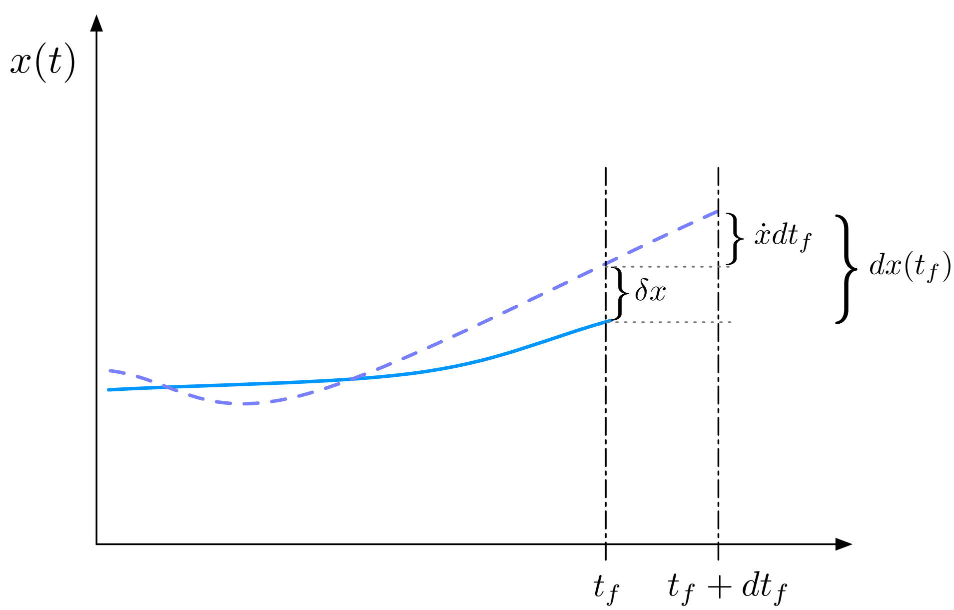


Figure 5.8 Free-Time Problem

- So in this problem, we have to worry about both types of change in $\boldsymbol{x}(t_f) \Rightarrow$ our problem will be slightly more complicated.
- How do we solve it? CALCULUS OF VARIATIONS

PROBLEM:

Dynamic System	$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}\left(\boldsymbol{x}(t), \boldsymbol{u}(t), t\right)$	$\boldsymbol{x}(t_0) = \boldsymbol{x}_0$
Cost	$J = \varphi\left\{\boldsymbol{x}\left(t_f\right)\right\} + \int_{t_0}^{t_f} L\left\{\boldsymbol{x}(t), \boldsymbol{u}(t), t\right\} dt$	
Terminal Constraints	$\boldsymbol{\psi}\left\{\boldsymbol{x}\left(t_f\right), t_f\right\} = \boldsymbol{c}$	$\boldsymbol{\psi}$ is $q \times 1,\ q \leq n$
Goal	Select $\boldsymbol{u}(t)$ and t_f to minimize J subject to terminal constraints.	

SOLUTION:

- Construct the augmented cost function,

$$\bar{J} = \varphi + \boldsymbol{\nu}^T\left\{\boldsymbol{\psi} - \boldsymbol{c}\right\} + \int_{t_0}^{t_f} \left\{L + \boldsymbol{\lambda}^T(t)\left(\boldsymbol{f} - \dot{\boldsymbol{x}}(t)\right)\right\} dt$$

$$= \Phi + \int_{t_0}^{t_f} \left\{L + \boldsymbol{\lambda}^T(t)\left(\boldsymbol{f} - \dot{\boldsymbol{x}}(t)\right)\right\} dt$$

- Now, we can take the differential:

$$\begin{aligned} d\bar{J} &= \frac{\partial \Phi}{\partial \boldsymbol{x}(t_f)} d\boldsymbol{x}(t_f) + \frac{\partial \Phi}{\partial t_f} dt_f \\ &+ \left\{H - \boldsymbol{\lambda}^T(t)\,\dot{\boldsymbol{x}}(t)\right\}\Big|_{t_f} dt_f - \left\{H - \boldsymbol{\lambda}^T(t)\,\dot{\boldsymbol{x}}(t)\right\}\Big|_{t_0} dt_0 \\ &+ \int_{t_0}^{t_f} \left\{\frac{\partial H}{\partial \boldsymbol{x}(t)} \delta \boldsymbol{x}(t) + \frac{\partial H}{\partial \boldsymbol{u}(t)} \delta \boldsymbol{u}(t) - \boldsymbol{\lambda}^T(t)\,\delta \dot{\boldsymbol{x}}(t)\right\} dt - \boldsymbol{\nu}^T \delta \boldsymbol{c} \end{aligned}$$

$\Rightarrow$

$$d\bar{J} = \frac{\partial \Phi}{\partial \boldsymbol{x}(t_f)} d\boldsymbol{x}(t_f) + \frac{\partial \Phi}{\partial t_f} dt_f$$

$$+L(t_f)dt_f - L(t_0)dt_0$$

$$+\int_{t_0}^{t_f} \left\{ \left[\frac{\partial H}{\partial \boldsymbol{x}\,(t)} + \dot{\boldsymbol{\lambda}}^T\,(t) \right] \delta\boldsymbol{x}\,(t) + \frac{\partial H}{\partial \boldsymbol{u}t\,()} \delta\boldsymbol{u}\,(t) \right\} dt$$

$$+\boldsymbol{\lambda}^T(t_0)\delta\boldsymbol{x}(t_0) - \boldsymbol{\lambda}^T(t_f)\delta\boldsymbol{x}(t_f)$$

$$-\boldsymbol{\nu}^T\delta\boldsymbol{c}$$

- What is new here?
- Since t_f is free, we must account for the fact that the final value of L may vary $\left(L(t_f)dt_f\right)$ as well as the fact that $\boldsymbol{x}(t_f)$ may vary in two ways $(d\boldsymbol{x} = \delta\boldsymbol{x} + \dot{\boldsymbol{x}}\,dt)$

$$d\bar{J} = \left\{ \frac{\partial \Phi}{\partial \boldsymbol{x}(t_f)} - \boldsymbol{\lambda}^T(t_f) \right\} \delta\boldsymbol{x}(t_f) + \left\{ \frac{\partial \Phi}{\partial t_f} + \frac{\partial \Phi}{\partial \boldsymbol{x}(t_f)} \dot{\boldsymbol{x}}(t_f) + L(t_f) \right\} dt_f$$

$$+\int_{t_0}^{t_f} \left\{ \left[\frac{\partial H}{\partial \boldsymbol{x}\,(t)} + \dot{\boldsymbol{\lambda}}^T\,(t) \right] \delta\boldsymbol{x}\,(t) + \frac{\partial H}{\partial \boldsymbol{u}\,(t)} \delta\boldsymbol{u}\,(t) \right\} dt$$

$$+\boldsymbol{\lambda}^T(t_0)\delta\boldsymbol{x}(t_0) - \boldsymbol{\nu}^T\delta\boldsymbol{c} - L(t_0)dt_0$$

- The rest of the process is the same as before:
  - Costate Equations: (choose $\boldsymbol{\lambda}\,(t)$ to eliminate $\delta\boldsymbol{x}\,(t)$)

$$\dot{\boldsymbol{\lambda}}^T\,(t) = -\frac{\partial H}{\partial \boldsymbol{x}\,(t)} \qquad \boldsymbol{\lambda}^T(t_f) = \frac{\partial \Phi}{\partial x(t_f)}$$

  - Stationarity Condition:

$$\frac{\partial H}{\partial \boldsymbol{u}\,(t)} = 0$$

– Constraints:

1. Dynamic $\Rightarrow$ $\dot{\boldsymbol{x}}(t) = \boldsymbol{f}\{\boldsymbol{x}(t), \boldsymbol{u}(t), t\}$
2. Terminal $\Rightarrow$ $\boldsymbol{\psi}\left\{\boldsymbol{x}(t_f), t_f\right\} = \boldsymbol{c}$

– Transversality Condition (introduced because t_f is free):

$$\frac{\partial \Phi}{\partial t_f} + \frac{\partial \Phi}{\partial \boldsymbol{x}(t_f)}\dot{\boldsymbol{x}}(t_f) + L(t_f) = 0$$

– But,

$$\boldsymbol{\lambda}^T(t_f) = \frac{\partial \Phi}{\partial \boldsymbol{x}(t_f)} \quad \text{and} \quad \dot{\boldsymbol{x}}(t_f) = \boldsymbol{f}\left\{\boldsymbol{x}(t_f), \boldsymbol{u}(t_f), t_f\right\}$$

$$\frac{\partial \Phi}{\partial t_f} + \left(L + \boldsymbol{\lambda}^T(t)\,\boldsymbol{f}\right)_{t_f} = 0$$

or

$$H(t_f) = -\frac{\partial \Phi}{\partial t_f}$$

- Are these results surprising? NO!

1. We can think of the unspecified terminal time problem as a *family* of fixed terminal time problems from which we must select the one which minimizes the cost.

   $\Rightarrow$ So all conditions derived previously for the fixed time problem must still apply.

2. But there must be another condition (Transversality Condition) available to determine the optimal value of t_f.

5.10: Example: Free-Time Problem

$$\dot{x}_1(t) = x_2(t) \qquad x_1(0) = x_2(0) = 0$$
$$\dot{x}_2(t) = -x_2(t) + u(t)$$

$$J = \frac{1}{2}\left\{x_1(t_f) - 5\right\}^2 + \frac{1}{2}\left\{x_2(t_f) - 2\right\}^2 + \frac{1}{2}\int_0^{t_f} u(t)^2\,dt$$

$$x_1(t_f) + 5x_2(t_f) = 15$$

- What equations do we use to solve this problem? Same as before...

$$\lambda_1 = \lambda_1(0)$$
$$\lambda_2 = \lambda_1(0)\left(1 - e^t\right) + \lambda_2(0)\,e^t$$

$$x_1(t) = \lambda_1(0)\left(-t - 1/2e^{-t} + 1/2e^t\right) + \lambda_2(0)\left(1 - 1/2e^{-t} - 1/2e^t\right)$$
$$x_2(t) = \lambda_1(0)\left(-1 + 1/2e^{-t} + 1/2e^t\right) + \lambda_2(0)\left(1/2e^{-t} - 1/2e^t\right)$$

$$u(t) = -\lambda_2(t)$$

- which give...

$$x_1(t_f) - 5 + \nu = \lambda_1(0)$$
$$x_2(t_f) - 2 + 5\nu = \lambda_1(0)\left(1 - e^{t_f}\right) + \lambda_2(0)\,e^{t_f}$$

$$\begin{aligned} x_1(t_f) &= \lambda_1(0)\left(-t_f - 1/2e^{-t_f} + 1/2e^{t_f}\right) \\ &\quad + \lambda_2(0)\left(1 - 1/2e^{-t_f} - 1/2e^{t_f}\right) \\ x_2(t_f) &= \lambda_1(0)\left(-1 + 1/2e^{-t_f} + 1/2e^{t_f}\right) \\ &\quad + \lambda_2(0)\left(1/2e^{-t_f} - 1/2e^{t_f}\right) \end{aligned}$$

$$x_1(t_f) + 5x_2(t_f) = 15$$

- We have 5 equations but now with t_f we have 6 unknowns:

$$x_1\left(t_f\right),\ x_2\left(t_f\right),\ \lambda_1\left(0\right),\ \lambda_2\left(0\right)\ \nu,\ t_f$$

- The sixth equation is obtain from the transversality condition:

$$H(t_f) = -\frac{\partial \Phi}{\partial t_f} = 0\ \Rightarrow$$

$$\frac{1}{2}u\left(t_f\right)^2 + \left(\lambda_1\left(t_f\right) - \lambda_2\left(t_f\right)\right)x_2\left(t_f\right) + \lambda_2\left(t_f\right)u\left(t_f\right) = 0$$

$$= \left\{\lambda_1(t_f) - \lambda_2(t_f)\right\}x_2(t_f) - \frac{1}{2}\lambda_2^2(t_f) = 0$$

- Are these problems easy to solve? In general — NO!
- A variety of numerical techniques exist – but this remains an active area of research.

5.11: Continuous Time Optimization: Minimum Time Problems

- A special case of the free terminal time problem is the *minimum time* problem.
  - The goal is to minimize the elapsed time needed to transfer a system from a specified initial state to a specified final condition.
  - For this problem to make sense, at least one state must be specified at $t = t_0$ and at least one constraint must be specified at $t = t_f$ (i.e., we must have to do *something* in minimum time, or we won't do anything!)
  - This is a *constrained, free terminal time* problem; so all of the techniques developed previously apply.
  - But what is the cost?

$$J = t_f - t_0 = \int_{t_0}^{t_f} 1 dt$$

so, $\varphi = 0$ and $L = 1$

Example: Brachistochrone Problem ("shortest time")

- This problem is a variant of the discrete-time brachistochrone problem introduced earlier in the course.
- Like before, a mass m moves in a constant gravity field of magnitude g starting from rest at the origin.
- In this version, however, we wish to find the *minimum time* path (instead of the maximum distance) to reach a specified final point $\left(x_b, \; y_b \right)$.
- As before, the control is the tangent angle θ.

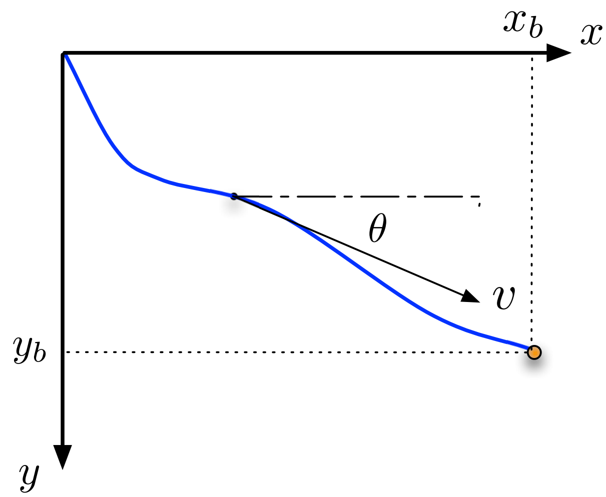


Figure 5.9 Brachistochrone Problem: Continous-Time

- Equations of motion are given by,

$$\dot{x} = v\cos\theta$$
$$\dot{y} = v\sin\theta$$

- But energy must be conserved, so

$$\frac{1}{2}mv^2 - mgy = 0$$

$$\Rightarrow \quad v = \sqrt{2gy}$$

- For minimum time, the cost is,

$$J = \int_{t_0}^{t_f} 1dt$$

- Solving,

$$\bar{J} = \nu_x\left\{x(t_f) - x_b\right\} + \nu_y\left\{y(t_f) - y_b\right\} + \int_{t_0}^{t_f}\left\{1 + \lambda_x(v\cos\theta - \dot{x}) + \lambda_y(v\sin\theta - \dot{y})\right\}dt$$

$$H = 1 + \lambda_x(v\cos\theta) + \lambda_y(v\sin\theta)$$

$$\dot{\lambda}_x = -\frac{\partial H}{\partial x} = 0$$
$$\dot{\lambda}_y = -\frac{\partial H}{\partial y} = \left\{\lambda_x \cos\theta + \lambda_y \sin\theta\right\} \frac{\partial v}{\partial y} = \frac{g}{v}\left\{\lambda_x \cos\theta + \lambda_y \sin\theta\right\}$$

$$\frac{\partial H}{\partial u} = \frac{\partial H}{\partial \theta} = -\lambda_x v \sin\theta + \lambda_y v \cos\theta = 0 \;\Rightarrow\; \frac{\lambda_y}{\lambda_x} = \tan\theta$$

$$H(t_f) = -\frac{\partial \Phi}{\partial t_f} = 0 \quad \text{and} \quad \frac{\partial H}{\partial t} = 0$$

$\Rightarrow H(t)$ is constant along the optimal path

- Therefore,

$$H(t) = 1 + \lambda_x v \cos\theta + \lambda_y \cdot v \sin\theta = 0$$

$$\begin{aligned} \dot{x} &= v\cos\theta & x(0) &= y(0) = 0 \\ \dot{y} &= v\sin\theta & & \\ x(t_f) &= x_b & y(t_f) &= y_b \\ \lambda_x(t_f) &= \nu_x & \lambda_y(t_f) &= \nu_y \end{aligned}$$

$$\left.\begin{aligned} \frac{\lambda_y}{\lambda_x} &= \tan\theta \\ \lambda_x v \cos\theta + \lambda_y v \sin\theta &= -1 \end{aligned}\right\} \Rightarrow \quad v\cos\theta + v\sin\theta\tan\theta = \frac{-1}{\lambda_x}$$

$$\lambda_x = \frac{-\cos\theta}{v}$$
$$\lambda_y = \frac{-\sin\theta}{v}$$

$$\dot{\lambda}_x = \frac{v(\sin\theta)\,\dot{\theta} + {}^{g}\!/_{v}\,(\cos\theta)\,\dot{y}}{v^2} = 0 \quad (\lambda_x \text{ is constant})$$

$$\Rightarrow \quad \dot{\theta} = -\frac{g}{v^2} \cdot \frac{\cos\theta}{\sin\theta} \cdot v\sin\theta = -\frac{g}{v}\cos\theta$$

- So, we have equations for $\dot{x},\ \dot{y},\ \dot{\theta}$, and 4 boundary conditions to solve for $x(t),\ y(t),\ \theta(t)$, and t_f ! But it's messy...
- Instead, let's treat θ as our *independent* variable.
  - $\lambda_x =$ constant $\Rightarrow$

$$\frac{\cos\theta(t)}{v(t)} = \frac{\cos\theta_f}{v_f}$$

$$\cos\theta = \frac{v}{v_f} \cdot \cos\theta_f = \sqrt{\frac{y}{y_b}} \cdot \cos\theta_f$$

$$\Rightarrow y(t) = y_b \cdot \frac{\cos^2\theta}{\cos^2\theta_f}$$

  - $\dot{x} = v\cos\theta \Rightarrow$

$$\frac{dx}{d\theta} \cdot \dot{\theta} = v\cos\theta$$

$$\frac{dx}{d\theta} = -\frac{v^2}{g} = -2y$$

$$\Rightarrow \frac{dx}{d\theta} = -\frac{2y_b}{\cos^2\theta_f} \cdot \cos^2\theta$$

$$x(t) = \int_{\theta(t)}^{\theta_f} \frac{dx}{d\theta} d\theta$$

$$\Rightarrow x(t) = x_b + \frac{y_b}{2\cos^2\theta_f}\left\{2\left(\theta_f - \theta\right) + \sin 2\theta_f - \sin 2\theta\right\}$$

$$y(0) = 0 \Rightarrow \theta(0) = 90^\circ$$

$$x(0) = 0 \;\Rightarrow\; \text{solve for } \theta_f$$

– if we know $x(t)$ and $y(t)$, we can find $\theta(t)$. A feedback law!

- Time-to-go can now be calculated by integrating $\dot{\theta} = g\lambda_x =$ constant to obtain

$$\theta_f - \theta(t) = g\lambda_x \left(t_f - t\right)$$

- Evaluating λ_x at t_f gives

$$t_f - t = \sqrt{\frac{2y\left(t_f\right)}{g}} \left(\frac{\theta(t) - \theta_f}{\cos\left(\theta_f\right)}\right)$$

- Two important observations:
  – The rate-of-chage of the optimal path angle is a constant
  – The initial path angle $\theta(0)$ is always $90°$
- The optimal path was computed for the case $x_b = y_b = 10$ and is shown in the accompanying plot.

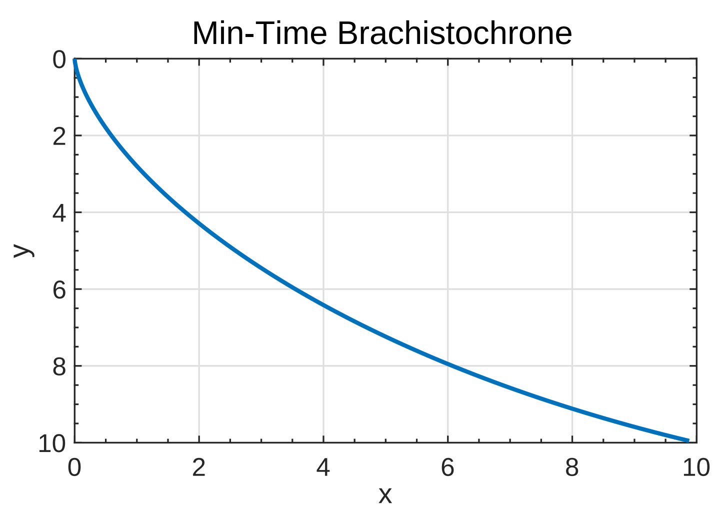


Figure 5.10 Brachistochrone Min-Time Trajectory

5.12: Minimum-Time Example: Zermelo's Problem

- Consider a ship travelling through a region of strong currents subject to the following conditions:
  1. Velocity of the current in the x-direction is a linear function of y
  2. Velocity of the current in the y-direction is zero
  3. Ship has constant speed (v) , but can change its heading (θ)
- Find the *minimum time path* from a given initial position to a specified final position.

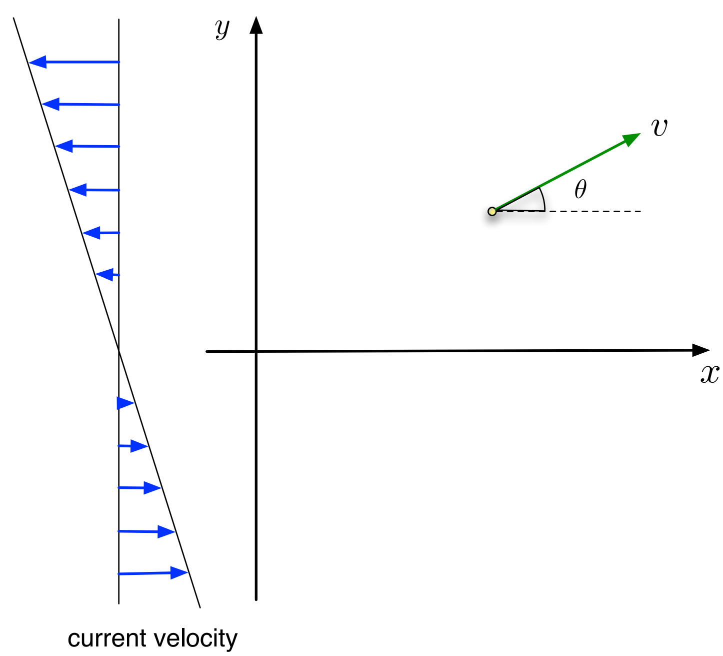


Figure 5.11 Zermelo's Problem

- Equations of motion:

$$\dot{x} = v\cos\theta + u$$
$$\dot{y} = v\sin\theta$$
$$u = -\frac{v}{h}\cdot y$$

where h is a normalizing constant.

- Cost is given by,

$$J = \int_{t_0}^{t_f} 1 dt$$

with terminal condition,

$$x(t_f) = y(t_f) = 0.$$

- Terminal constraint and Hamiltonian are given by,

$$\Phi = \nu_x x(t_f) + \nu_y y(t_f)$$
$$H = 1 + \lambda_x \left\{ v\cos\theta - \frac{v}{h}\cdot y \right\} + \lambda_y v \sin\theta$$

- The costate equations may be written,

$$\dot{\lambda}_x = -\frac{\partial H}{\partial x} = 0, \qquad \lambda_x(t_f) = \nu_x$$
$$\dot{\lambda}_y = -\frac{\partial H}{\partial y} = \lambda_x \cdot \frac{v}{h}, \qquad \lambda_y(t_f) = \nu_y$$

$$\lambda_x = k_1$$
$$\lambda_y = \frac{k_1 v}{h}\cdot t + k_2$$

- And the state equations,

$$\dot{x} = v\cos\theta - \frac{v}{h}\cdot y \quad x(0) = x_0$$
$$\dot{y} = v\sin\theta \quad y(0) = y_0$$

- The stationarity condition becomes,

$$\frac{\partial H}{\partial \theta} = -\lambda_x v \sin\theta + \lambda_y v\cos\theta = 0 \;\Rightarrow\; \tan\theta = \frac{\lambda_y}{\lambda_x}$$

- And the transversality condition,

$$H(t_f) = -\frac{\partial \Phi}{\partial t_f} = 0 \;\Rightarrow\; \lambda_x \left\{ v\cos\theta - \frac{v}{h}\cdot y \right\} + \lambda_y v \sin\theta = -1$$

- Terminal Constraints:

$$x(t_f) = y(t_f) = 0$$

- It's a messy process to try to identify everything in terms of t; so let's eliminate t and make θ the independent variable.
- Since H is not an explicit function of time, $H(t) =$ constant,

$$H(t) = H(t_f) = 0$$

$$\begin{bmatrix} v\cos\theta - \frac{v}{h}\cdot y & v\sin\theta \\ -v\sin\theta & v\cos\theta \end{bmatrix} \begin{bmatrix} \lambda_x \\ \lambda_y \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

$$\lambda_x = \frac{-\cos\theta}{v\left(1 - \frac{y}{h}\cos\theta\right)} \qquad \lambda_y = \frac{-\sin\theta}{v\left(1 - \frac{y}{h}\cos\theta\right)}$$

$$\dot{\lambda}_x = -\frac{\partial H}{\partial x} = 0 = \frac{\partial \lambda_x}{\partial \theta}\cdot\dot{\theta} + \frac{\partial \lambda_x}{\partial y}\cdot\dot{y} = 0$$

$$\dot{\theta} = \frac{-\frac{\partial \lambda_x}{\partial y}}{\frac{\partial \lambda_x}{\partial \theta}}\cdot\dot{y} \qquad\qquad \dot{y} = v\sin\theta$$

$\Rightarrow$

$$\dot{\theta} = \frac{v}{h}\cdot\cos^2\theta$$

$$\dot{x} = v\cos\theta - \frac{v}{h}\cdot y = \frac{dx}{d\theta}\cdot\dot{\theta} \;\Rightarrow\; \frac{dx}{d\theta} = \frac{(h\cos\theta - y)}{\cos^2\theta}$$

$$\dot{y} = v\sin\theta = \frac{dy}{d\theta}\cdot\dot{\theta} \;\Rightarrow\; \frac{dy}{d\theta} = \frac{h\sin\theta}{\cos^2\theta}$$

$$\int_y^{y_f} dy = \int_\theta^{\theta_f} \frac{h\sin\theta}{\cos^2\theta}d\theta \quad\Rightarrow\quad y(t) = h\left\{\sec\theta(t) - \sec\theta(t_f)\right\}$$

$$\int_x^{x_f} dx = \int_\theta^{\theta_f} \left\{h\sec\theta - h\sec^3\theta + h\sec\theta_f\sec^2\theta\right\}d\theta$$

$$\int_\theta^{\theta_f} \sec\theta_f\sec^2\theta d\theta = \sec\theta_f\left\{\tan\theta_f - \tan\theta\right\}$$

$$\int_\theta^{\theta_f} \sec\theta d\theta = \ln\left\{\frac{\sec\theta_f + \tan\theta_f}{\sec\theta + \tan\theta}\right\}$$

$$\int_\theta^{\theta_f} \sec^3\theta d\theta = \frac{1}{2}\frac{\sin\theta}{\cos^2\theta}\Bigg|_\theta^{\theta_f} + \frac{1}{2}\int_\theta^{\theta_f}\sec\theta d\theta$$

$\Rightarrow$

$$x(t) = \frac{h}{2}\left\{\sec\theta(t_f)\left\{\tan\theta(t) - \tan\theta(t_f)\right\} + \tan\theta(t)\left\{\sec\theta(t_f) - \sec\theta(t)\right\}\right.$$

$$\left. + \ln\left[\frac{\tan\theta(t) + \sec\theta(t)}{\tan\theta(t_f) + \sec\theta(t_f)}\right]\right\}$$

- So if $x(t)$ and $y(t)$ are known, $\theta(t)$ and $\theta(t_f)$ can be calculated
- Feedback Control Law:

$$\lambda_x = \text{constant} \;\Rightarrow\; \frac{\cos\theta(t)}{v\left\{1 - \frac{y(t)}{h}\cos\theta(t)\right\}} = \frac{\cos\theta(t_f)}{v}$$

$\Rightarrow$

$$\cos\theta(t) = \frac{\cos\theta(t_f)}{1 + \frac{y(t)}{h}\cos\theta(t_f)}$$

- The value of $\theta(t_f)$ can be identified using initial conditions, but the process isn't easy!
  - It is necessary here to solve for $\theta\left(t_f\right)$ and $\theta\,(0)$ simultaneously.
  - One approach is to generate a family of paramtetric curves for various values of $\theta\left(t_f\right)$ and $\theta\,(0)$, a representative plot is shown.

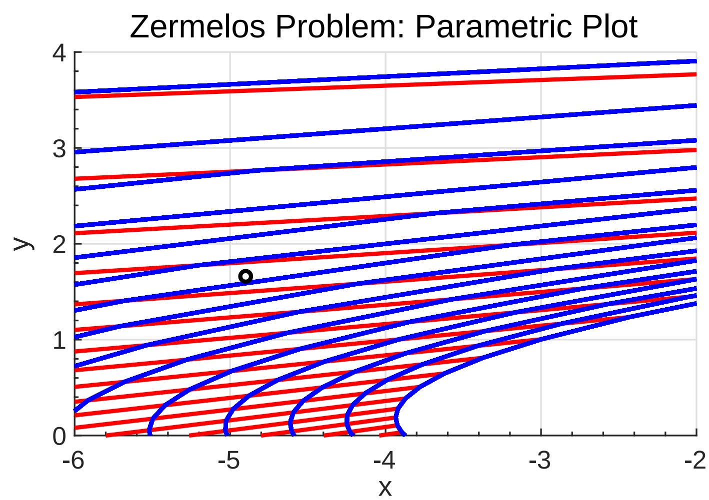


Figure 5.12 Paramtric Plot for Zermelo's Problem

  - From such a plot, one can read off corresponding values of $\theta\left(t_f\right)$ and $\theta\,(0)$ for given initial x and y values.
- Time-to-go can also be found as:

$$\dot{\theta} = \frac{v}{h}\cos^2\theta \quad \Rightarrow \quad \frac{v}{h}\left(t_f - t\right) = \tan\theta_f - \tan\theta$$

- An example was run using as initial conditions $x_0 = -4.9$ and $y_0 = 1.66$; h was set equal to one.
- For these values, reading off the parametric plot, it was found that

$$\theta\,(0) = 4.995\,\text{rad}$$
$$\theta\left(t_f\right) = 1.160\,\text{rad}$$

- Using these values, the optimal trajectory was computed using the equations obtained above for $x(t)$ and $y(t)$; results are shown in the accompanying plot.

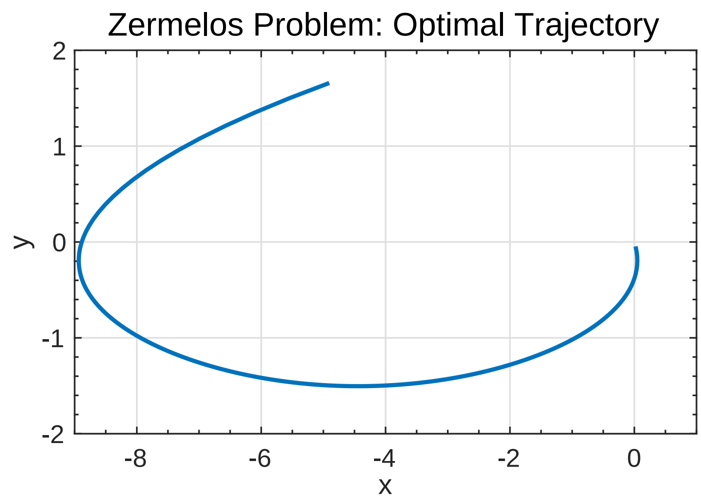


Figure 5.13 Zermelo's Problem: Optimal Trajectory

5.13: Discrete-TIme Optimization: Minimum Time Problem

- As we've seen in the previous sections, free-time and minimum-time problems are difficult to solve analytically; so in many cases, we use numerical techniques to obtain a solution.
- Here it is helpful to examine a discrete-time version of the min-time problem with terminal constraints in light of our goal to obtain a computer algorithm to solve it.

Problem:

For the discrete-time system with time interval Δ and terminal constraints $\boldsymbol{\psi}$,

$$\boldsymbol{x}[k+1] = \boldsymbol{f}\,\{\boldsymbol{x}[k],\, u[k],\, \Delta\} \quad \boldsymbol{x}[0] \text{ given}$$

$$J = N\Delta \quad (N \text{ fixed})$$

$$\boldsymbol{\psi}\,\{\boldsymbol{x}[N]\} = \boldsymbol{c}$$

$$\text{Find } \{u[0],\, u[1],\, \ldots,\, u[N-1]\} \text{ and } \Delta \text{ to minimize } J$$

NOTE:

- As before, at least one terminal condition must be specified to define the problem (i.e., minimum time to do what?)
- $\boldsymbol{x}[k+1]$ will generally be an explicit function of time Δ:

$$\boldsymbol{x}[k+1] = \boldsymbol{f}\,(\boldsymbol{x}[k],\, u[k],\, \Delta)$$

- The solution process varies from the one we've used so far because now we want to develop a computer algorithm to solve the problem.

- First, let's examine the effects of changing $u[k]$ and Δ on each of the terminal constraints ψ_i.
  - First, define a cost associated with each constraint,

$$J_i = \psi_i - c_i$$

  - Next, define a corresponding augmented cost function,

$$\bar{J}_i = \psi_i - c_i + \sum_{k=0}^{N-1} \lambda_i^T[k+1]\{\boldsymbol{f} - \boldsymbol{x}[k+1]\}$$

- Now, we can take the first variation of $\bar{J}_i$:

$$\begin{aligned}\delta\bar{J}_i &= \left\{\frac{\partial\psi_i}{\partial\boldsymbol{x}[N]} - \boldsymbol{\lambda}_i^T[N]\right\}\delta\boldsymbol{x}[N] \\ &+ \sum_{k=1}^{N-1}\left\{\boldsymbol{\lambda}_i^T[k+1]\frac{\partial\boldsymbol{f}}{\partial\boldsymbol{x}[k]} - \boldsymbol{\lambda}_i[k]\right\}\delta\boldsymbol{x}[k] \\ &+ \sum_{k=0}\boldsymbol{\lambda}_i^T[k+1]\frac{\partial f}{\partial\boldsymbol{u}[k]}\delta\boldsymbol{u}[k] + \sum_{k=0}^{N-1}\boldsymbol{\lambda}_i^T[k+1]\frac{\partial\boldsymbol{f}}{\partial\Delta}\delta\Delta\end{aligned}$$

$$\delta J_i = \left\{\frac{\partial\psi_i}{\partial x[N]} - \lambda_i^T[N]\right\}\delta x[N]$$

  - Here, $\boldsymbol{\lambda}_i^T[k+1]\dfrac{\partial\boldsymbol{f}}{\partial\boldsymbol{u}[k]}$ indicates the effect of changing $\boldsymbol{u}[k]$ on the i^{th} constraint and $\boldsymbol{\lambda}_i^T[k+1]\dfrac{\partial\boldsymbol{f}}{\partial\Delta}$ indicates the effect of changing Δ on the i^{th} constraint.

- The following difference equation defines λ_i:

$$\boldsymbol{\lambda}_i[k] = \left\{\frac{\partial\boldsymbol{f}}{\partial\boldsymbol{x}[k]}\right\}^T\boldsymbol{\lambda}_i[k+1] \qquad \boldsymbol{\lambda}_i[N] = \frac{\partial\psi_i}{\partial x[N]}^T$$

- Having developed expressions for $\delta \bar{J}_i$, we can adjoin these to the first variation of the original cost and attempt to set the resulting equation to zero:

$$\delta J + \sum_{i=1}^{q} \nu_i \delta \bar{J}_i = N\delta\Delta + \left[\sum_{i=1}^{q} \nu_i \left\{ \sum_{k=0}^{N-1} \boldsymbol{\lambda}_i^T[k+1] \frac{\partial \boldsymbol{f}}{\partial \Delta} \right\} \right] \delta\Delta + \sum_{k=0}^{N-1} \left\{ \sum_{i=1}^{q} \nu_i \boldsymbol{\lambda}_i^T[k+1] \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}[k]} \right\} \delta \boldsymbol{u}[k]$$

  – Or in matrix notation,

$$\delta J = \boldsymbol{\nu}^T \delta \bar{J} = \left(N + \boldsymbol{\nu}^T H_\Delta \right) \delta\Delta + \sum_{k=0}^{N-1} \boldsymbol{\nu}^T H_u[k] \delta \boldsymbol{u}[k]$$

  where H_Δ is a column vector whose i^{th} element is

$$\sum_{k=0}^{N-1} \boldsymbol{\lambda}_i^T[k+1] \frac{\partial \boldsymbol{f}}{\partial \Delta}$$

  and $H_u[k]$ is a matrix whose i^{th}row is $\boldsymbol{\lambda}_i^T[k+1] \dfrac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}[k]}$

  – Using the results above, we now have enough information to establish and implement a *gradient algorithm* (as we've done before) to solve the problem:

- STEP 1: Guess Δ and $\{u[k];\ k = 0, 1, \ldots, N-1\}$
- STEP 2: Compute $\{\boldsymbol{x}[k];\ k = 1, 2, \ldots, N\}$ using the dynamic constraints defined by the difference equations:

$$\boldsymbol{x}[k+1] = \boldsymbol{f}\ \{\boldsymbol{x}[k],\ u[k],\ \Delta\} \qquad \boldsymbol{x}[0] \text{ given}$$

- STEP 3: Compute $\{\lambda_i[k];\ k = 0, 1, \ldots, N-1, \quad i = 1, 2, \ldots, q\}$ using the costate difference equations:

$$\boldsymbol{\lambda}_i[k] = \left\{\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}[k]}\right\}^T \boldsymbol{\lambda}_i[k+1] \qquad \boldsymbol{\lambda}_i[N] = \frac{\partial x_i}{\partial \boldsymbol{x}[N]}^T$$

- STEP 4: Compute H_Δ and $\{H_u[k];\ k = 0,\ 1,\ \ldots,\ N-1\}$ using expressions developed previously.
- STEP 5: Based on the available information, update Δ and $\{\boldsymbol{u}[k];\ k = 0,\ 1,\ \ldots,\ N-1\}$ and return to Step 2.
- How?

$$\delta\Delta = -K_\Delta \left\{N + \boldsymbol{\nu}^T H_\Delta\right\} \qquad \delta\boldsymbol{u}[k] = -K_u H_u^T[k]\boldsymbol{\nu}$$

  – $K_\Delta > 0$ and $K_u > 0$ should be selected based on your physical understanding of the problem so that $\delta\Delta$ and $\delta\boldsymbol{u}[k]$ do not violate our first-order assumptions.
- Why? If $\delta\Delta$ and $\delta\boldsymbol{u}[k]$ are chosen in this manner, then

$$\delta J + \boldsymbol{\nu}^T \delta J = -K_\Delta \left(N + \boldsymbol{\nu}^T H_\Delta\right)^2 - K_u \left\{\sum_{k=0}^{N-1} \boldsymbol{\nu}^T H_u[k] H_u^T[k]\boldsymbol{\nu}\right\}$$

  – So the cost will be reduced by choosing these values for δ_Δ and $\delta\boldsymbol{u}[k]$.
- But to compute δ_Δ and $\delta\boldsymbol{u}[k]$, we need $\boldsymbol{\nu}$!

$$\begin{aligned}\delta\bar{J} &= H_\Delta \delta\Delta + \sum_{k=0}^{N-1} H_u[k]\delta u[k] \\ &= -K_\Delta N H_\Delta - K_\Delta H_\Delta H_\Delta^T \boldsymbol{\nu} - \sum_{k=0}^{N-1} K_u H_u[k] H_u^T[k]\boldsymbol{\nu} \\ &= -K_\Delta \boldsymbol{q} - K_\Delta Q \boldsymbol{\nu}\end{aligned}$$

where

$$q = NH_{\Delta} \quad Q = H_{\Delta}H_{\Delta}^{T} + \frac{K_u}{K_{\Delta}}\sum_{k=0}^{N-1} H_u[k]H_u^{T}[k]$$

- So,

$$\boldsymbol{\nu} = -Q^{-1}\left\{\boldsymbol{q} + \frac{\partial \bar{J}}{K_{\Delta}}\right\}$$

- What is $\delta\bar{J}$?

$$\bar{J} = \boldsymbol{\psi} - \boldsymbol{c}$$

$$\bar{J}_{opt} = \bar{J} + \delta\bar{J} \;\Rightarrow\; \delta\bar{J} = -\bar{J}$$

To determine $\boldsymbol{\nu}$, calculate $\boldsymbol{\psi} - \boldsymbol{c}$ for given $\boldsymbol{x}[N]$.

QUESTION: When do we stop?

- Looking at the first variation of the cost, we find that

$$N + \boldsymbol{\nu}^{T}H_{\Delta} = 0 \qquad \boldsymbol{\nu}^{T}H_u[k] = 0$$

- So, stop when

$$\left|N + \boldsymbol{\nu}^{T}H_{\Delta}\right| < \epsilon_1 \qquad \text{and} \qquad \frac{1}{N}\left\{\sum_{k=0}^{N-1}\boldsymbol{\nu}^{T}H_u[k]H_u^{T}[k]\boldsymbol{\nu}\right\} < \epsilon_2$$

5.14: Continuous-Time Optimization: Equality Path Constraints

- Up to now, the only constraints that we've included in the optimization process (apart from the dynamic constraints imposed by the system) have been *end point* constraints.
- Our goal in this section is to investigate the solution process in which constraints exist *along the entire trajectory* or at intermediate points (so-called "path constraints")

Integral Equality Contraints

- The first of these problems is the INTEGRAL CONSTRAINT problem:

$$\dot{\boldsymbol{x}} = \boldsymbol{f}\{\boldsymbol{x},\, \boldsymbol{u},\, t\} \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0$$

$$J = \varphi\left\{\boldsymbol{x}_f,\, t_f\right\} + \int_{t_0}^{t_f} L\{\boldsymbol{x},\, \boldsymbol{u},\, t\}\, dt$$

$$\boldsymbol{\psi}\left\{\boldsymbol{x}_f,\, t_f\right\} = \boldsymbol{c}$$

$$\int_{t_0}^{t_f} N\{\boldsymbol{x},\, \boldsymbol{u},\, t\}\, dt = k \quad \text{(NEW constraint)}$$

- To solve this problem, we define a new additional state, $x_{n+1}(t)$, which satisfies the following equations:

$$\dot{x}_{n+1} = N\{\boldsymbol{x},\, \boldsymbol{u},\, t\} \qquad x_{n+1}(0) = 0 \qquad x_{n+1}(t_f) = k$$

- By augmenting the state vector with this additional state, we can transform the new integral constraint into a form that we already know how to handle
- The solution process is now identical to the one we've already developed:

1. Adjoin the constraints to the cost

$$\bar{J} = \varphi + \boldsymbol{\nu}^T \{\boldsymbol{\psi} - \boldsymbol{c}\} + \nu_{q+1}\left\{x_{n+1}(t_f) - k\right\}$$
$$+ \int_{t_0}^{t_f} \left\{L + \boldsymbol{\lambda}^T(\boldsymbol{f} - \dot{\boldsymbol{x}}) + \gamma\left(N - \dot{x}_{n+1}\right)\right\} dt$$

2. Take the first variation of $\bar{J}$ (remembering to integrate by parts):

$$\delta\bar{J} = \left\{\frac{\partial\varphi}{\partial\boldsymbol{x}(t_f)} + \boldsymbol{\nu}^T\frac{\partial\boldsymbol{\psi}}{\partial\boldsymbol{x}(t_f)} - \boldsymbol{\lambda}^T(t_f)\right\}\delta\boldsymbol{x}(t_f)$$
$$+ \left\{\nu_{q+1} - \gamma(t_f)\right\}\delta x_{n+1}(t_f)$$
$$+ \int_{t_0}^{t_f}\left\{\left(\frac{\partial H}{\partial\boldsymbol{x}} + \dot{\boldsymbol{\lambda}}\right)\delta\boldsymbol{x} + \left(\frac{\partial H}{\partial x_{n+1}} + \dot{\gamma}\right)\delta x_{n+1} + \frac{\partial H}{\partial\boldsymbol{u}}\delta\boldsymbol{u}\right\} dt$$

where $H = L + \boldsymbol{\lambda}^T\boldsymbol{f} + \gamma N$

NOTE: $\displaystyle\int_{t_0}^{t_f}\boldsymbol{\lambda}\delta\dot{\boldsymbol{x}}\,dt = \boldsymbol{\lambda}^T\delta\boldsymbol{x}\Big|_{t_0}^{t_f} - \int_{t_0}^{t_f}\dot{\boldsymbol{\lambda}}^T\delta\boldsymbol{x}\,dt$

3. Set each coefficient in $\delta\bar{J}$ to zero to identify the equations required to solve this problem:

$$\dot{\boldsymbol{\lambda}}^T = -\frac{\partial H}{\partial\boldsymbol{x}}, \qquad \lambda^T(t_f) = \frac{\partial\varphi}{\partial\boldsymbol{x}(t_f)} + \boldsymbol{\nu}^T\frac{\partial\boldsymbol{\psi}}{\partial\boldsymbol{x}(t_f)}$$

$$\dot{\gamma} = -\frac{\partial H}{\partial x_{n+1}} = 0, \qquad \gamma(t_f) = \nu_{q+1}$$

$$\frac{\partial H}{\partial\boldsymbol{u}} = 0$$

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, t), \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0$$

$$\boldsymbol{\psi}\left(\boldsymbol{x}_f, t_f\right) = \boldsymbol{c}$$

$$x_{n+1}(t_f) = k$$

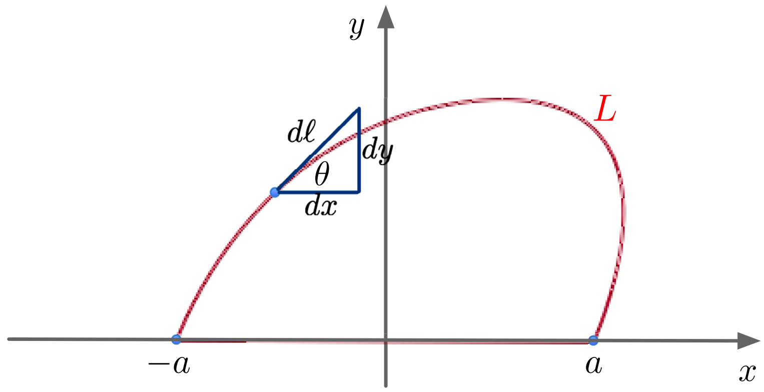
NOTE: $2n + m + q + 1$ equations in $2n + m + q + 1$ unknowns

5.15: Maximum Area with Fixed Perimeter: Dido's Problem

- Dido was the founder of Carthage (present-day Tunisia) who, upon arriving at the coast in 814 BC, asked for a parcel of land.
- She was granted her request under the condition that the land she was given could be encompassed by the hide of an ox.
- Giving this some thought, Dido cut the ox-hide into a long thin strip and used it to encircle the land – which became known as Carthage and she its Queen.
- The problem that bears her name ("Dido's Isoperimetric Problem") is to find the closed curve which has the maximum area for a given perimeter length.

Figure 5.14 *Dido Purchases Land for the Foundation of Carthage.* Engraving by Matthaus Merian the Elder, in Historiche Chronica Frankfurt a.M., 1630.

PROBLEM SET-UP:



- Independent variable: x (Note that 'dynamic' type problems do not require time as the independent variable)
- Dependent variable: y (In this problem both independent and dependent variables are spatial)
- Control variable: θ (In the same spirit as Zermelo's problem)
- Dynamic constraint: $\dfrac{dy}{dx} = \tan\theta$ [NOTE: $-\pi/2 < \theta \le \pi/2$]
  - Constraint defines the 'control input' for the problem.
- Integral constraint: $L = \displaystyle\int_{-a}^{a} \frac{1}{\cos\theta} dx$ $[dx = d\ell \cos\theta\ ]$
  - Constraint relates the total perimeter length to the input.
- Terminal constraint: $y(x_f) = y(a) = y(-a) = 0$
  - Constraint ensures area bounds begin and end at the 'coastline'.
- Cost function: $J = -\displaystyle\int_{-a}^{a} y\, dx$ $\Rightarrow$ minimize total area, $-A$

SOLUTION:

1. Define state:

$$z(x) = \text{length}$$

$$\frac{dz(x)}{dx} = \frac{1}{\cos(\theta)} = \sec\theta \qquad z(a) = L$$

2. Adjoin constraints to cost function:

$$\bar{J} = \nu_y y(a) + \nu_z \{z(a) - L\} + \int_{-a}^{a} \left\{ -y + \lambda_y \left(\tan\theta - \frac{dy}{dx} \right) + \gamma \left(\sec\theta - \frac{dz}{dx} \right) \right\} dx$$

3. Take the first variation of $\bar{J}$:

$$\delta\bar{J} = \left(\nu_y - \lambda_y\right)\delta y(a) + (\nu_z - \gamma)\,\delta z(a) + \int_{-a}^{a} \left\{ \left(\frac{\partial H}{\partial y} + \dot{\lambda}_y \right)\delta y + \left(\frac{\partial H}{\partial z} + \dot{\gamma} \right)\delta z + \frac{\partial H}{\partial\theta}\delta\theta \right\} dx$$

where $H = -y + \lambda_y \tan\theta + \gamma \sec\theta$

4. Identify the equations to be solved:

$$\begin{aligned} \dot{\lambda}_y &= 1; & \lambda_y(a) &= \nu_y & \Rightarrow \lambda_y &= x + k_1 \\ \dot{\gamma} &= 0; & \gamma(a) &= \nu_z & \Rightarrow \gamma &= \nu_z \end{aligned}$$

$$\frac{\partial H}{\partial\theta} = \lambda_y \sec^2\theta + \gamma \tan\theta \sec\theta = 0 \quad \Rightarrow \quad \sin\theta = -\frac{\lambda_y}{\gamma}$$

$$\begin{aligned} dy &= \tan\theta dx & \lambda_y &= x + k_1 & &= -\gamma\sin\theta \\ dy &= -\gamma\sin\theta d\theta & x &= -\gamma\sin\theta - k_1 \\ y &= \gamma\cos\theta + k_2 & dx &= -\gamma\cos\theta d\theta \end{aligned}$$

$$\int_{-a}^{a} \left(\frac{1}{\cos\theta} \right) dx = \int_{\theta_i}^{\theta_f} -\gamma d\theta = -\gamma\left(\theta_f - \theta_i\right) = L$$

- Unknowns: $k_1,\ k_2,\ \gamma,\ \theta_f,\ \theta_i$
- Equations:

$$x(\theta_i) = -a \qquad x(\theta_f) = a$$
$$y(\theta_i) = 0 \qquad y(\theta_f) = 0$$
$$L = -\gamma\left(\theta_f - \theta_i\right)$$

$$\begin{aligned}
&(1) \quad \gamma \sin\theta_i + k_1 = a \\
&(2) \quad -\gamma \sin\theta_f - k_1 = a \\
&(3) \quad \gamma \cos\theta_i + k_2 = 0 \\
&(4) \quad \gamma \cos\theta_f + k_2 = 0 \\
&(5) \quad -\gamma\left(\theta_f - \theta_i\right) = L
\end{aligned}$$

(3) & (4) $\rightarrow\ \theta_f = \pm\theta_i$, but (5) $\Rightarrow\ \theta_f = -\theta_i;\quad \gamma = \dfrac{L}{2\theta_i}$

(1) & (2) $\rightarrow\ k_1 = 0;\quad a = \gamma\sin\theta_i$

$$\Rightarrow \quad \frac{\sin\theta_i}{\theta_i} = \frac{2a}{L}$$

Since $-\dfrac{\pi}{2} < \theta < \dfrac{\pi}{2}$, this equation $\Rightarrow\ P < \pi a$

$$\begin{aligned}
k_1 &= 0 \\
\gamma &= \frac{L}{2\theta_i} \\
\theta_f &= -\theta_i \\
k_2 &= -\frac{L\cos\theta_i}{2\theta_i}
\end{aligned}$$

- What does all of this mean?

$$x = -\frac{L}{2\theta_i}\sin\theta \qquad y = \frac{L}{2\theta_i}\left\{\cos\theta - \cos\theta_i\right\}$$

$$\Rightarrow \quad x^2 + \left\{ y + \frac{L}{2\theta_i} \cos\theta_i \right\}^2 = \left(\frac{L}{2\theta_i} \right)^2$$

– This is the equation of a circle!

- Therefore, the rope forms a *circular arc* of radius $L/2\theta_i$ centered at:

$$\begin{aligned} x &= 0 \\ y &= -\frac{L}{2\theta_i} \cos\theta_i \end{aligned}$$

- So far, we've examined path constraints by looking at "integral contraints".
- We'll continue the process by looking at equality constraints that must be satisfied along the entire optimal path.

5.16: Control- and State-Only Equality Constraints

- The standard (fixed-time, terminal constraint) problem is the same, but now we add an additional constraint on the controls:

$$\boldsymbol{G}\{u,\, t\} = \boldsymbol{k} \qquad t_o \le t \le t_f$$

[NOTE: $m \ge 2$ or else completely specified by $\boldsymbol{G}$]

$$\bar{J} = \varphi + \boldsymbol{\nu}^T(\boldsymbol{\psi} - \boldsymbol{c}) + \int_{t_0}^{t_f} \left\{L + \boldsymbol{\lambda}^T(\boldsymbol{f} - \dot{\boldsymbol{x}}) + \boldsymbol{\mu}^T(\boldsymbol{G} - \boldsymbol{k})\right\} dt$$

$$H = L + \boldsymbol{\lambda}^T \boldsymbol{f} + \boldsymbol{\mu}^T \boldsymbol{G}$$

- Since $\boldsymbol{G}$ is not a function of $\boldsymbol{x}$, all of the equations developed previously are valid except

$$\frac{\partial H}{\partial \boldsymbol{u}} = \frac{\partial L}{\partial \boldsymbol{u}} + \boldsymbol{\lambda}^T \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}} + \boldsymbol{\mu}^T \frac{\partial \boldsymbol{G}}{\partial \boldsymbol{u}} = 0$$

$$\boldsymbol{G}(\boldsymbol{u},\, t) = \boldsymbol{k}$$

- These two sets of equations provide enough information to identify $\boldsymbol{u}$ and $\boldsymbol{\mu}$

Control and State Equality Constraints

- Again, the standard problem is the same, but now our path constraints take the form:

$$\boldsymbol{G}\{\boldsymbol{x},\, \boldsymbol{u},\, t\} = \boldsymbol{k}$$

- $\bar{J}$ and H are identical to those shown above for the control-only equality constraints, so the equations required to solve this problem are:

$$\dot{\boldsymbol{\lambda}} = -\frac{\partial H}{\partial \boldsymbol{x}}^T = -\left\{\frac{\partial L}{\partial \boldsymbol{x}} + \boldsymbol{\lambda}^T \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}} + \boldsymbol{\mu}^T \frac{\partial \boldsymbol{G}}{\partial \boldsymbol{x}}\right\}^T$$

$$\boldsymbol{\lambda}(t_f) = \frac{\partial \Phi}{\partial \boldsymbol{x}(t_f)} \qquad \text{where } \Phi = \varphi + \boldsymbol{\nu}^T [\boldsymbol{\psi} - \boldsymbol{c}]$$

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, t), \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0$$

$$\frac{\partial H}{\partial \boldsymbol{u}} = \frac{\partial L}{\partial \boldsymbol{u}} + \boldsymbol{\lambda}^T \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}} + \boldsymbol{\mu}^T \frac{\partial \boldsymbol{G}}{\partial \boldsymbol{u}} = 0$$

$$\boldsymbol{G}(\boldsymbol{x}, \boldsymbol{u}, t) = \boldsymbol{k}$$

$$\psi\left\{\boldsymbol{x}(t_f), \boldsymbol{u}, t\right\} = \boldsymbol{c}$$

- Unknowns: $\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}, \boldsymbol{\mu}, \boldsymbol{u}$

State-Only Equality Constraints

- Once again, the standard problem is the same; but now our path constraints are a function of $\boldsymbol{x}$ and t only

$$\boldsymbol{G}(\boldsymbol{x}, t) = \boldsymbol{k}$$

  – This is a somewhat more complicated problem because some of the elements of $\boldsymbol{x}(t)$ depend on other elements of $\boldsymbol{x}(t)$ as well as the previous $\boldsymbol{x}$ and $\boldsymbol{u}$.

- There are a number of different ways to solve this problem:

  – METHOD 1:

    ◦ Solve for one subset of the states as a function of the remaining states and time

- ◦ Reduce the dimension of the state vector using the solution derived above
- ◦ Problem with this approach $\rightarrow$ the choice of state subsets is not unique, so some choices may produce more difficulties than others

– METHOD 2:

- ◦ Convert the state-only constraints into control and state constraints
  - · Since $\boldsymbol{G}(\boldsymbol{x},\ t) = \boldsymbol{k}$ along the optimal path, ${}^{\partial \boldsymbol{G}}/_{\partial t}$ must be zero along the optimal path
$$\frac{\partial \boldsymbol{G}}{\partial t} = \frac{\partial \boldsymbol{G}}{\partial t} + \frac{\partial \boldsymbol{G}}{\partial \boldsymbol{x}}\frac{d\boldsymbol{x}}{dt} = \frac{\partial \boldsymbol{G}}{\partial t} + \frac{\partial \boldsymbol{G}}{\partial \boldsymbol{x}}\boldsymbol{f} = 0$$
  - · In general, $\boldsymbol{f}$ is a function of $\boldsymbol{u}$, so ${}^{d\boldsymbol{G}}/_{dt} = 0$ is a control and state constraint.
- ◦ But ${}^{d\boldsymbol{G}}/_{dt} = 0$ even when $\boldsymbol{G} \neq \boldsymbol{k}$, so we must add additional terminal constraints,
$$\boldsymbol{G}\left\{\boldsymbol{x}(t_f),\ t_f\right\} = \boldsymbol{k}$$
to ensure tha the proper constraint is applied
- ◦ If ${}^{dG}/_{dt}$ is not a function of u, additional derivatives can be taken and additional terminal constraints added.

– METHOD 3: A computational alternative using "soft" constraints

- ◦ Define a new cost function:
$$\bar{J} = J + K\int_{t_0}^{t_f}\{\boldsymbol{G} - \boldsymbol{k}\}^T\{\boldsymbol{G} - \boldsymbol{k}\}\,dt$$
- ◦ Select K to establish the proper trade-off between the various elements of the cost.

5.17: Continuous-Time Optimization: Inequality Constraints

- In many problems, we may not need to force the satisfaction of an equality constraint but instead may be forced to live with inequality constraints driven by physical attributes of the problem (e.g., limited fuel)
  - In most instances, these inequality constraints apply only to the available control variables.
  - So, we'll focus our attention on control-only inequality constrained problems:

$$J = \varphi + \int_{t_0}^{t_f} L dt$$

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, u, t)$$

$$\boldsymbol{\psi}\left\{\boldsymbol{x}(t_f), t_f\right\} = \boldsymbol{k}_1$$

$$\boldsymbol{c}\left\{\boldsymbol{u}(t), t\right\} \le k_2$$

STANDARD CALCULUS OF VARIATIONS APPROACH

$$J = \varphi + \boldsymbol{\nu}^T \{\boldsymbol{\psi} - \boldsymbol{k}_1\} + \int_{t_0}^{t_f} \left\{L + \boldsymbol{\lambda}^T(\boldsymbol{f} - \dot{\boldsymbol{x}}) + \boldsymbol{\mu}^T(\boldsymbol{c} - \boldsymbol{k}_2)\right\} dt$$

$$H = L + \boldsymbol{\lambda}^T \boldsymbol{f} + \boldsymbol{\mu}^T \boldsymbol{c}$$

- All of the equations developed previously still apply:

$$\dot{\boldsymbol{\lambda}}^T = -\frac{\partial H}{\partial \boldsymbol{x}} \qquad \boldsymbol{\lambda}^T(t_f) = \frac{\partial \Phi}{\partial \boldsymbol{x}(t_f)}$$

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}, t) \qquad \boldsymbol{x}(t_0) = \boldsymbol{x}_0$$

$$\frac{\partial H}{\partial \boldsymbol{u}} = 0$$

$$\boldsymbol{\psi} = \boldsymbol{k}_1$$

$$\boldsymbol{c} \leq \boldsymbol{k}_2$$

NOTE:

$$\boldsymbol{\mu} = 0 \quad \text{if constraints are inactive}$$
$$\boldsymbol{\mu} > 0 \quad \text{if constraints are active}$$

ALTERNATIVE APPROACH: PONTRYAGIN'S MINIMUM PRINCIPLE

- Russian mathematician Pontryagin demonstrated that the optimal control must minimize the function

$$H = \boldsymbol{\lambda}^T \boldsymbol{f}$$

  for all admissable controls.

- So? We already knew that we had to solve $\partial H / \partial \boldsymbol{u} = 0$
  - In many inequality constrained problems, finding the $\boldsymbol{u}$ which minimizes H is almost obvious.
  - And Pontryagin proved that this $\boldsymbol{u}$ is the optimal one, so we don't need to add the complexities associated with additional Lagrange multipliers!

Example:

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -x_2 + u$$

$$J = \int_{t_0}^{t_f} \frac{1}{2}\left(x_1^2 + u^2\right) dt$$

t_f specified $\quad\quad$ $\boldsymbol{x}(t_f)$ free

1. No Control Constraints:

$$H = \frac{1}{2}\left(x_1^2 + u^2\right) + \lambda_1 x_2 + \lambda_2 (u - x_2)$$

$$\begin{aligned}\dot{\lambda}_1 &= -x_1, \qquad \lambda_1(t_f) = 0\\ \dot{\lambda}_2 &= \lambda_2 - \lambda_1, \qquad \lambda_2(t_f) = 0\\ u &= -\lambda_2\end{aligned}$$

- So, we augment the states to solve for $\boldsymbol{x}$ and $\boldsymbol{\lambda}$

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{\lambda}_1\\ \dot{\lambda}_2\end{bmatrix} = \begin{bmatrix}0 & 1 & 0 & 0\\ 0 & -1 & 0 & -1\\ -1 & 0 & 0 & 0\\ 0 & 0 & -1 & -1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \lambda_1\\ \lambda_2\end{bmatrix}$$

  – Recall that $u = -\lambda_2$

$$\Rightarrow \quad \boldsymbol{z}(t) = \mathcal{L}^{-1}\left\{(sI - A)^{-1}\right\}\boldsymbol{z}(0)$$

  – And since $x_1(t_0),\ x_2(t_0),\ \lambda_1(t_f)$, and $\lambda_2(t_f)$ are known, $\lambda_1(0)$ and $\lambda_2(0)$ can be identified and used to describe $x(t)$, $\lambda(t)$, and $u(t)$.

2. Constrained Control: What if $|u| \le 1$?

$$H = \frac{1}{2}\left(x_1^2 + u^2\right) + \lambda_1 x_2 + \lambda_2 u - \lambda_2 x_2$$

- Using Pontryagin's Principle, the control that minimizes H is:

$$u = -\lambda_2, \qquad \text{provided } |u| \le 1$$

  – What if $\lambda_2 > 1$? Pick $u = -1$ to minimize H

  – What if $\lambda_2 < -1$? Pick $u = +1$ to minimize H

  – What if $-1 < \lambda_2 < 1$? Pick $u = -\lambda_2$ to minimize H

- The solution to this problem is not the same as that for the unconstrained problem because u may not be *continuous*
  - So, we may need to solve the problem in parts by piecing together constrained and unconstrained arcs.

5.18: Introduction to Linear Constraints

- A special case of the control inequality constraint problem in which Pontryagin's Principle plays a very important role occurs when the dynamic system constraints and the control variable inequality constraints are all *linear*:

$$\dot{\boldsymbol{x}} = A\boldsymbol{x} + Bu \qquad -1 \le u(t) \le 1$$

- We'll let u be a scalar in the follwing development to keep things simple, but the ideas are easily extended.

Example: Single-Axis Satellite Attitude Control Using Reaction Jets

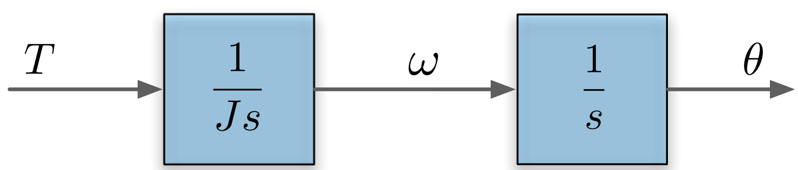


Figure 5.15 Single-Axis Satellite Attitude Control

- Define:

$$\begin{aligned} x_1 &= \theta \\ x_2 &= \omega \end{aligned}$$

- Equations of motion:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \frac{1}{2}T \end{aligned}$$

where T is the commanded input

- So, we can write,

$$\dot{\boldsymbol{x}} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 0 \\ \dfrac{1}{J} \end{bmatrix} u$$

$$u = T \quad \text{and} \quad -1 \le \frac{T}{T_{max}} \le 1$$

- How do we solve this problem? It depends on what we want to do

1. Minimum time $\Rightarrow$
   - Since the system is linear and the performance index is linear, we should expect that the optimal solution requires a control that lies on the boundary of the feasible region.
   - In addition, one or more changes in control may be required during operation $\rightarrow$ the control may suddently change from one point on the boundary to another [BANG-BANG CONTROL].
2. Minimum fuel $\Rightarrow$
   - Same linear system, but different linear cost $\rightarrow$ now saving fuel is more important than saving time..
   - At certain times, it may be beneficial to turn the reaction jets off [BANG-OFF-BANG CONTROL].
3. Minimum energy $\Rightarrow$
   - Quadratic cost $\rightarrow$ variable control in the feasible region.

- For the reaction jet problem, we have the state equation,

$$\begin{bmatrix} \dot{\theta} \\ \dot{\omega} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \theta \\ \omega \end{bmatrix} + \begin{bmatrix} 0 \\ \dfrac{1}{J} \end{bmatrix} T$$

where the torque generated by the reaction jets is limited to the range $-T_m \le T \le T_m$.

MINIMUM TIME PROBLEM

$$J = \int_{t_0}^{t_f} dt$$

with terminal constraints $\Rightarrow$ $\boldsymbol{x}(t_f) = 0$

$$\bar{J} = \boldsymbol{\nu}^T \boldsymbol{x}(t_f) + \int_{t_0}^{t_f} \left\{1 + \boldsymbol{\lambda}^T (A\boldsymbol{x} + B\boldsymbol{u} - \dot{\boldsymbol{x}})\right\} dt$$

$$\dot{\boldsymbol{\lambda}}^T = -\frac{\partial H}{\partial \boldsymbol{x}} = -\boldsymbol{\lambda}^T A \qquad \lambda^T(t_f) = \frac{\partial \Phi}{\partial x(t_f)} = \nu^T$$

$$\dot{\boldsymbol{x}} = A\boldsymbol{x} + B\boldsymbol{u} \qquad \boldsymbol{x}(0) = \boldsymbol{x}_0$$

$$H(t_f) = 1 + \boldsymbol{\lambda}^T(t_f)\left\{A\boldsymbol{x}(t_f) + B\boldsymbol{u}(t_f)\right\} = -\frac{\partial \Phi}{\partial t_f} = 0$$

$$\boldsymbol{x}(t_f) = 0$$

And

$$\frac{\partial H}{\partial \boldsymbol{u}} = 0 \quad (???)$$

But,

$$\frac{\partial H}{\partial \boldsymbol{u}} = \boldsymbol{\lambda}^T B \neq 0 \quad (unless\ \boldsymbol{\lambda}^T = 0)$$

So what does this mean? What do we do?

- The dilemma that has arisen is a direct result of the linear nature of the problem that we've set up
  - I the control were unconstrained, this condition $(\partial H / \partial \boldsymbol{u} = 0)$ would suggest that there is always a $\delta \boldsymbol{u}$ that can be selected to reduce the cost.

- But remember, our control is constrained.

- To identify the proper $\boldsymbol{u}$, we'll use Pontryagin's Principle,

$$\min_{\boldsymbol{u}} H = \min_{\boldsymbol{u}} \left\{1 + \boldsymbol{\lambda}^T A\boldsymbol{x} + \boldsymbol{\lambda}^T B\boldsymbol{u}\right\}$$

  - I can't do anything about $1 + \boldsymbol{\lambda}^T A\boldsymbol{x}$, but I can choose $\boldsymbol{u}$ so that $\boldsymbol{\lambda}^T B\boldsymbol{u}$ is as small as possible:

$$\boldsymbol{u} = -T_m \text{sgn}\left(\boldsymbol{\lambda}^T A\right)$$

  - $\boldsymbol{\lambda}^T A$ is a function of time and is often referred to as the "switching function" since it will determine when $\boldsymbol{u} = +T_m$ and $\boldsymbol{u} = -T_m$

- Additional details of the solution can now be obtained by examining the remaining equations:

$$\begin{aligned} \dot{\boldsymbol{\lambda}}^T = -\boldsymbol{\lambda}^T A \quad \Rightarrow \quad \dot{\lambda}_1 &= 0 & \Rightarrow \lambda_1 &= \text{constant} \\ \dot{\lambda}_2 &= -\lambda_1 & \Rightarrow \lambda_2 &= \lambda_2(0) - \lambda_1 t \end{aligned}$$

$$\lambda^T B = \frac{\lambda_2}{J} \quad \Rightarrow \quad u(t) \text{ depends on } \lambda_2(t)$$

  - Note that since $\lambda_2(t)$ is linear, $\boldsymbol{\lambda}^T A$ can change signs at most once $\Rightarrow$ the control, $u(t)$, will switch at most one time
  - Can we get more information about the control? YES

$$H(t_f) = 0 \quad \Rightarrow \quad \left\{\frac{\lambda_2(t_f)}{J}\right\} u(t_f) = -1$$

NOTE: $\boldsymbol{x}(t_f) = 0$ by the forced constraints

$\Rightarrow$

$$\lambda_2(t_f) = \frac{-1}{T_m} \quad \text{when } u(t_f) = T_m$$

$$\lambda_2(t_f) = \frac{1}{T_m} \quad \text{when } u(t_f) = -T_m$$

- Based on the results above, what possible optimal control strategies exist?

1. $u = -T_m$ for all $t \geq 0$
2. $u = -T_m$ for $t < t_f$; switches to $u = +T_m$ for $t \geq t_f$
3. $u = +T_m$ for $t < t_f$; switches to $u = -T_m$ for $t \geq t_f$
4. $u = +T_m$ for all $t \geq 0$

   – Which of these strategies do we use? ... It depends on θ and ω

FEEDBACK

$$\dot{\theta} = \omega$$
$$\dot{\omega} = \frac{u}{J}$$

$\Rightarrow$

$$\frac{d\omega}{d\theta} = \frac{u}{J\omega}$$
$$\omega d\omega = \frac{u}{J} d\theta$$
$$\omega^2 = \frac{2u}{J}\theta + c_0$$

$\Rightarrow$

$$\omega^2 = \pm\frac{2T_m}{J}\theta + c_0$$

which gives a family of parabolas

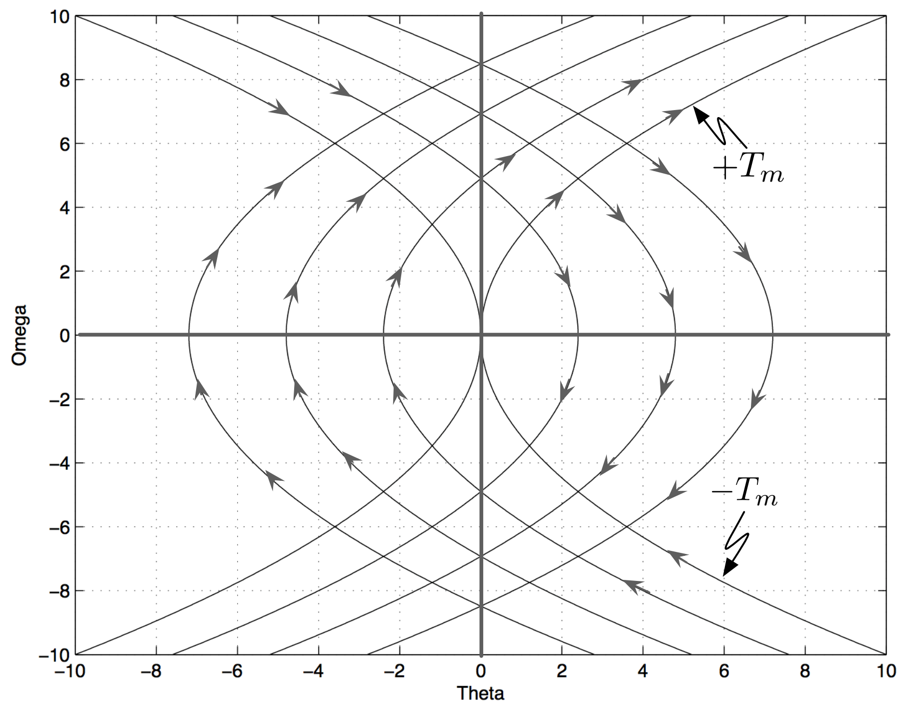


Figure 5.16 Family of Parabolas

- Two of these parabolas go through the origin and these are the ones we want to get on in order to get to $\boldsymbol{x}(t_f) = 0$

1. if $\omega < 0$, $\theta = \dfrac{J}{2T_m}\omega^2$
2. if $\omega > 0$, $\theta = -\dfrac{J}{2T_m}\omega^2$

$$\theta = -\frac{J}{2T_m}\omega\,|\omega| \quad \text{switching curve}$$

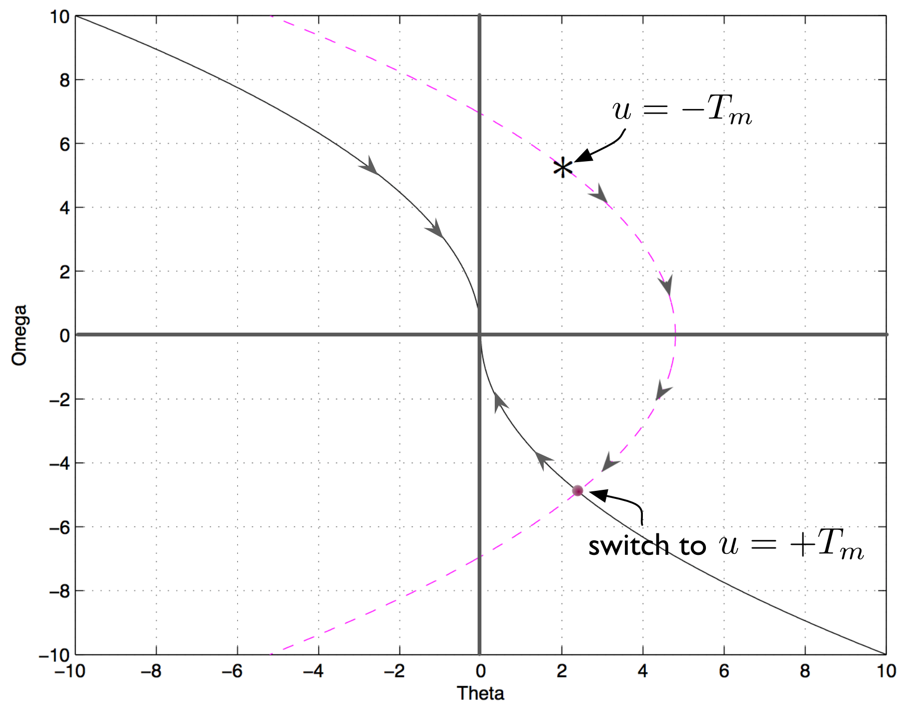


Figure 5.17 Switching Curve

- The FEEDBACK LAW is:
  1. $u = -T_m$ when $\theta > -\dfrac{J}{2T_m}\omega\,|\omega|$ or $\theta = -\dfrac{J}{2T_m}\omega\,|\omega|$ and $\theta < 0$
  2. $u = +T_m$ when $\theta < -\dfrac{J}{2T_m}\omega\,|\omega|$ or $\theta = -\dfrac{J}{2T_m}\omega\,|\omega| > 0$
- Can we solve for the times required to perform this maneuver?
  - YES, but using the dynamics of the satellite and knowledge of the control strategy.
  - Assume we start at $t = 0$ with some initial conditions $(\theta_0,\ \omega_0)$.
  - Since u is constant from 0 to t_s,.

$$\theta(t_s) = \theta_0 + \omega_0 t_s + \frac{u}{J}\frac{t_s^2}{2}$$

$$\omega(t_s) = \omega_0 + \frac{u}{J}t_s$$

$$u = \pm T_m$$

– At t_s^- (just before the switch),

$$\theta(t_s) = -\frac{J}{2u}\omega^2(t_s)$$

$$\theta_0 + \omega_0 t_s + \frac{u}{J}\frac{t_s^2}{2} = -\frac{J}{2u}\left\{\omega_0^2 + \frac{2u}{J}\omega_0 t_s + \left(\frac{u}{J}\right)^2 t_s^2\right\}$$

$$-\frac{J}{2u}\omega_0^2 - \omega_0 t_s - \frac{u}{2J}t_s^2 = \theta_0 + \omega_0 t_s + \frac{u}{2J}t_s^2$$

$$\frac{u}{J}t_s^2 + 2\omega_0 t_s + \theta_0 + \frac{J}{2u}\omega_0^2 = 0$$

$$t_s^2 + \frac{2J}{u}\omega_0 t_s + \left(\frac{J}{u}\right)^2\omega_0^2 - \left(\frac{J}{u}\right)^2\omega_0^2 + \frac{J}{u}\theta_0 + \frac{1}{2}\left(\frac{J}{u}\right)^2\omega_0^2 = 0$$

$$\left(t_s + \frac{J}{u}\omega_0\right)^2 = \left(\frac{J}{u}\right)^2\left\{\frac{\omega_0^2}{2} - \frac{u}{J}\theta_0\right\}$$

$$\Rightarrow \quad t_s = -\frac{J}{u}\left\{\omega_0 \pm \sqrt{\frac{\omega_0^2}{2} - \frac{u}{J}\theta_0}\right\}$$

– After the switch at t_s,

$$\omega(t_f) = \omega(t_s) - \frac{u}{J}\left(t_f - t_s\right) = 0$$

– But $\omega(t_f) = \omega_0 + \frac{u}{J}t_s \;\Rightarrow$

$$t_f = 2t_s + \frac{J}{u}\omega_0$$

MINIMUM FUEL PROBLEM:

- Same as above except

$$J = \int_{t_0}^{t_f} |u| \, dt$$

Note: t_f is specified and $t_f > t_{min}$!

MINIMUM ENERGY PROBLEM:

- Same as above except

$$J = \int_{t_0}^{t_f} u^2 dt$$

  - Because J is quadratic in u, the standard optimization techniques may be applied to solve this problem

$$\frac{\partial H}{\partial u} = 0 \quad \Rightarrow \quad \text{solve for } u$$

    - But remember, we must ensure that u doesn't violate its constraints

Appendix 5.A: Hamilton's Principle

- In mechanics, the motion of a conservative system from time t_0 to t_f is such that the integral

$$J = \int_{t_0}^{t_f} [T(t) - V(t)]\, dt$$

has a stationary value.

- Here we define:

$$\begin{aligned} T(t) &= \text{kinetic energy} \\ V(t) &= \text{potential energy} \\ \boldsymbol{x}(t) &= [q_1(t), q_2(t), \ldots, q_n(t)] \Rightarrow \text{vector of generalized coordinates} \\ \boldsymbol{u}(t) &= [\dot{q}_1(t), \dot{q}_2(t), \ldots, \dot{q}_n(t)] \Rightarrow \text{vector of generalized velocities} \end{aligned}$$

- A dynamic constraint is then given by:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{u}(t)$$

- The Hamiltonian is then

$$\begin{aligned} H(t) &= [T(t) - V(t)] + \boldsymbol{\lambda}^T(t)\,\boldsymbol{u}(t) \\ &= L(t) + \boldsymbol{\lambda}^T(t)\,\boldsymbol{u}(t) \end{aligned}$$

- Once again the necessary conditions for stationarity may be written,

$$\frac{\partial H(t)}{\partial \boldsymbol{u}(t)} = 0$$

$$\dot{\boldsymbol{\lambda}}^T(t) = -\frac{\partial H(t)}{\partial \boldsymbol{x}(t)}, \quad \boldsymbol{\lambda}(t_f) = \boldsymbol{0}$$

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{u}(t), \quad \boldsymbol{x}(0) = \boldsymbol{x}_0$$

- Now,

$$\frac{d}{dt}\left\{\frac{\partial H(t)}{\partial \boldsymbol{u}(t)}\right\} = 0$$

$$\frac{\partial H(t)}{\partial \boldsymbol{u}(t)} = \frac{\partial L(t)}{\partial \boldsymbol{u}(t)} + \boldsymbol{\lambda}^T(t)$$

- So we can write,

$$\begin{aligned}\frac{d}{dt}\left\{\frac{\partial H(t)}{\partial \boldsymbol{u}(t)}\right\} &= \frac{d}{dt}\left\{\frac{\partial L(t)}{\partial \boldsymbol{u}(t)}\right\} + \dot{\boldsymbol{\lambda}}^T(t) \\ &= \frac{d}{dt}\left\{\frac{\partial L(t)}{\partial \boldsymbol{u}(t)}\right\} - \frac{\partial L(t)}{\partial \boldsymbol{x}(t)} = 0\end{aligned}$$

$\Rightarrow$ Equations of Motion

- Consider,

$$\frac{dH(t)}{dt} = \frac{\partial L(t)}{\partial t} + \frac{\partial H(t)}{\partial \boldsymbol{x}(t)}\frac{\partial \boldsymbol{x}(t)}{\partial t} + \frac{\partial H(t)}{\partial \boldsymbol{u}(t)}\frac{\partial \boldsymbol{u}(t)}{\partial t} + \dot{\boldsymbol{\lambda}}^T(t)\,\boldsymbol{f}(t)$$

- At the optimal solution,

$$\frac{\partial H(t)}{\partial \boldsymbol{u}(t)} = 0$$

$$\frac{\partial H(t)}{\partial \boldsymbol{x}(t)} = -\dot{\boldsymbol{\lambda}}^T(t)$$

- If J is not an explicit function of t, then

$\partial L(t)/\partial t = 0 \quad \Rightarrow \quad dH(t)/dt = 0 \quad \Rightarrow \quad H(t) =$ constant

(mostly blank)