# Optimal Control
## Lecture 28: Indirect Solution Methods

Benoît Chachuat <benoit@mcmaster.ca>

McMaster University

Department of Chemical Engineering

Spring 2009



## Direct vs. Indirect Methods

### Direct Methods

Discretize the control problem, then apply NLP techniques to the resulting finite-dimensional optimization problem

- A priori knowledge of the solution structure not required
- **But,** approximate solution only (due to control parameterization)

### Indirect Methods (or PMP-based Methods or Variational Methods)

Seek a solution to the (closed system of) necessary conditions of optimality

- Discretization of the control profile not needed
- **But,** need to guess the optimal solution structure too

**Recommendation:** First approximate the solution with a direct approach, then refine this solution with an indirect approach!



## Regular, Terminal-Constrained Problems

$$\min_{(\mathbf{u},t_f)\in\mathcal{C}[t_0,T]^{n_u}\times\mathbb{R}} \int_{t_0}^{t_f} \ell(t,\mathbf{x}(t),\mathbf{u}(t))\,dt + \phi(t_f,\mathbf{x}(t_f))$$

$$\text{s.t. } \dot{\mathbf{x}}(t)=\mathbf{f}(t,\mathbf{x}(t),\mathbf{u}(t)); \quad \mathbf{x}(t_0)=\mathbf{x}_0$$

$$\psi_k(t_f,\mathbf{x}(t_f))=0, \quad k=1,\ldots,n_\psi$$

### Necessary Conditions for $(\mathbf{u}^*, t_f^*, \mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ to be Optimal

- Euler-Lagrange Equations ($\mathcal{H} \triangleq \ell + \boldsymbol{\lambda}^\mathsf{T}\mathbf{f}$):

$$\dot{\mathbf{x}} = \mathcal{H}_{\boldsymbol{\lambda}}, \qquad \dot{\boldsymbol{\lambda}} = -\mathcal{H}_{\mathbf{x}}, \qquad \mathbf{0} = \mathcal{H}_{\mathbf{u}}, \qquad t_0 \le t \le t_f$$

- Transversal Conditions:

$$[\mathbf{x}-\mathbf{x}_0]_{t_0} = \mathbf{0}, \qquad \left[\boldsymbol{\lambda} - \phi_{\mathbf{x}} + \boldsymbol{\nu}^\mathsf{T}\psi_{\mathbf{x}}\right]_{t_f} = \mathbf{0}$$

$$\left[\mathcal{H} + \phi_t + \boldsymbol{\nu}^\mathsf{T}\psi_t\right]_{t_f} = 0, \text{ if } t_f \text{ is free}$$

$$[\psi]_{t_f} = \mathbf{0}, \text{ and } \psi \text{ satisfy a regularity condition}$$



## Indirect Methods

General Idea: Split the NCOs into 2 subsets:

- Those NCOs that are enforced at each iteration
- The remaining NCOs that are modified, at each iteration, via successive linearization, until convergence
  - E.g., quasi-linearization, control vector iteration, indirect shooting

### Indirect Shooting Approach

1. Guess the values of $\boldsymbol{\lambda}^*(t_0)$, $\boldsymbol{\nu}^*$ and $t_f^*$
2. Iteratively update these estimates to meet the transversal conditions

$$\mathbf{F}(\boldsymbol{\lambda}_0^k, \boldsymbol{\nu}^k, t_f^k) \triangleq \begin{pmatrix} \boldsymbol{\lambda} + \phi_{\mathbf{x}} + {\boldsymbol{\nu}^k}^\mathsf{T}\psi_{\mathbf{x}} \\ \psi \\ \ell + \boldsymbol{\lambda}^\mathsf{T}\mathbf{f} + \phi_t + {\boldsymbol{\nu}^k}^\mathsf{T}\psi_t \end{pmatrix}_{t=t_f^k} \longrightarrow \mathbf{0}$$

subject to:

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

$$\dot{\boldsymbol{\lambda}}(t) = -\ell_{\mathbf{x}}(t,\mathbf{x}(t),\mathbf{u}(t)) - \boldsymbol{\lambda}(t)^\mathsf{T}\mathbf{f}_{\mathbf{x}}(t,\mathbf{x}(t),\mathbf{u}(t)); \quad \boldsymbol{\lambda}(t_0)=\boldsymbol{\lambda}_0^k$$

$$\mathbf{0} = \ell_{\mathbf{u}}(t,\mathbf{x}(t),\mathbf{u}(t)) + \boldsymbol{\lambda}(t)^\mathsf{T}\mathbf{f}_{\mathbf{u}}(t,\mathbf{x}(t),\mathbf{u}(t))$$

## Indirect Shooting Approach: Algorithm

**Initialization:**

- Choose initial estimates $\boldsymbol{\lambda}_0^0$, $\boldsymbol{\nu}^0$, $t_f^0$, termination tolerance $\epsilon > 0$; Set $k \leftarrow 0$

**Main Step:**

1. Calculate the defect $\mathbf{F}(\boldsymbol{\lambda}_0^k, \boldsymbol{\nu}^k, t_f^k)$; If $\|\mathbf{F}(\boldsymbol{\lambda}_0^k, \boldsymbol{\nu}^k, t_f^k)\| < \epsilon$, **Stop**
2. Calculate the defect gradient $\boldsymbol{\nabla}_{\boldsymbol{\lambda}_0}\mathbf{F}(\boldsymbol{\lambda}_0^k, \boldsymbol{\nu}^k, t_f^k)$, $\boldsymbol{\nabla}_{\boldsymbol{\nu}}\mathbf{F}(\boldsymbol{\lambda}_0^k, \boldsymbol{\nu}^k, t_f^k)$ and $\boldsymbol{\nabla}_{t_f}\mathbf{F}(\boldsymbol{\lambda}_0^k, \boldsymbol{\nu}^k, t_f^k)$
3. Calculate the search direction via solution of the linear system,
$$\begin{pmatrix} \boldsymbol{\nabla}_{\boldsymbol{\lambda}_0}\mathbf{F}(\boldsymbol{\lambda}_0^k, \boldsymbol{\nu}^k, t_f^k)^{\mathsf{T}} \\ \boldsymbol{\nabla}_{\boldsymbol{\nu}}\mathbf{F}(\boldsymbol{\lambda}_0^k, \boldsymbol{\nu}^k, t_f^k)^{\mathsf{T}} \\ \nabla_{t_f}\mathbf{F}(\boldsymbol{\lambda}_0^k, \boldsymbol{\nu}^k, t_f^k)^{\mathsf{T}} \end{pmatrix}^{\mathsf{T}} \begin{pmatrix} \mathbf{d}_{\boldsymbol{\lambda}}^k \\ \mathbf{d}_{\boldsymbol{\nu}}^k \\ d_{t_f}^k \end{pmatrix} = -\mathbf{F}(\boldsymbol{\lambda}_0^k, \boldsymbol{\nu}^k, t_f^k)$$
4. Update the estimates,
$$\boldsymbol{\lambda}_0^{k+1} \leftarrow \boldsymbol{\lambda}_0^k + \mathbf{d}_{\boldsymbol{\lambda}}^k, \quad \boldsymbol{\nu}^{k+1} \leftarrow \boldsymbol{\nu}^k + \mathbf{d}_{\boldsymbol{\nu}}^k, \quad t_f^{k+1} \leftarrow t_f^k + d_{t_f}^k$$
5. Increment $k \leftarrow k+1$, and return to step 1



## Indirect Shooting Approach: Application

**Class Exercise:** Consider the optimal control problem:

$$\text{minimize:} \quad \mathcal{J}(u) \triangleq \int_0^1 \tfrac{1}{2}u(t)^2 \, dt$$
$$\text{subject to:} \quad \dot{x}(t) = u(t)[1 - x(t)]; \quad x(0) = -1; \quad x(1) = 0$$

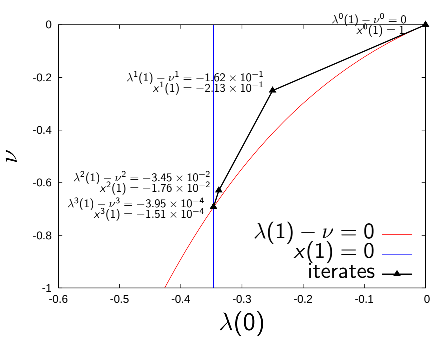
1. Formulate the defect $\mathbf{F}(\boldsymbol{\lambda}_0^k, \boldsymbol{\nu}^k, t_f^k)$ in the indirect shooting approach

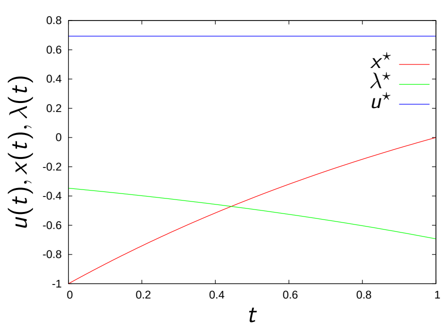


## Indirect Shooting Approach: Application

**Class Exercise:** Consider the optimal control problem:

$$\text{minimize:} \quad \mathcal{J}(u) \triangleq \int_0^1 \tfrac{1}{2}u(t)^2 \, dt$$
$$\text{subject to:} \quad \dot{x}(t) = u(t)[1 - x(t)]; \quad x(0) = -1; \quad x(1) = 0$$






## Indirect Shooting Approach: Remarks

- The methods applies readily in the case of a fixed terminal time, or in the absence of terminal constraints
  - Keep only the relevant transversal conditions!
- The defect gradient $\boldsymbol{\nabla}_{\boldsymbol{\lambda}_0}\mathbf{F}(\boldsymbol{\lambda}_0^k, \boldsymbol{\nu}^k, t_f^k)$, $\boldsymbol{\nabla}_{\boldsymbol{\nu}}\mathbf{F}(\boldsymbol{\lambda}_0^k, \boldsymbol{\nu}^k, t_f^k)$ and $\boldsymbol{\nabla}_{t_f}\mathbf{F}(\boldsymbol{\lambda}_0^k, \boldsymbol{\nu}^k, t_f^k)$ is needed to perform the Newton iteration
  - Use forward or reverse sensitivity analysis
  - Or, apply a quasi-Newton method with Broyden update scheme to estimate the gradient
- **Difficulty 1.** Find good initial estimates $(\boldsymbol{\lambda}_0^0, \boldsymbol{\nu}^0, t_f^0)$ of $(\boldsymbol{\lambda}_0^*, \boldsymbol{\nu}^*, t_f^*)$
  - Very high sensitivity towards the transversal conditions
  - Calculate good estimates via a direct discretization approach!
- **Difficulty 2.** Hard to apply in the presence of singular arcs and/or path constraints
  - Optimal arc sequence not known a priori
  - Optimality conditions at junction times give rise to multi-point boundary value problems (MPBVP)