OPTIMIZATION OF DIFFERENTIAL-ALGEBRAIC EQUATION SYSTEMS

L. T. Biegler
Chemical Engineering Department
Carnegie Mellon University
Pittsburgh, PA 15213
biegler@cmu.edu
http://dynopt.cheme.cmu.edu

May 26, 2000

DAE Optimization Outline

I Introduction

Process Examples

II Parametric Optimization

- Gradient Methods
  - Perturbation
  - Direct - Sensitivity Equations
  - Adjoint Equations

III Optimal Control Problems

- Optimality Conditions
- Model Algorithms
  - Control Vector Iteration
  - Parameterization

IV Simultaneous Solution Strategies

- Relation to "Black-Box" Methods
- Properties and Approach
- Roles of Finite Elements
- High Index DAEs
- Process Case Studies

INTRODUCTION

Task: Optimize a system described by a differential-algebraic system of equations.

Problem: $\min \phi(z(t_f))$

$$\begin{aligned} \text{s.t.} \quad & \dot{z} = f(z,u,p),\ z(0) = z_0 \\ & g(z,u,p,t) \le 0,\ t \in S \\ & h(z,u,p,t) = 0,\ t \in V \end{aligned}$$

State variables, z, and control variables, u, are now functions of time.

ϕ - objective functional at final time
g,h - path constraints over time domains S and V
t_f - final time, variable or fixed
p - "constant" variables or parameters (n_p)

Potentially infinite number of variables and constraints.

Examples in Process Engineering

- Optimizing Process Control

- Parameter Estimation (kinetic data)
- Optimal Startup, Shutdown, Time Scheduling
- Reactor Design and Optimization
- Analysis of Dynamic Systems

Batch Reactor Temperature Profiles

Consider a nonlinear batch reactor example (Ray, 1981) with temperature as the control variable. It is desired to maximize and intermediate product after a fixed reaction time. Here we consider the following reaction:

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C$$

The problem is nonlinear in the rate equations for the concentration of A. Letting c_1 and c_2 represent the concentration of A and B, respectively, the optimal control problem is:

$$\text{Max} \quad c_2(1.0)$$

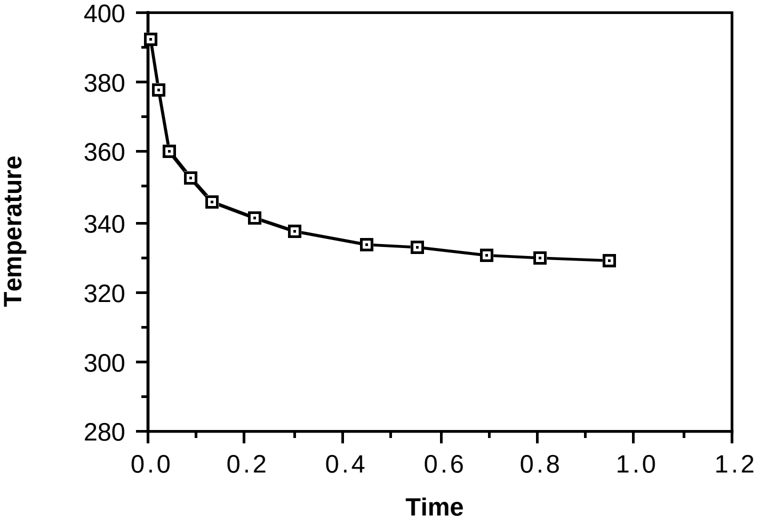
$$\text{s.t.} \quad \frac{dc_1}{dt} = -k_1(T)\, c_1^2$$

$$\frac{dc_2}{dt} = k_1(T)\, c_1^2 - k_2(T)\, c_2$$

$$k_i(T) = A_{i0} \exp^{[-E_i / RT]} \qquad i = 1,2$$

$$c_1(0) = 1.0 \ , \ c_2(0) = 0$$

$$298 \le T \le 398$$



Parameter Estimation Example

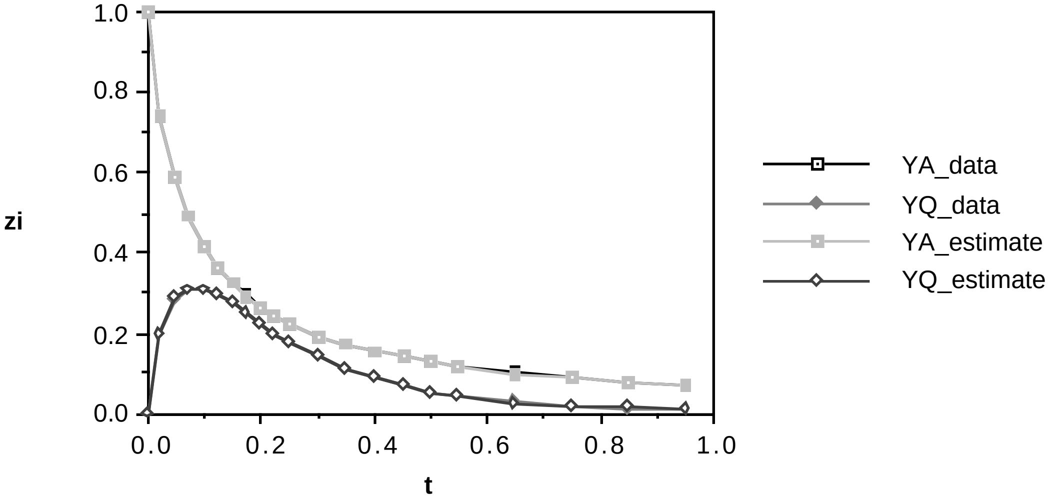
Catalytic Cracking of Gasoil (Tjoa, 1991)

$$A \xrightarrow{p_1} Q \qquad Q \xrightarrow{p_2} S \qquad A \xrightarrow{p_3} S$$

$$z_A' = -(p_1+p_3)\, z_A^2$$
$$z_Q' = -p_1 z_A^2 - p_2\, z_Q$$
$$z_A(0) = 1,\ z_Q(0) = 0$$

number of states and ODEs: 2
number of parameters: 3
no control profiles
constraints: $p^L \le p \le p^U$
Objective Function: Ordinary Least Squares

$$(p_1, p_2, p_3)^0 = (6, 4, 1)$$
$$(p_1, p_2, p_3)^* = (11.95, 7.99, 2.02)$$
$$(p_1, p_2, p_3)^{true} = (12, 8, 2)$$



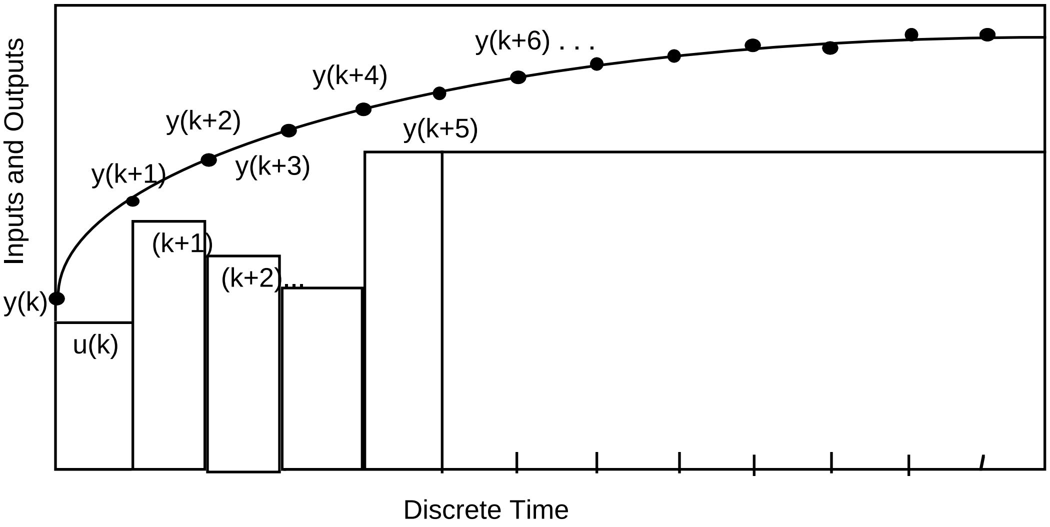
Nonlinear Model Predictive Control (NMPC)

Optimization of discretized control profile for a receding time horizon - solved on-line

$$\begin{array}{lcl} z' = f(z, u, t) & & z_{k+1} = F(z_k, u_k, t_{k+1}) \\ y = g(z) & \Longrightarrow & y_k = F(z_k, u_k, t_k) \\ z(0) = z_k & & \end{array}$$

number of states and ODEs: n_Z
no parameters (in control phase)
n_u control profiles
constraints: $u^L \leq u_k \leq u^U \quad y^L \leq y_k \leq y^U$
(+ stability constraints)

Objective Function: Linear Quadratic



Batch Process Optimization

Optimization of dynamic batch process operation resulting from reactor and distillation column DAE models

$$z' = f(z, y, u, p)$$
$$g(z, y, u, p) = 0$$

number of states and DAEs: $n_z + n_y$

parameters for equipment design (reactor, column)

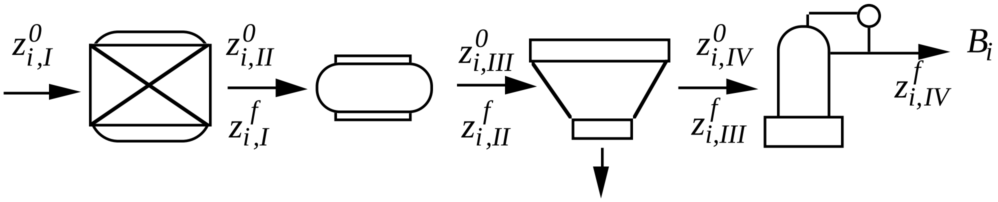
n_u control profiles for optimal operation

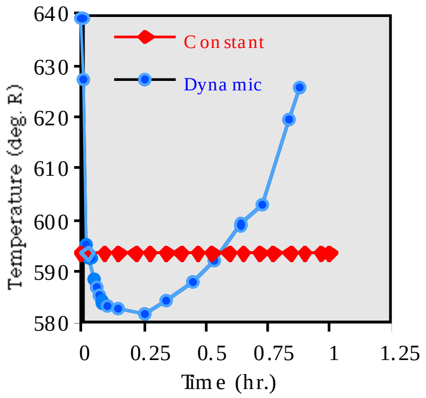
constraints: $u^L \leq u(t) \leq u^U$ $z^L \leq z(t) \leq z^U$

$y^L \leq y(t) \leq y^U$ $p^L \leq p \leq p^U$

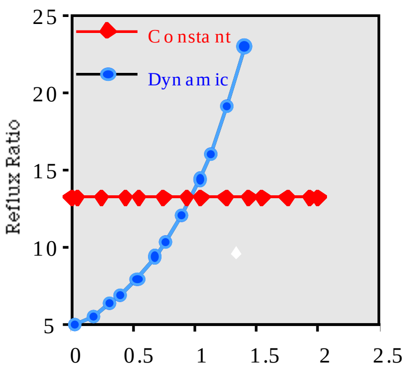
Objective Function: amortized economic function at end of cycle time t_f

$$\boxed{\begin{array}{l} A + B \rightarrow C \\ C + B \rightarrow P + E \\ P + C \rightarrow G \end{array}}$$





optimal reactor temperature policy



optimal column reflux ratio

Reactor Design Example

Plug Flow Reactor Optimization

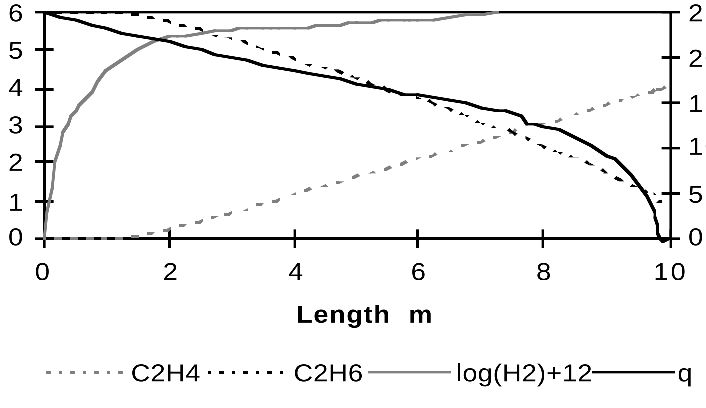
The cracking furnace is an important example in the olefin production industry, where various hydrocarbon feedstocks react. Consider a simplified model for ethane cracking (Chen et al., 1996). The objective is to find an optimal profile for the heat flux along the reactor in order to maximize the production of ethylene.

$$\begin{gathered} \text{Max} \quad F_{exit}^{C_2H_4} \\ \text{s.t. DAE} \\ T_{exit} \leq 1180\text{K} \end{gathered}$$

The reaction system includes six molecules, three free radicals, and seven reactions. The model also includes the heat balance and the pressure drop equation. This gives a total of eleven differential equations.

$$\begin{array}{ll} C_2H_6 \rightarrow 2CH_3\bullet & CH_3\bullet + C_2H_6 \rightarrow CH_4 + C_2H_5\bullet \\ H\bullet + C_2H_6 \rightarrow H_2 + C_2H_5\bullet & C_2H_5\bullet \rightarrow C_2H_4 + H\bullet \\ C_2H_5\bullet + C_2H_4 \rightarrow C_3H_6 + CH_3\bullet & 2C_2H_5\bullet \rightarrow C_4H_{10} \\ H\bullet + C_2H_4 \rightarrow C_2H_5\bullet & \end{array}$$

Concentration and Heat Addition Profile

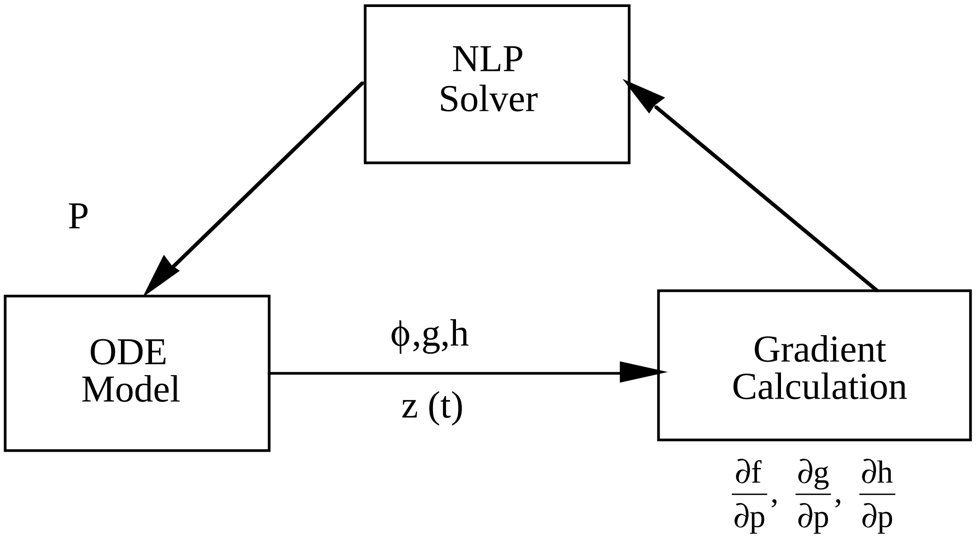


PARAMETER OPTIMIZATION

Consider a simpler problem without control profiles: e.g., equipment design with DAE models - reactors, absorbers, heat exchangers

$$\begin{aligned} \text{Min} \quad & \phi\,(z(t_f)) \\ z' = f(z,p),\ z\,(0) = z_0 & \\ & g\,(z(t_f)) \le 0, \\ & h\,(z(t_f)) = 0 \end{aligned}$$

By treating the ODE model as a "black-box" a sequential algorithm can be constructed that can be treated as a nonlinear program.



Task: How are gradients calculated for optimizer?

GRADIENT CALCULATION

- Perturbation
- Sensitivity Equations
- Adjoint Equations

Pertubation

Calculate approximate gradient by solving ODE model (n_p + 1) times

Let $\Psi = \phi$, g and h (at $t = t_f$)

$$\frac{\partial \psi}{\partial p_i} = \frac{1}{\Delta p_i}\{\Psi (p_i + \Delta p_i) - \Psi (p_i)\}$$

- Very simple to set up
- Leads to poor performance of optimizer and poor detection of optimum unless roundoff error ($O(1/\Delta p_i)$ and truncation error ($O(\Delta p_i)$) are small.
- Work is proportional to np (expensive)

Sensitivity Equations

From model:

$$\frac{\partial}{\partial p}\left\{\dot{z} = f(z,p),\ x(0) = z_o\right\}$$

$$\Rightarrow \frac{\partial \dot{z}}{\partial p} = \frac{d}{dt}\left(\frac{\partial z}{\partial p}\right) = \frac{\partial f}{\partial p} + \left(\frac{\partial z}{\partial p}\right)\frac{\partial f}{\partial z}$$

$$\left(\frac{\partial z}{\partial p}\right)\Big|_{t=0} = \frac{\partial z_o}{\partial p}$$

(n_z x n_p sensitivity equations)

- z and (∂z/∂p) can be integrated forward simultaneously.
- for implicit ODE solvers, (∂z/∂p) can be carried forward after converging on z
- linear sensitivity equations

For $\psi = \phi, g, h$

$$\frac{\partial \psi}{\partial p} = \left(\frac{\partial z}{\partial p}\right)_{t_f}\left(\frac{\partial \psi}{\partial z_{t_f}}\right)$$

Note: sensitivity equations are efficient for problems with many more constraints than parameters ($1 + n_g + n_h > n_p$)

Example: Sensitivity Equations

$$\dot{z}_1 = z_1^2 + z_2^2$$

$$\dot{z}_2 = z_1 z_2 + p_2 z_1$$

$$z1(0) = 5$$
$$z2(0) = p1$$

$$\left(\frac{\partial \dot{z}_1}{\partial p_1}\right) = 2z_1\left(\frac{\partial z_1}{\partial p_1}\right) + 2z_2\left(\frac{\partial z_2}{\partial p_1}\right)$$

$$\left(\frac{\partial \dot{z}}{\partial p_2}\right) = 2z_1\left(\frac{\partial z_1}{\partial p_2}\right) + 2z_2\left(\frac{\partial z_2}{\partial p_2}\right)$$

$$\left(\frac{\partial \dot{z}_2}{\partial p_1}\right) = \left(z_2 + p_2\right)\left(\frac{\partial z_1}{\partial p_1}\right) + z_1\left(\frac{\partial z_2}{\partial p_1}\right)$$

$$\left(\frac{\partial \dot{z}_2}{\partial p_2}\right) = z_1 + \left(z_2 + p_2\right)\left(\frac{\partial z_1}{\partial p_2}\right) + z_1\left(\frac{\partial z_2}{\partial p_2}\right)$$

$$\left(\frac{\partial z_1}{\partial p_1}\right)_0 = 0 \quad \left(\frac{\partial z_2}{\partial p_1}\right)_0 = 1$$

$$\left(\frac{\partial z_1}{\partial p_2}\right) = 0 \quad \left(\frac{\partial z_2}{\partial p_2}\right) = 0$$

Adjoint Equations

Adjoint or Dual approach to sensitivity

Adjoin model to objective function <u>or</u> constraint
($\Psi = \phi$,g or h)

- $\psi(t_f) - \int_o^{t_f} \lambda^T(\dot{z} - f(z,p))\, dt = \psi$

  (λ(t)) serve as multipliers on ODE's)

Now, integrate by parts

- $\psi(t_f) - \lambda^T z \Big|_o^{t_f} + \int_o^{t_f} \left(\dot{\lambda}^T z + \lambda^T f(z,p)\right) dt = \psi$

Reduced gradient analogy -

Find $\frac{d\psi}{dp}$ subject to feasibility of ODE's

$$d\psi = \left(\frac{\partial \psi}{\partial z(t_f)}^T - \lambda_{t_f}^T\right) \delta z\,(t_f) + \lambda_o^T \delta z_o$$

$$+\int_o^{t_f} \left\{\dot{\lambda}^T + \lambda^T \left(\frac{\partial f}{\partial z}\right)^T)\delta z\right\} + \left\{\lambda^T \frac{\partial f}{\partial p}\right\} dp\, dt$$

Now, set all terms <u>not</u> in dp to zero.

Adjoint System

δz_0 (fixed) $\Rightarrow$ λ_0 arbitrary
δz_0 (parameter) $\Rightarrow$ $\lambda_0^T\, dz_0/dp_j$ (see example)

δz_{t_f}: $\dfrac{\partial \psi}{\partial z_{t_f}} = \lambda(t_f)$ final condition

$\delta z(t)$: $\dot{\lambda} = -\dfrac{\partial f}{\partial z}\lambda$ adjoint equations

$\delta z(t)$: $\dot{\lambda} = -\dfrac{\partial f}{\partial z}\lambda$, $I(0) = 0$

$$\Rightarrow \qquad \frac{d\psi}{dp} = I(t_f) + \lambda(0)\, \partial z_0 / \partial p$$

- Integrate model equations forward
- Integrate adjoint equations backward and evaluate integral, $I(t_f)$

Notes:

- $n_z (n_g + n_h + 1)$ adjoint equations must be solved backward (one for each objective and constraint function)

- for implicit ODE solvers, profiles (and even matrices) can be stored and carried backward after solving forward for z (Sargent & Sullivan, 1977)

- this approach is more efficient on problems where: $n_p > 1 + n_g + n_h$

Example: Adjoint Equations

$$\dot{z}_1 = z_1^2 + z_2^2$$

$$\dot{z}_2 = z_1 z_2 + p_2 z_1$$

$$z_1(0) = 5$$

$$z_2(0) = p_1$$

Adjoints: Form $\lambda^T f = \lambda_1(z_1^2 + z_2^2) + \lambda_2(z_1 z_2 + p_2 z_1)$

$$\dot{\lambda}_1 = -2z_1\,\lambda_1 - (z_2 + p_2)\,\lambda_2$$

$$\dot{\lambda}_2 = -2z_2\,\lambda_1 - z_1\,\lambda_2$$

$$\lambda_1(t_f) = \partial\psi/\partial z_1$$

$$\lambda_2(t_f) = \partial\psi/\partial z_2$$

$$\frac{d\psi}{dp_1} = \lambda_2(0) + \int_o^{t_f} (0)\,dt \qquad \frac{d\psi}{dp_2} = \int_0^{t_f} \lambda_2 z_1\,dt$$

Parameter Estimation/Optimization

Example: Catalytic Cracking of Gasoil (Tjoa, 1991)

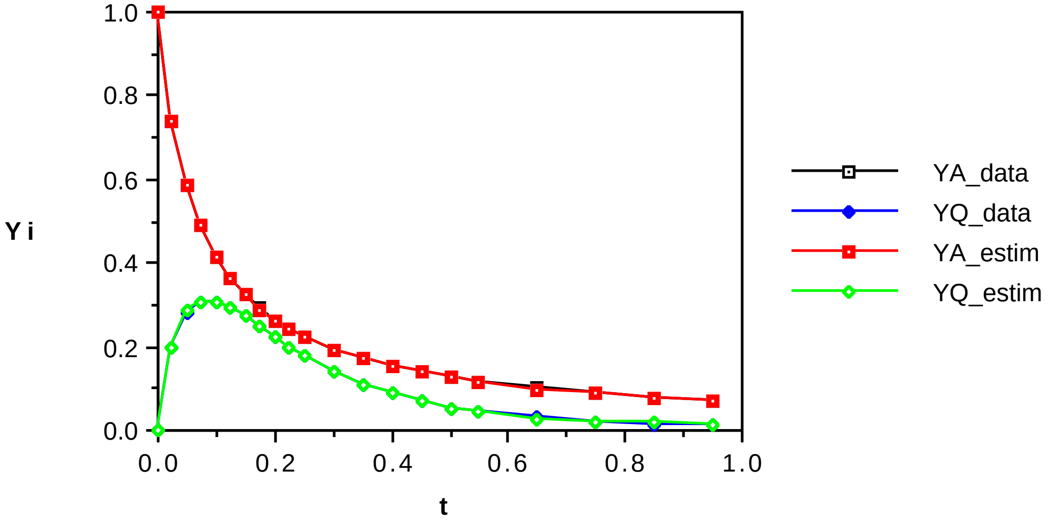
$$A \xrightarrow{k_1} Q \qquad Q \xrightarrow{k_2} S \qquad A \xrightarrow{k_3} S$$

$$y_A' = -(k_1+k_3)\, y_A^2$$
$$y_Q' = -k_1 y_A^2 - k_2\, y_Q$$
$$y_A(0) = 1,\; y_Q(0) = 0$$

Apply Trust Region method (GREG) using sensitivity equations (DDASAC):

$$(k_1, k_2, k_3)^0 = (6, 4, 1)$$
$$(k_1, k_2, k_3)^* = (11.95, 7.99, 2.02)$$
$$(k_1, k_2, k_3)^{true} = (12, 8, 2)$$

Converges in 5 iterations (11 function calls)



Parameter Optimization

Example: Hot Spot Reactor (Finlayson)

$$\min_{T_P, T_R, L, T_s, q(t), \overline{T}(t)} \quad \Phi = L - \int_0^L \left(\overline{T}(t) - T_s / T_R \right) dt$$

s.t.

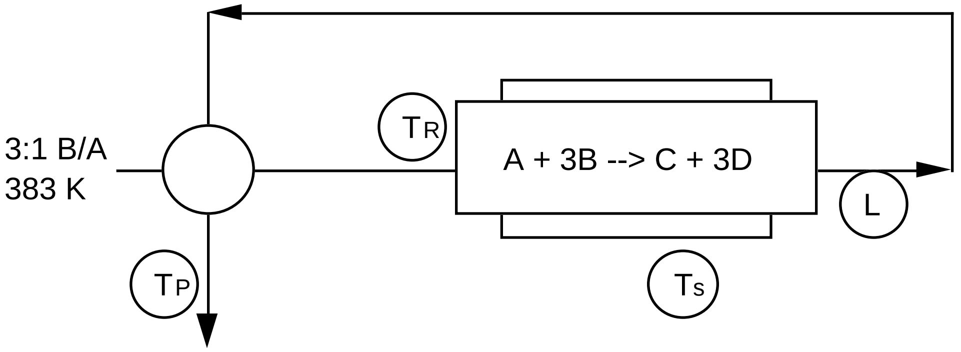
$$\frac{dq(t)}{dt} = 0.3(1 - q(t))\exp[20 - 20 / \overline{T}(t)] \qquad q(0) = 0$$

$$\frac{d\overline{T}(t)}{dt} = -1.5(\overline{T}(t) - T_s / T_R) + 2/3 \frac{dq(t)}{dt} \qquad \overline{T}(0) = 1$$

$$t \in [0, L]$$

$$C_{ph} = \Delta H_{feed}(T_R, 110^oC) - \Delta H_{product}(T_P, T(L)) = 0$$

$$T_P \geq 120^oC, \quad T(L) \geq T_R + 10^oC$$



where $\overline{T} = T(t)/T_R$

T_P = specified product temperature

T_R = reactor inlet, reference temperature

L = reactor length

T_S = steam sink temperature

$q(t)$ = reactor conversion profile

$\overline{T}(t)$ = normalized reactor temperature profile

Cases considered:

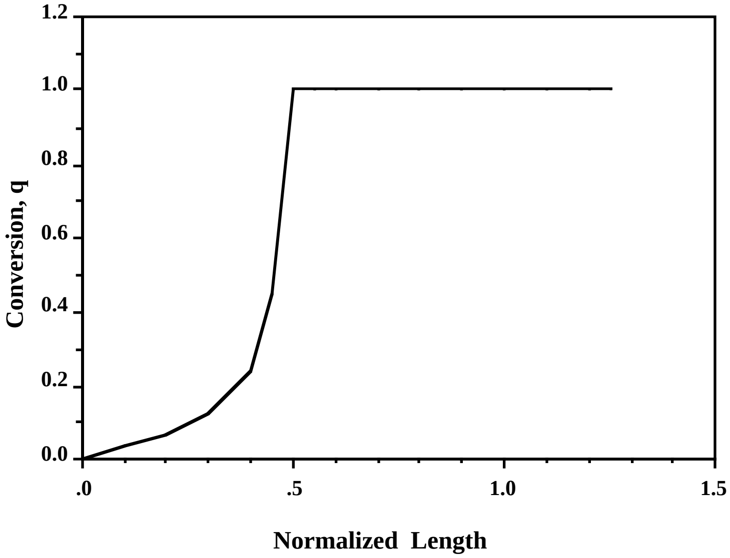
- Hot Spot - no state variable constraints
- Hot Spot with $\overline{T}(t) \leq 1.45$

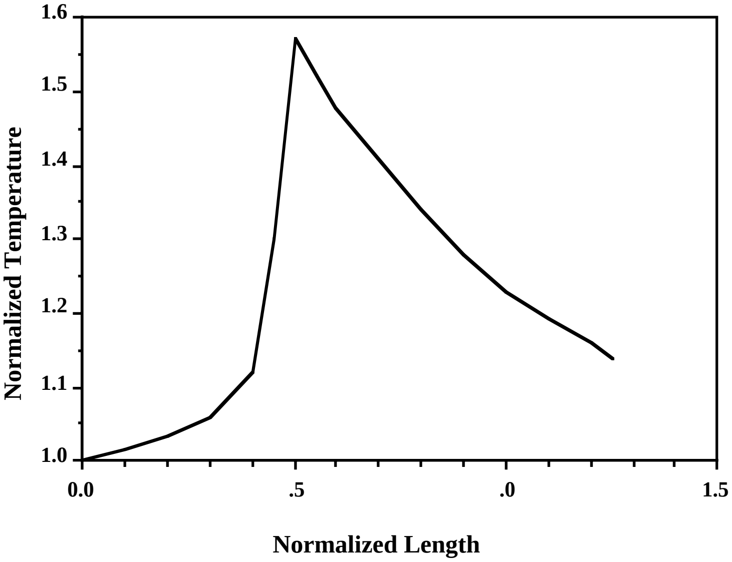
Unconstrained Case

Method: SQP (perturbation derivatives)

	L(norm)	T_R(K)	T_S(K)	T_P(K)
Initial:	1.0	462.23	425.26	250
Optimal:	1.25	500	470.1	188.4

13 SQP iterations / 2.67 CPU min. (µVax II)





Constrained Temperature Case: could not be solved with sequential method

'TRICKS' TO GENERALIZE CLASSES OF PROBLEMS

1. Variable Final Time (Miele, 1980)

   Define $t = p_{n+1}\,\tau$

   $0 \leq \tau \leq 1$

   $p_{n+1} = tf$

   Let $\dot{z} = \dfrac{1}{p_{n+1}}\dfrac{dz}{d\tau} = f(z,p)$

   $\Rightarrow\ dz/d\tau = (p_{n+1})\, f(z,p)$

2. Converting Path Constraints to Final Time

   Define measure of infeasibility as a new variable, z n_Z+1 (Sargent & Sullivan, 1977):

   $$z_{n_{z+1}}(tf) = \sum_j \int_s \max(0, g_j)^2 dt + \sum \int_v h_j(t)^2 dt$$

   or

   $$\dot{z}_{nz+1} = \sum_j \max(0, g_j)^2 + \sum h_j(t)^2$$

   (over respective time domains)

   and

   $$\dot{z}_{nz+1}(t_f) \leq \varepsilon$$

   - Simple solution, but can cause numerical problems due to degeneracy

Profile Optimization (Optimal Control)

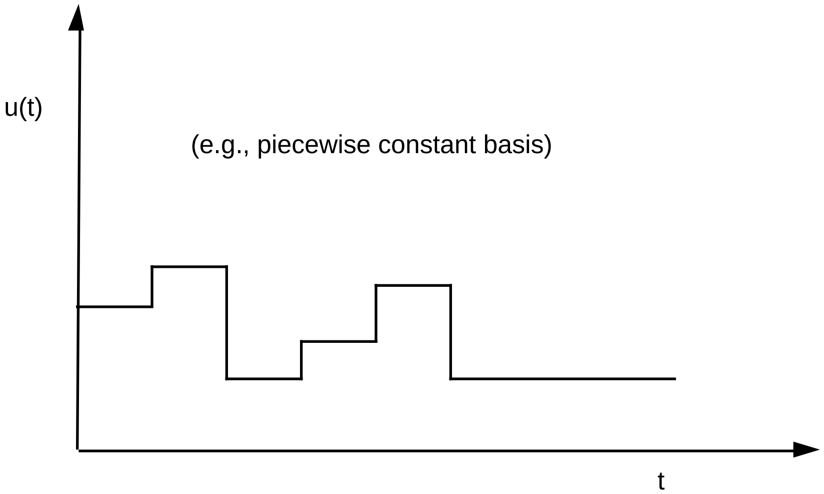
Examples:

- Optimal Feed Strategy (Schedule) in Batch Reactor
- Optimal Startup and Shutdown Policy
- Optimal Control of Transients and Upsets

Approaches to Solution:

1. Represent profile as parametric approximation
(piecewise constant, linear, polynomial, etc.)

- Apply NLP to discretization as with parametric optimization
- Obtain gradients through adjoints (Sargent and Sullivan, Goh and Teo) or sensitivity equations (Morshedi, Vassiliadis)



$u(t) = \sum_i \sum_j u_{ij}\, \varphi_{ij}(t)$ (parameterization w/ bases $\varphi_{ij}(t)$)

2. Optimize continuous profiles directly

Derive optimality conditions for infinite dimensional problem

Generalize NLP algorithms to infinite dimensional case

Optimality Conditions (Bound constraints on u(t))

$$
\begin{aligned}
&\text{Min} \quad \phi\,(z(t_f)) \\
&\text{s.t.} \\
&\quad z' = f(z,u),\ z\,(0) = z_0 \\
&\quad g\,(z(t_f)) \le 0 \\
&\quad h\,(z(t_f)) = 0 \\
&\quad a \le u(t) \le b
\end{aligned}
$$

Form Lagrange function - adjoin objective function and constraints:

$$\phi = \phi(z(t_f)) + g(z(t_f))\,\mu + h(z(t_f))\,\nu -$$

$$\int_0^{t_f} \lambda^T(\dot{z} - f(z,u,t)) + \alpha_a(a - u(t)) + \alpha_b(u(t) - b)dt$$

Integrate by parts:

$$\phi = \phi(z(t_f)) + g(z(t_f))\,\mu + h(z(t_f))\,\nu -$$

$$\lambda^T(t_0)z(t_0) - \lambda^T(t_f)z(t_f) + \int_0^{t_f} \dot{\lambda}^T z + \lambda^T f(z,u,t) + \alpha_a(a - u(t)) + \alpha_b(u(t) - b)dt$$

Variational Conditions from Lagrange Function:

$$\delta\phi = \left\{\frac{\partial\phi}{\partial z} + \frac{\partial g}{\partial z}\mu + \frac{\partial h}{\partial z}\gamma - \lambda(t_f)\right\}^T \delta z(t_f)$$

$$+ \lambda(0)^T \delta z(0) + \int_0^{t_f} \left\{\left(\dot{\lambda} + \frac{\partial f}{\partial z}\lambda\right)^T \delta z(t) + \left(\frac{\partial f}{\partial u}\lambda + \alpha_b - \alpha_a\right)^T \delta u(t)\right\} dt$$

At optimum, $\delta\phi \geq 0$. Since u is the control variable, let all other terms vanish.

$\Rightarrow$ $\delta z(t_f)$: $\lambda(t_f) = \left\{\frac{\partial\phi}{\partial z} + \frac{\partial g}{\partial z}\mu + \frac{\partial h}{\partial z}\gamma\right\}_{t=t_f}$

$\delta(t_f)$: $\lambda(0) = 0$ (if $z(t_f)$ is not specified)

$\delta z(t)$: $\dot{\lambda} = -\frac{\partial f}{\partial z}\lambda$

Define Hamiltonian, $H = \lambda^T f(z,u)$

For u not at bound: $\frac{\partial f}{\partial u}\lambda = \frac{\partial H}{\partial u} = 0$

For u at bounds: $\frac{\partial H}{\partial u} = \alpha_a - \alpha_b$

$\alpha_a^T(a - u(t))$

$\alpha_b^T(u(t) - b)$

$u_a \leq u(t) \leq u_b$

$\alpha_a \geq 0, \alpha_b \geq 0$

Upper bound, $u(t) = b$, $\frac{\partial H}{\partial u} = -\alpha_b \leq 0$

Lower bound, $u(t) = a$, $\frac{\partial H}{\partial u} = \alpha_a \geq 0$

Example: Travel a fixed distance (rest-to-rest) in minimum time.

Min t_f	Min x_3
s.t.	s.t.
$\ddot{x} = u$	$\dot{x}_1 = x_2$
$\dot{x}(0) = 0$	$\dot{x}_2 = u$
$\dot{x}(t_f) = 0$	$\dot{x}_3 = 1$
$x(0) = 0$	$x_1(0) = 0,\ x_1(t_f) = L$
$x(t_f) = L$	$x_2(0) = 0,\ x_2(t_f) = 0$
$a \le u \le b$	$a \le u \le b$

Hamiltonian: $H = \lambda^T f = \lambda_1 x_2 + \lambda_2 u + \lambda_3$

Adjoints:

$\dot{\lambda}_1 = 0$ $\qquad$ $\lambda_1(0), \lambda_2(0)$ unspecified

$\dot{\lambda}_2 = -\lambda_1$

$\dot{\lambda}_3 = 0$ $\qquad$ $\lambda_3(t_f) = 1$

$$\frac{\partial H}{\partial u} = \lambda_2$$

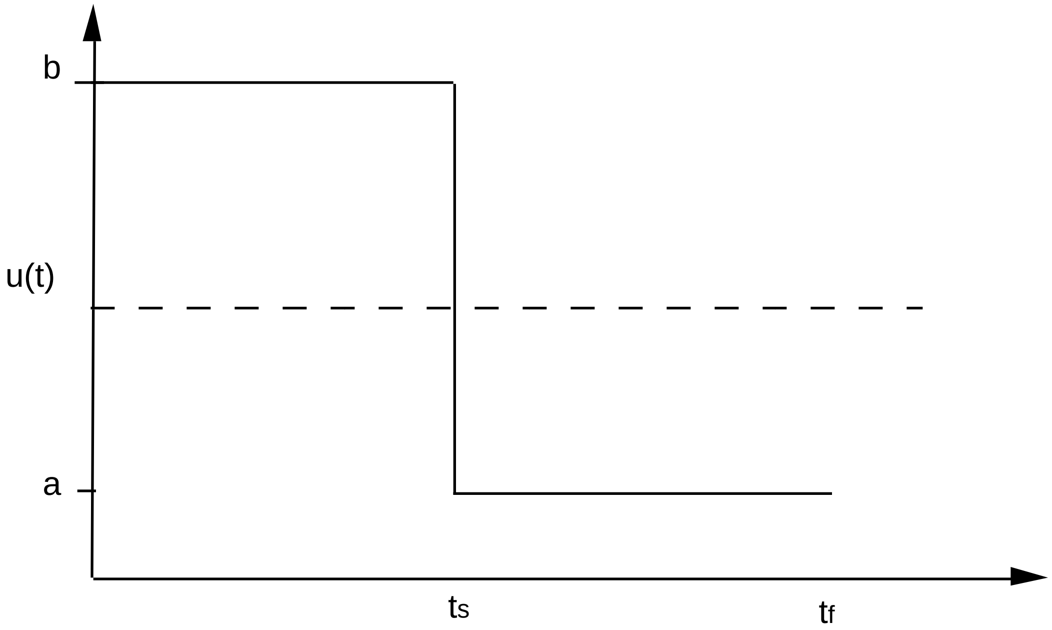
$$\Rightarrow \quad \lambda_1 = c_1 \qquad \lambda_2 = c_1(t_f - t) + c_2$$

Only condition that makes sense is:

$$c_1 < 0,\ c_2 > 0 \qquad \text{or for: } t_s = t_f + c_2/c_1$$

$$\frac{\partial H}{\partial u} = \lambda_2 \begin{cases} < 0 & t < t_s \Rightarrow u = b \\ > 0 & t > t_s \Rightarrow u = a \end{cases}$$

Optimal Profile



From state equations:

$$x_1(t) = \begin{cases} 1/2\ bt^2, t < t_s \\ 1/2\left(bt_s^2 - a\left(t_s - t_f\right)^2\right),\ t \geq t_s \end{cases}$$

$$x_2\ (t) = \begin{cases} bt,\ \ t < t_s \\ bt_s + a\left(t - t_s\right),\ \ t \geq t_s \end{cases}$$

Apply boundary conditions at $t = t_f$:

$$x_1(t_f) = 1/2\ (bt_s^2 - a\ (t_s - t_f)^2) = L$$
$$x_2(t_f) = bt_s + a\ (t_f - t_s) = 0$$

$$\Rightarrow \qquad t_s = \left[\frac{2L}{b(1 - b/a)}\right]^{1/2} \qquad t_f = (1 - b/a)\left[\frac{2L}{b(1 - b/a)}\right]^{1/2}$$

Note:

- Problem is linear in u.
- Frequently these problems have "bang-bang" character.

General Algorithm for Optimal Control

(Control Vector Iteration, Bryson and Ho, 1969)

Given:

$$\min \quad \phi(t_f)$$

s.t.

$$z' = f(z,u),\ z(0) = z_0$$
$$g(t_f) \le 0$$
$$h(t_f) = 0$$
$$a \le u(t) \le b$$

0. Initialize u(t) at some <u>feasible</u> point.

1. At iteration i with $u^i(t)$, solve $z' = f(z,u^i)$ forward in time.

2. Determine multipliers, μ and ν at final time (linear equations), for $g(t_f)$ and $h(t_f)$. Set:

$$\lambda(t_f) = \left\{ \frac{\partial \phi}{\partial z} + \frac{\partial g}{\partial z}\mu + \frac{\partial h}{\partial z}\gamma \right\}_{t=t_f}$$

3. Solve adjoint equations backward in time:

$$\dot{\lambda} = \frac{-\partial H}{\partial z} = -\frac{\partial f}{\partial z}\lambda$$

4. Evaluate $\frac{\partial H}{\partial u}$ over $t \in [0, t_f]$

Determine if correct sign of $\frac{\partial H}{\partial u}$ for u(t) is at bounds.

If $\|\partial H / \partial u\| \le \varepsilon$ for u(t) <u>not</u> at bounds, stop.

Control Vector Iteration (continued)

5. Update u(t) at each point in time.

If $u_j(t)$ at bound with correct sign of $\frac{\partial H}{\partial u}$, set $u_j^{i+1}(t) = u_j^i(t)$.

For remaining u(t) between bounds, take step:

$$u^{i+1}(t) = u_j^i(t) - \alpha\, M(t) \frac{\partial H}{\partial u}(t)$$

where α, stepsize along search direction,
M(t), matrix chosen depending on method:

- M = I, steepest descent
- M, based on Newton or quasi-Newton methods
- M, based on conjugate gradient steps

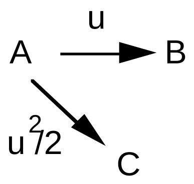
6. Set i = i+1, go to step 1.

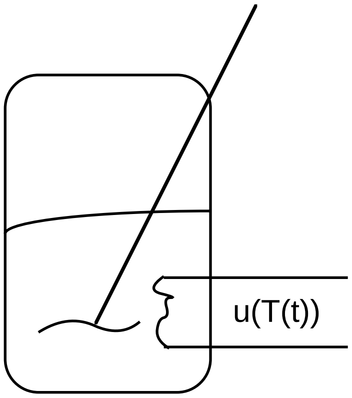
Notes:

- Control Vector Iteration requires an outer iteration loop to satisfy final time constraints

- Inequalities that include state variables are very difficult to satisfy directly (penalty functions are usually tried)

Example: Batch reactor - temperature profile





Maximize yield of B after one hour's operation by manipulating a transformed temperature, u(t).

$\Rightarrow$ Minimize $-z_B(1.0)$

s.t.

$$z'_A = -(u+u^2/2)\, z_A$$
$$z'_B = u\, z_A$$
$$z_A(0) = 1$$
$$z_B(0) = 0$$
$$0 \le u(t) \le 5$$

Adjoint Equations:

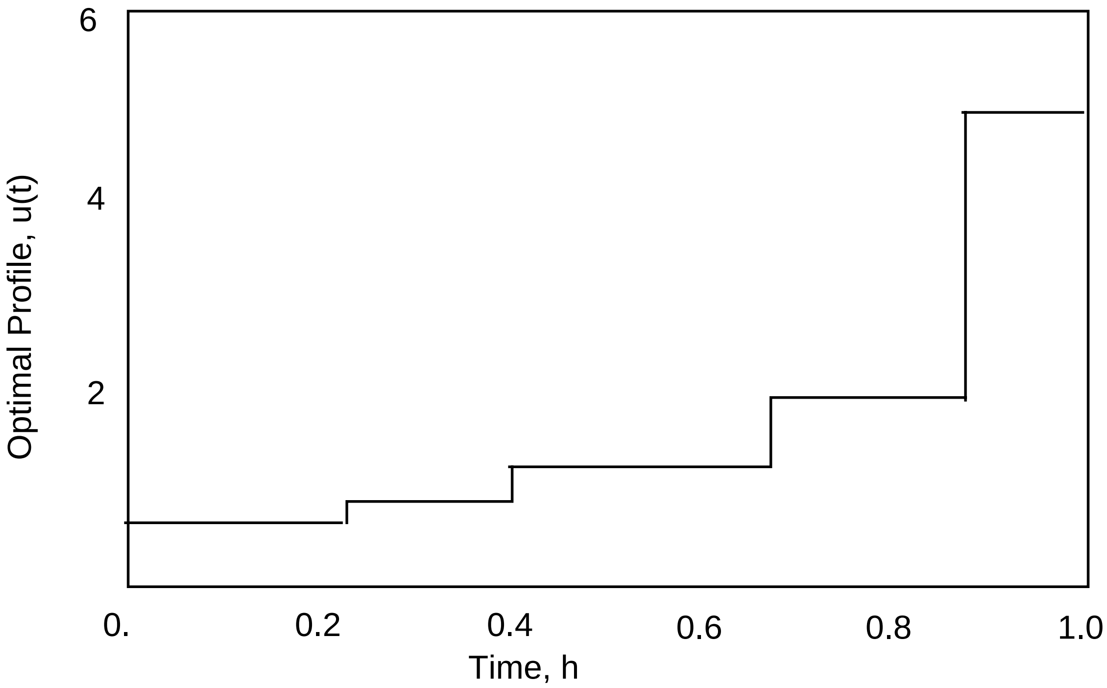
$$H = -\lambda_A(u+u^2/2)\, z_A + \lambda_B\, u\, z_A$$
$$\partial H/\partial u = \lambda_A\, (1+u)\, z_A + \lambda_B\, z_A$$
$$\lambda'_A = \lambda_A(u+u^2/2) + \lambda_A\, u,\ \lambda_A(1.0) = 0$$
$$\lambda'_B = 0,\ \lambda_B(1.0) = -1$$

Cases Considered

1. NLP Approach - piecewise constant and linear profiles.

2. Control Vector Iteration

Results of Optimal Temperature Program
Batch Reactor

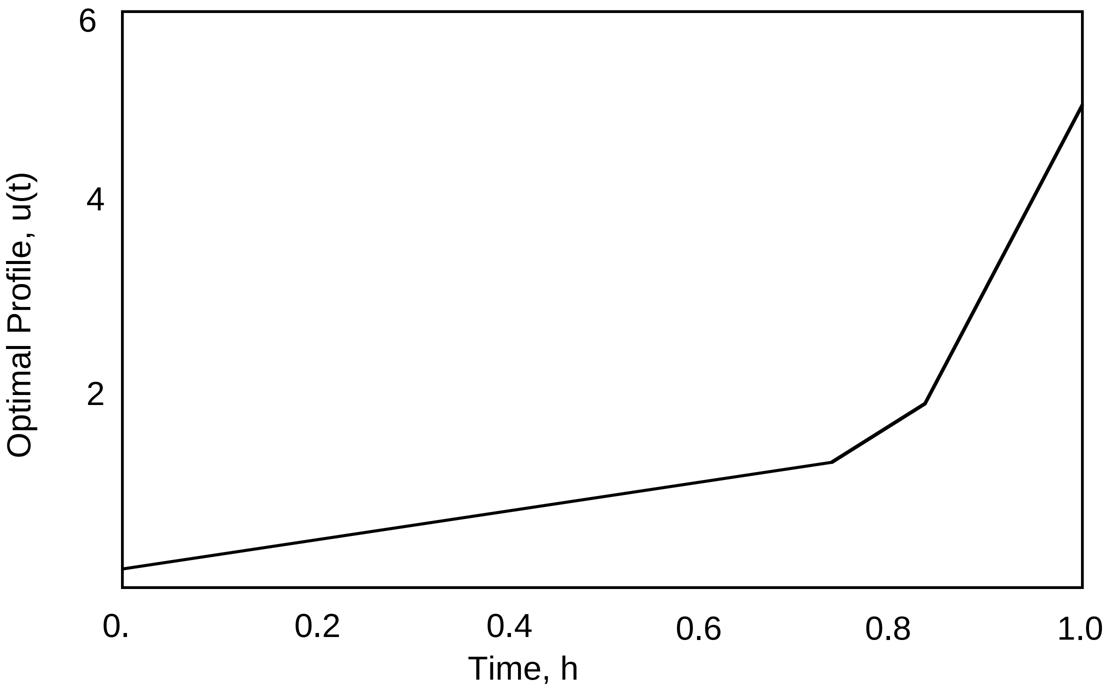


Results

Piecewise Constant Approximation with Variable Time Elements

Optimum B/A: 0.57105

Results of Optimal Temperature Program Batch Reactor

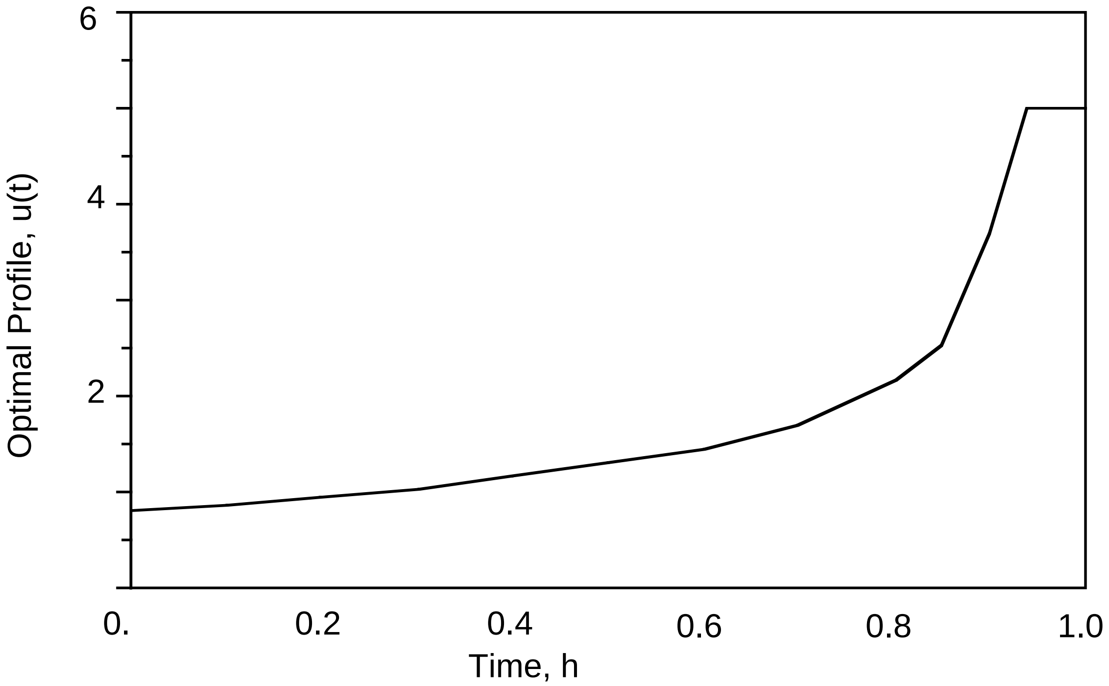


Results:

Piecewise Linear Approximation with Variable Time Elements
Optimum B/A: 0.5726

Equivalent # of ODE solutions: 32

Results of Optimal Temperature Program
Batch Reactor



Results:

Control Vector Iteration with Conjugate Gradients

Optimum (B/A): 0.5732
Equivalent # of ODE solutions: 58

Summary

1. The sequential NLP approach is an extension of parameter optimization. Derivatives obtained through adjoint or sensitivity equations.

2. Variational approach is direct extension of finite-dimensional optimization.

3. Both methods require repeated solution of ODE model and adjoint equations.

Difficulties with Sequential Optimization

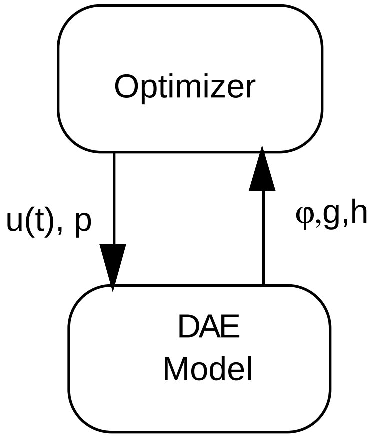
1. Path constraints are very difficult to handle:

    - for NLP approach, may have numerical problems and inexact profiles
    - for Variational Approach, CVI algorithm will not work for path constraints.

2. End conditions (g(tf) or h(tf)) lead to more difficult and larger otpimization problems for CVI. They are not difficult to handle for the sequential approach

3. Sequential optimization is not recommended for unstable systems. State variables blow up at intermediate iterations for control variables and parameters.

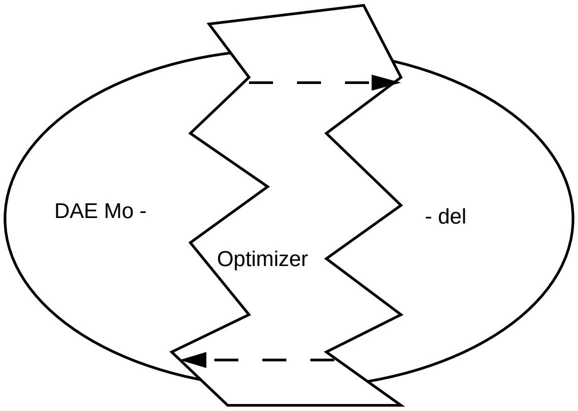
Approaches to Dynamic Optimization



- repeated solution of DAE model
- difficult to specify state variable constraints in model solver
- DAE solver can fail at intermediate points

Sequential Strategies
(Feasible Path)

Simultaneous Approach



Apply simultaneous solution and optimization algorithm (e.g. SQP, rSQP, MINOS)

- Model is solved only once, at optimum solution

- Decomposition can accommodate unstable systems

- State variable constraints are incorporated directly into optimization problem.

Discretization of Differential Equations - Orthogonal Collocation

For

$$\dot{z} = f(z,u,p), \quad z(t) = z_o$$

Approximate z and u by Lagrange interpolation polynomials (order N+1 and N, respectively)

$$z_{N+1}(t) = \sum_{i=0}^{N} z_i l_i(t) \qquad l_i(t) = \prod_{j=0,i}^{N} \frac{(t - t_j)}{(t_i - t_j)}$$

Note $z_{N+1}(t_i) = z_i$ (and $z_{N+1}(t_0) = z_0$)

Similarly, let:

$$u_N(t) = \sum_{i=1}^{N} u_i l_i(t) \qquad l_i(t) = \prod_{j=1,i}^{N} \frac{(t - t_j)}{(t_i - t_j)}$$

Now, at t_i chosen as roots of N^{th} degree Legendre polynomial (shifted to $t \in [0, t_f]$), apply orthogonal collocation:

$$r(t_i) = \sum z_j l_j'(t_i) - f(z_i, u_i\ p) = 0 \qquad i = 1,N$$

Collocation Example

Consider Lagrange polynomials:

$$z_{N+1}(t) = \sum_{i=0}^{N} z_i l_i(t) \qquad l_i(t) = \prod_{j=0,i}^{N} \frac{(t - t_j)}{(t_i - t_j)}$$

For two point collocation:

$$t_0 = 0, \quad t_1 = 0.21132, \quad t_2 = 0.78868$$

$l_0(t) = (t^2 - t + 1)/6$ $\qquad$ $l_0'(t) = t/3 - 1/6$
$l_1(t) = -8.195t^2 + 6.4483t$ $\qquad$ $l_0'(t) = 6.4483 - 16.39\ t$
$l_2(t) = 2.19625t^2 - 0.4641t$ $\qquad$ $l_0'(t) = 4.392t - 0.46412$

Solve $z' = z^2 - 3z + 2,\ z(0) = 0$

$$z_0 l_0'(t_1) + z_1 l_1'(t_1) + z_2 l_2'(t_1) = z_1^2 - 3z_1 + 2$$
$$(2.9857\ z_1 + 0.46412\ z_2 = z_1^2 - 3z_1 + 2)$$

$$z_0 l_0'(t_2) + z_1 l_1'(t_2) + z_2 l_2'(t_2) = z_2^2 - 3z_2 + 2$$
$$(-6.478\ z_1 + 3\ z_2 = z_2^2 - 3z_2 + 2)$$

$$z_0 = 0, \quad z_1 = 0.291\ (0.319), \quad z_2 = 0.7384\ (0.706)$$

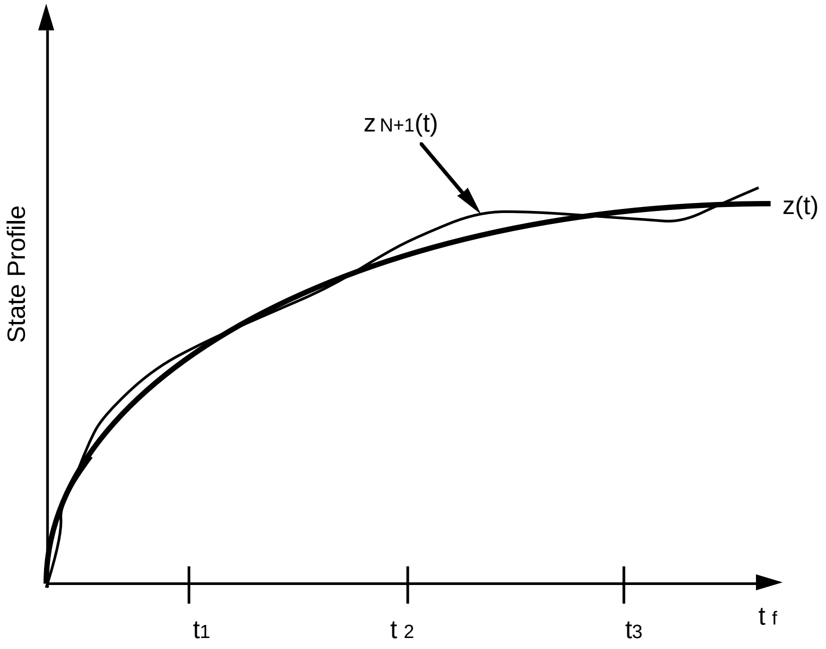
$$z(t) = 1.5337\ t - 0.76303\ t^2$$

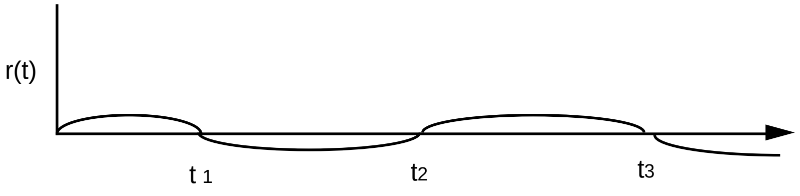
Convert Optimal Control Problem

$$\begin{aligned} &\text{Min } \phi(z(t_f)) \\ &\text{s.t.} \quad z' = f(z, u, p) \\ &\qquad g(z(t), u(t), p) \le 0 \\ &\qquad h(z(t), u(t), p) = 0 \end{aligned}$$

to Nonlinear Program

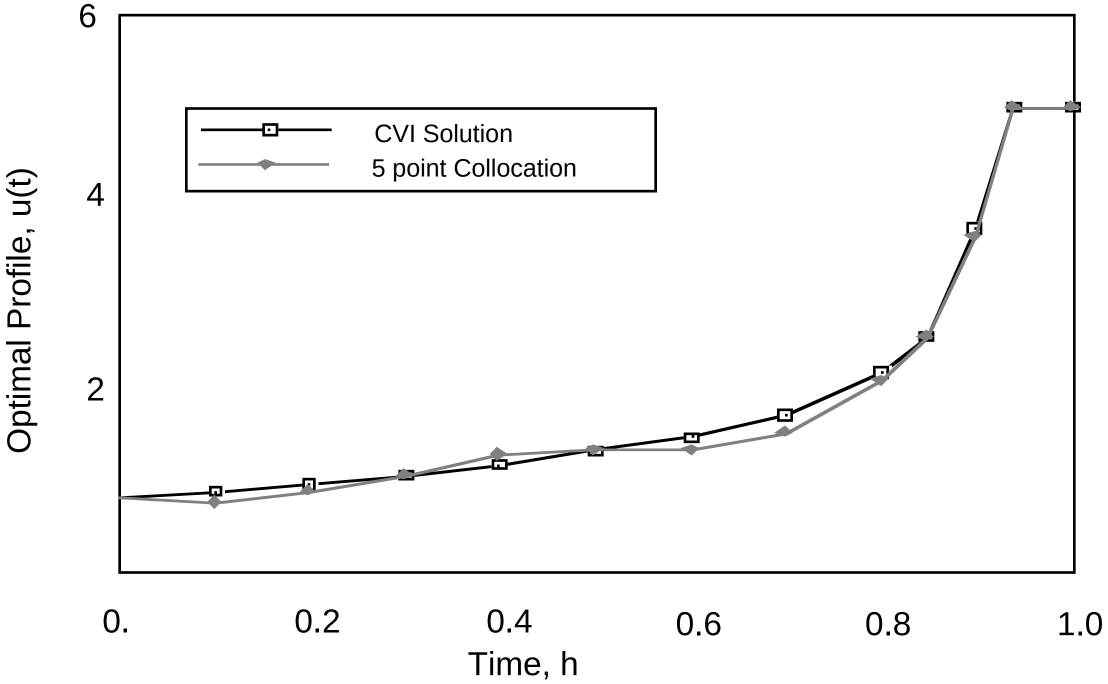
$$\begin{aligned} &\text{Min } \phi(z_{N+1}(t_f)) \\ &\text{s.t.} \quad \sum z_j \, l_j'(t_i) - f(z_i, u_i\ p) = 0 \qquad i = 1,\ldots N \\ &\qquad g(z_i, u_i, p) \le 0 \\ &\qquad h(z_i, u_i, p) = 0 \end{aligned}$$





How accurate is approximation?

Results of Optimal Temperature Program
Batch Reactor (Revisited)



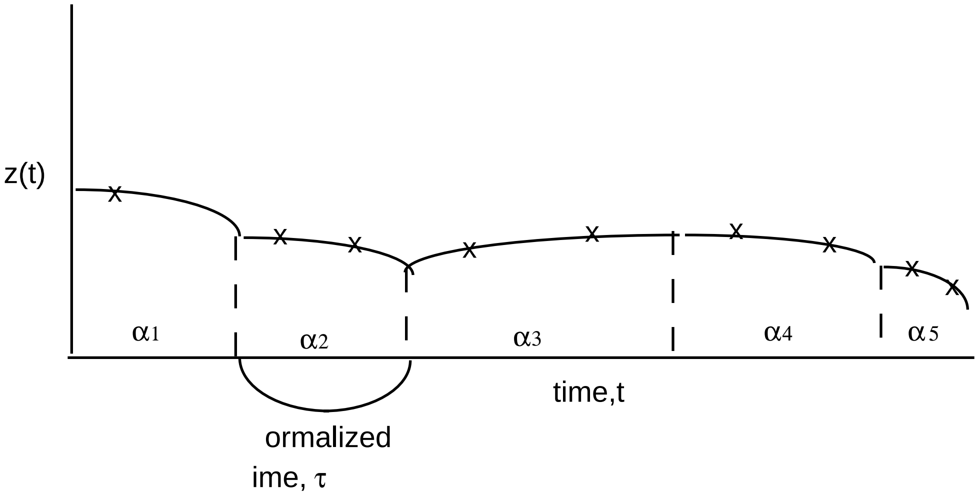
<u>Results</u> - NLP with Orthogonal Collocation

Optimum B/A - <u>0.5728</u>
# of ODE Solutions - <u>0.7</u> (Equivalent)

Extension to Finite Element Collocation

Motivation:

- Control approximation error
- Allow for discontinuities in control profiles
- Identify regions of active state and control variable bounds



Collocation points: $t_{ij} = \sum_{k=0}^{j-1} \propto_k + \propto_j \tau_i$

τ_i - normalized Legendre roots $\in$ (0, 1]

Since $\frac{dz}{d\tau} = \frac{dz}{d\tau} = \frac{1}{\propto_j} \frac{dz}{dt}$, ODE's and collocation equations are:

$$\frac{dz}{d\tau} = \alpha_j \ f(z,u,p)$$

and

$$r(t_{ij}) = \sum z_{kj} \ l'_k(\tau_i) - \propto_j f\left(z_{ij}, u_{ij} \ p\right) = 0$$

⇒ Extend to larger optimization problem (NLP).
Enforce continuity in state profiles.

SIMULTANEOUS APPROACH NLP FORMULATION with OCFE

Finite elements, α_j, fixed.

Min $\phi(z_{N,NE}(t_f))$

s.t. (Collocation Equations)

$$\sum_k z_{kj}\, l_k'(\tau_i) - \propto_j f(z_{ij}, u_{ij}, p) = 0$$

$z_{01} = z(0)$ $\quad i = 1,N, \; j = 1, NE - 1$

(Continuity of States)

$z_j(\tau=1) = z_{0,j+1}$ $\quad j = 1,NE - 1$

(Other Constraints Final Time, Paths)

$g(z_{ij},u_{ij},p) \leq 0$ $\quad g(z_f,p) \leq 0$

$h(z_{ij},u_{ij},p) = 0$ $\quad h(z_f,p) = 0$

What are the advantages of this approach?

Theoretical Properties of the Solution Method

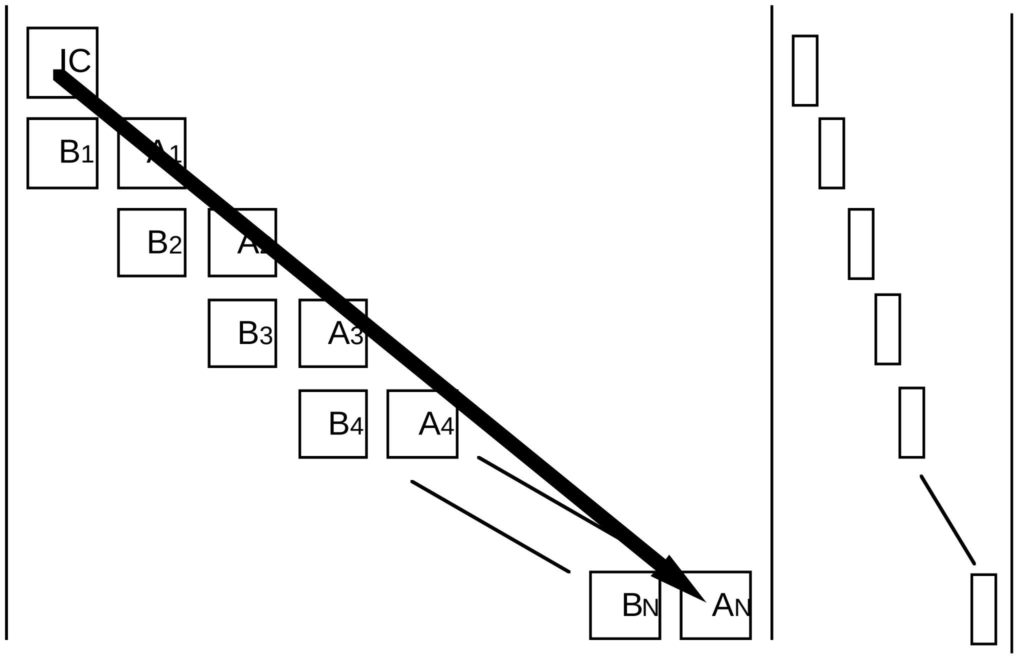
A. Stability and Accuracy of Orthogonal Collocation

- Equivalent to performing a *fully implicit* Runge-Kutta integration of the ODE models at Gaussiam points
- A high order (2N) method which uses N collocation points
- Algebraically stable (i.e., possesses A, B, AN and BN stability)

B. Analysis of the Optimality Conditions

- An equivalence has been established between the Kuhn-Tucker conditions of NLP and the variational necessary conditions
- Rates of convergence have been established for the NLP method

Forward Decomposition for Dynamic Optimization



- Rapid decomposition - fast competitor to sequential approach
- Element by element: same speed as solving IV problems
- Compatible with rSQP decomposition
- Requires dynamic system with stable modes

Instabilities in DAE Models

This example cannot be solved with sequential methods (Bock, 1983):

$$dy_1/dt = y_2$$

$$dy_2/dt = \tau^2 y_1 + (\pi^2 - \tau^2) \sin(\pi t)$$

The characteristic solution to these equations is given by:

$$y_1(t) = \sin(\pi t) + c_1 \exp(-\tau t) + c_2 \exp(\tau t)$$

$$y_2(t) = \pi \cos(\pi t) - c_1 \tau \exp(-\tau t) + c_2 \tau \exp(\tau t)$$

and both c_1 and c_2 can be set to zero by either of the following equivalent conditions:

a) $y_1(0) = 0$ $y_2(0) = \pi$

b) $y_1(0) = 0$ $y_2(1) = 0$

If we now add roundoff errors ε_1 and ε_2 to the conditions in 5a) and 5b) we see significant differences in the sensitivities of the solutions.

For the initial value case a), the sensitivity to the *analytic* solution profile is seen by large changes in the profiles $y_1(t)$ and $y_2(t)$ given by:

$$y_1(t) = \sin(\pi t) + (\varepsilon_1 - \varepsilon_2/\tau) \exp(-\tau t)/2$$
$$+(\varepsilon_1 + \varepsilon_2/\tau) \exp(\tau t)/2$$
$$y_2(t) = \pi \cos(\pi t) - (\tau \varepsilon_1 - \varepsilon_2) \exp(-\tau t)/2$$
$$+ (\tau \varepsilon_1 + \varepsilon_2) \exp(\tau t)/2$$

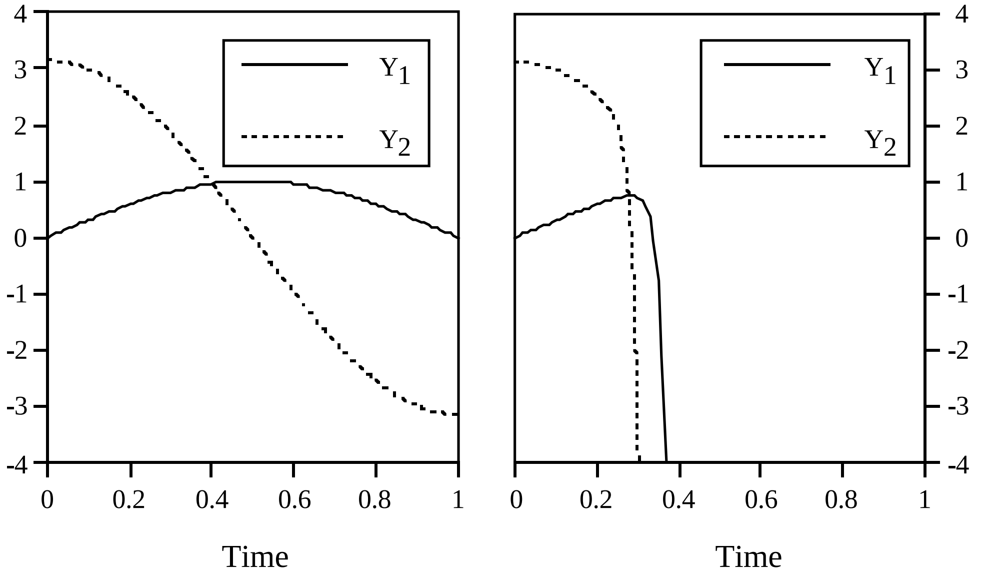
Therefore, even if ε_1 and ε_2 are at the level of machine precision ($< 10^{-13}$), a large value of τ and t will lead to unbounded solution profiles.

On the other hand, for the boundary value problem, case b), the errors affect the *analytic* solution profiles in the following way:

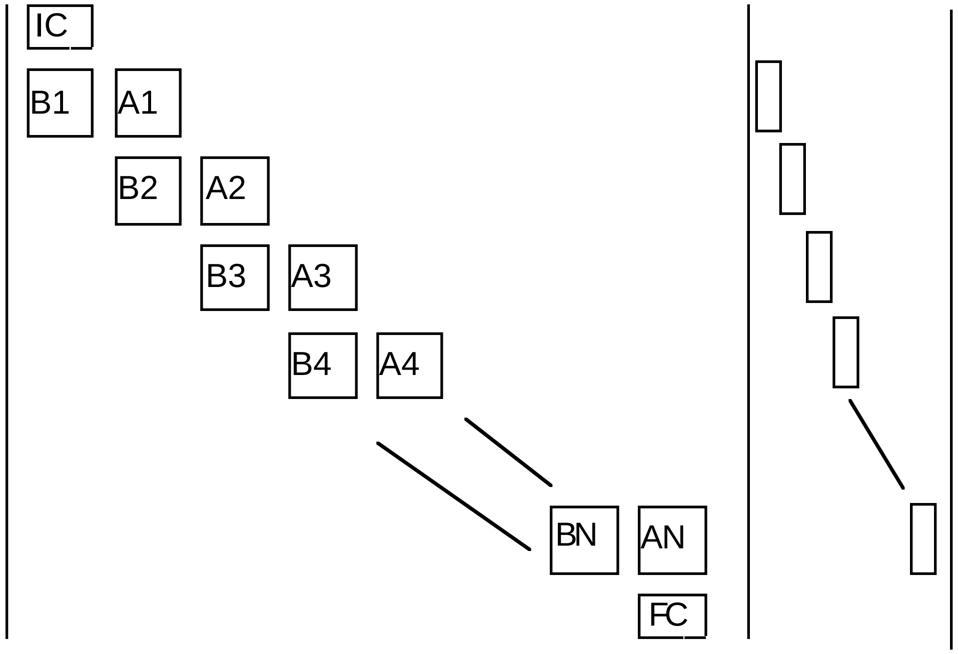
$$y_1(t) = \sin(\pi t) + [\varepsilon_1 \exp(\tau) - \varepsilon_2] \exp(-\tau t)/[\exp(\tau) - \exp(-\tau)] \\ +[\varepsilon_1 \exp(-\tau) - \varepsilon_2] \exp(\tau t)/[\exp(\tau) - \exp(-\tau)]$$

$$y_2(t) = \pi \cos(\pi t) - \tau [\varepsilon_1 \exp(\tau) - \varepsilon_2] \exp(-\tau t)/[\exp(\tau) - \exp(-\tau)] + \\ \tau [\varepsilon_1 \exp(-\tau) - \varepsilon_2] \exp(\tau t)/[\exp(\tau) - \exp(-\tau)]$$

Notice that the absolute errors in these profiles never exceed $\tau (\varepsilon_1 + \varepsilon_2)$, and as a result a solution to the BVP is readily obtained, as shown in the above studies.



BVP Problem Decomposition



- bound unstable modes with boundary conditions
- can be done implicitly by determining stable pivot sequences in simultaneous approach
- unstable modes determined automatically (deHoog and Mattheij)
- leads to repartitioning of control and state variables

Additional Features

- perform local repartitioning in each element
- leads to element-by-element decomposition
- computational and storage efficiency of forward strategy

Bock Problem (with $\tau = 50$)

- Sequential approach blows up (starting within 10^{-9} of optimum
- Simultaneous optimization requires 4 SQP iterations

Ethane Cracking Furnace Optimization

(Chen et al., 1996).

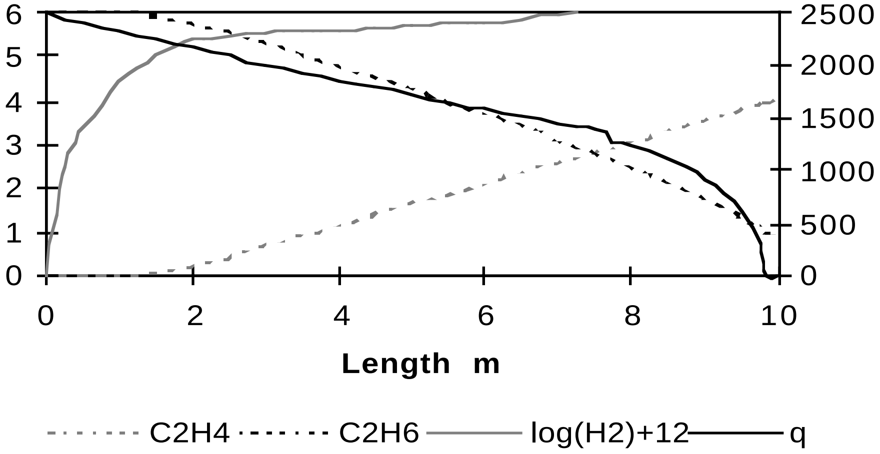
- Maximize Ethylene Production
- Manipulate Heat Flux Profile
- PFR Model with Unstable Modes
- Reactions:

$$C_2H_6 \rightarrow 2CH_3\bullet \qquad CH_3\bullet + C_2H_6 \rightarrow CH_4 + C_2H_5\bullet$$

$$H\bullet + C_2H_6 \rightarrow H_2 + C_2H_5\bullet \qquad C_2H_5\bullet \rightarrow C_2H_4 + H\bullet$$

$$C_2H_5\bullet + C_2H_4 \rightarrow C_3H_6 + CH_3\bullet \qquad 2C_2H_5\bullet \rightarrow C_4H_{10}$$

$$H\bullet + C_2H_4 \rightarrow C_2H_5\bullet$$



No. of ODEs	Discr'd Eq/Vars.	Basis selection	Iter'ns	CPU (s) HP 9000
11	1618/ 1631	MA28	66	282.5
11	1618/ 1631	MA28s	77	287.3
11	1618/ 1631	Forward	Failed	

CONTROLLING ELEMENT LENGTH

Role of Finite Elements

- Control approximation error in state profiles (sufficiently many elements).
- Allow for discontinuity in control profiles and active bounds (if present).

Approximation Error Criteria

- <u>Equidistribution</u> (for parameter optimization only)

For a given number of elements, find minimum approximation error, based on residual norm.

$$\underset{\propto_j}{\text{Min}} \left\{ \underset{j}{\text{Max}} (\text{error})_j \right\} \quad (\text{error})_j = C^1 \left\| r_j(t) \propto_j \right\|$$

$$\Rightarrow \quad \left\| r_j(t) \propto_j \right\| - \left\| r_{j+1}(t) \propto_{j+1} \right\| = 0 \quad j = 1, NE -1$$

- <u>Direct Specification of Tolerance</u>

For unspecified number of elements, find satisfactory approximation error, based on residual norm.

$$(\text{Error})_j = C^1 \left\| r_j(t) \propto_j \right\| \leq \varepsilon$$

Need to choose correct number of elements, NE, as an integer variable.

COMBINED NLP FORMULATION

Simultaneous Approach

Min $\phi(z_{N,NE}(t_f))$

s.t. (<u>Collocation</u> <u>Equations</u>)

$$\sum_k z_{kj}\, l_k'(\tau_i) - \propto_j f\left(z_{ij}, u_{ij}, p\right) = 0$$

$z_{01} = z(0)$ $\quad$ $i = 1,N,\ \ j = 1,\ NE - 1$

(<u>Continuity of States</u>)

$z_j(\tau=1) = z_{0,j+1}$ $\quad$ $j = 1,NE - 1$

(<u>Placement of Elements</u>)

$d_j(\alpha, z_{ij}, u_{ij}, p) \leq \varepsilon$

- element placement based on error tolerance

(<u>Other Constraints Final Time, Paths</u>)

$g\ (z_{ij}, u_{ij}, p) \leq 0$ $\quad$ $g\ (z_f, p) \leq 0$

$h\ (z_{ij}, u_{ij}, p) = 0$ $\quad$ $h\ (z_f, p) = 0$

How does this fomulation compare with traditional approaches?

Optimization - Collocation with Moving Finite Elements

Example: Hot Spot Reactor (Finlayson)

$$\min_{T_P, T_R, L, T_S, q(t), \bar{T}(t)} \quad \Phi = L - \int_0^L \left(\bar{T}(t) - T_s / T_R\right) dt$$

s.t.

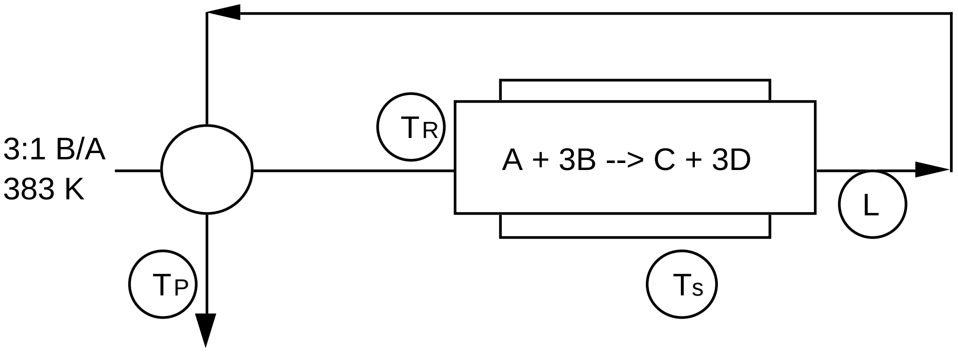
$$\frac{dq(t)}{dt} = 0.3(1 - q(t))\exp[20 - 20 / \bar{T}(t)] \qquad q(0) = 0$$

$$\frac{d\bar{T}(t)}{dt} = -1.5(\bar{T}(t) - T_s / T_R) + 2/3 \frac{dq(t)}{dt} \qquad \bar{T}(0) = 1$$

$$t \in [0, L]$$

$C_{ph} = \Delta H_{feed}(T_R, 110^oC) - \Delta H_{product}(T_P, T(L)) = 0$

$T_P \geq 120^oC, \; T(L) \geq T_R + 10^oC$



where $\bar{T}$ = $T(t)/T_R$

T_P = specified product temperature

T_R = reactor inlet, reference temperature

L = reactor length

T_S = steam sink temperature

$q(t)$ = reactor conversion profile

$\bar{T}(t)$ = normalized reactor temperature profile

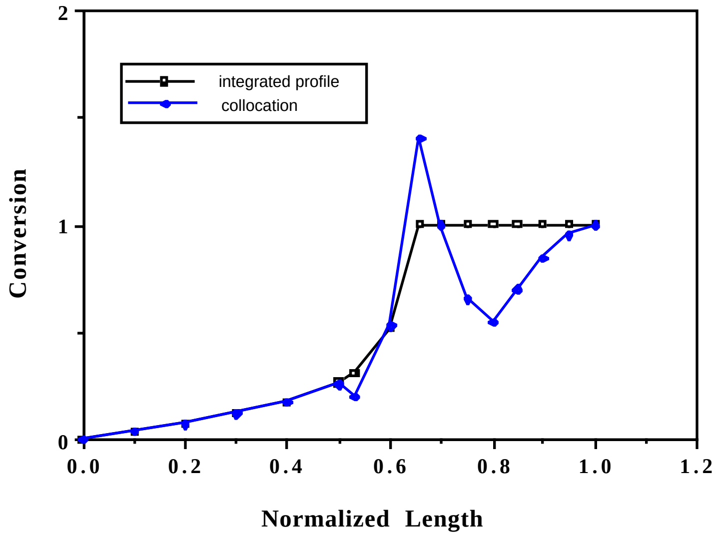
Cases considered:

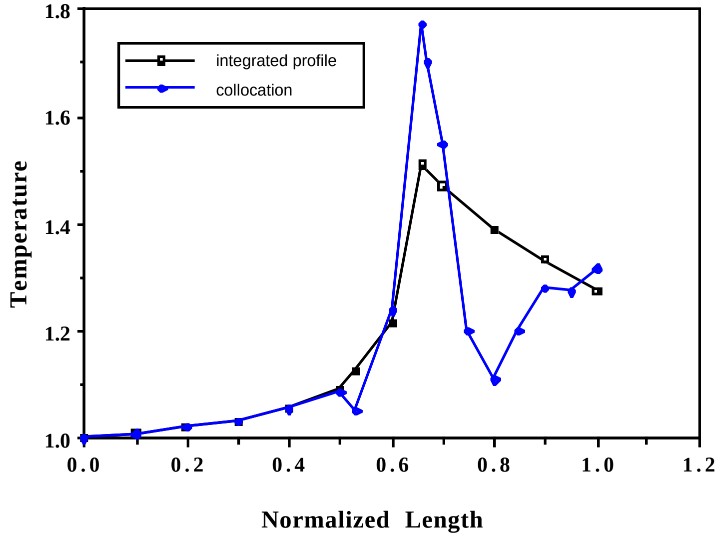
- Hot Spot - no state variable constraints
- Hot Spot with $\bar{T}(t) \leq 1.45$

Base Case

Method: OCFE at initial point with 6 equally spaced elements

	L(norm)	T_R(K)	T_S(K)	T_P(K)
Base Case:	1.0	462.23	425.26	250





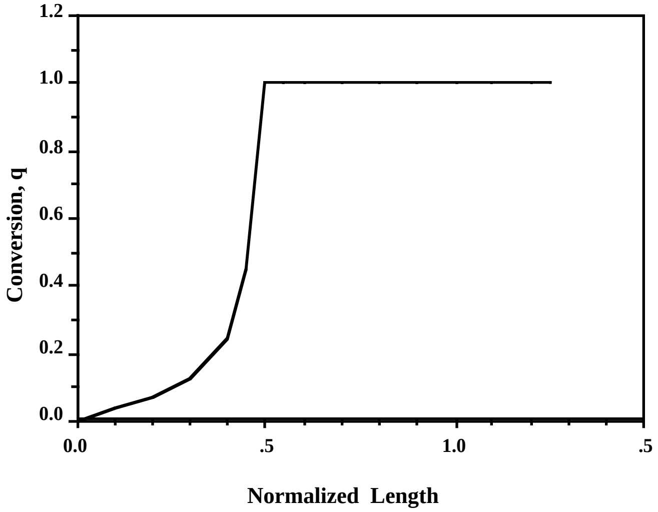
Unconstrained Case

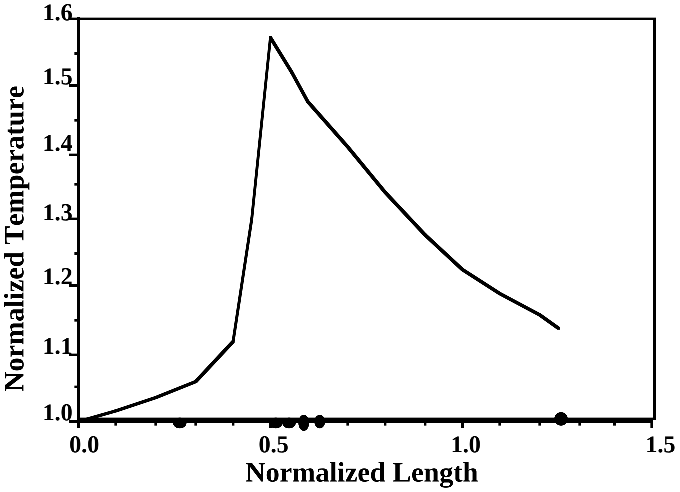
Method: OCFE combined formulation with rSQP identical to integrated profiles at optimum

	L(norm)	T_R(K)	T_S(K)	T_P(K)
Initial:	1.0	462.23	425.26	250
Optimal:	1.25	500	470.1	188.4

123 CPU s. (μVax II)

$\phi^* = -171.5$





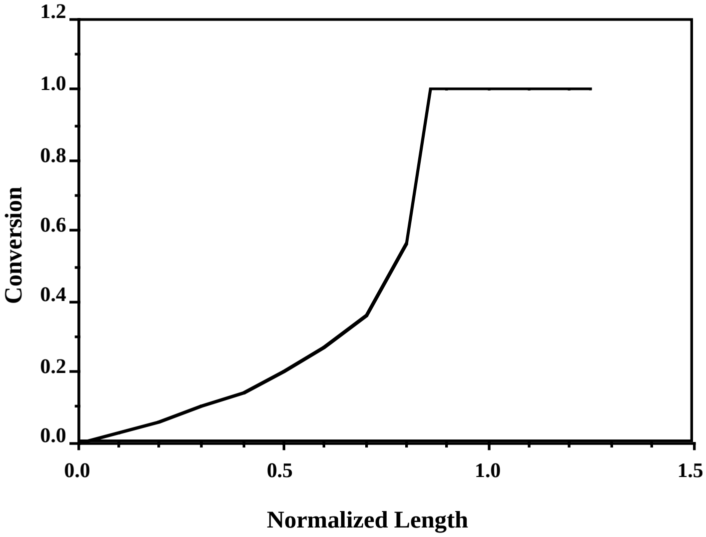
Temperature Constrained Case, $\overline{T} \leq 1.45$

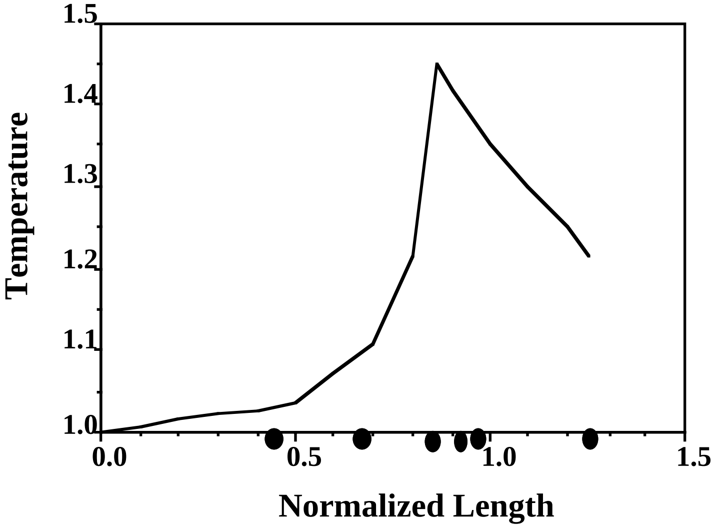
Method: OCFE combined formulation with rSQP, identical to integrated profiles at optimum

	L(norm)	T_R(K)	T_S(K)	T_P(K)
Initial:	1.0	462.23	425.26	250
Optimal:	1.25	500	450.5	232.1

57 CPU s. (μVax II)

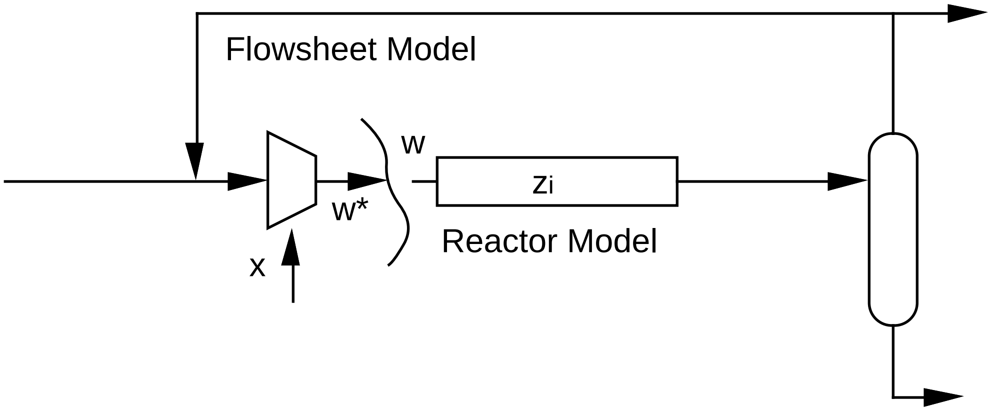
$\phi^* = -148.5$





Flowsheet Optimization with Reactors

- complicated functions for DAE's
- many levels of iteration
- implicit method necessary due to stiffness...
- complex thermodynamic models
- slowly converging algorithms



Problem Formulation

$$\min_{x, z_i, w} \quad \phi(x, z_i, w)$$

s.t.

Tear Equations $\quad h(x,y) = w - w^*(x,y) = 0$

Reactor Model: $\quad r(t_i) = \Sigma_i z_i - f(z_i,x,w) = 0$

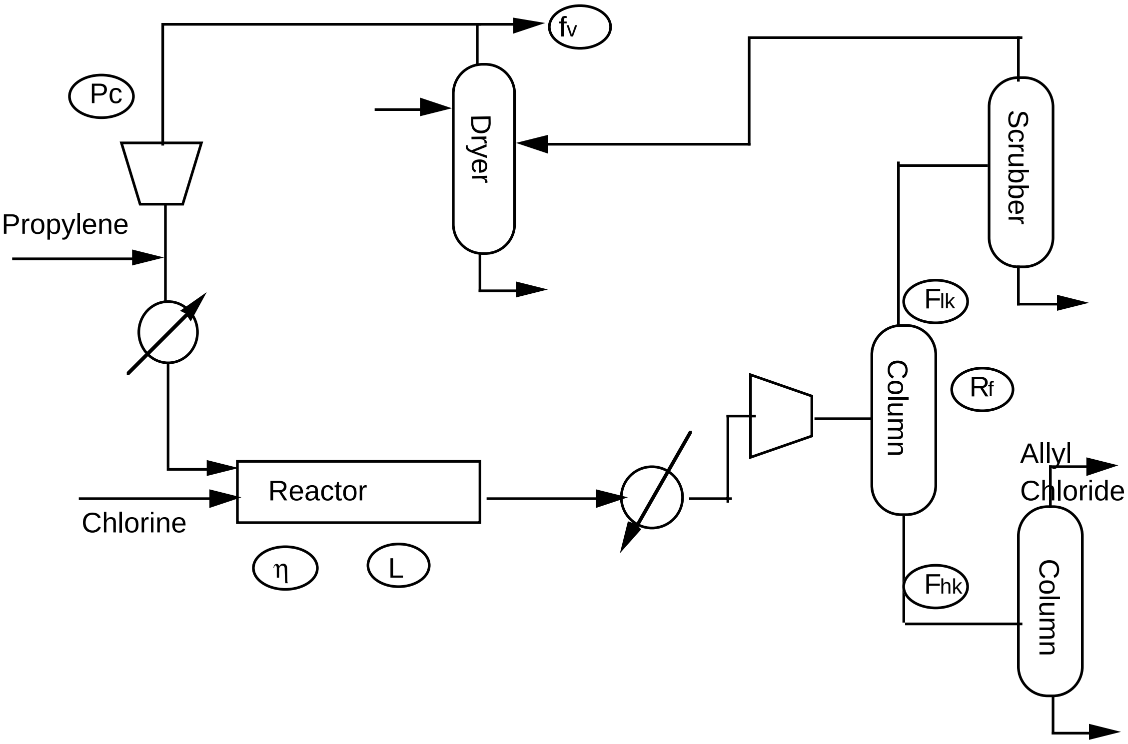
Flowsheet Constraints $\quad g(x,y) \leq 0$

$c(x,y) = 0$

Reactor Constraints $\quad g_r (x,z_i) \leq 0$

$c_r (x,z_i) = 0$

Propylene Chlorination Optimization



Max Net sales
s.t. Purity (99% chlorides)
Pressure constraints

Reactor in Flowsheet Optimization - Cases

- Fixed conversion model (OPTFIX)
- DAE model - solved in "Black-Box" strategy using LSODE (OPTODE)
- DAE model - solved simultaneously with orthogonal collocation and SQP (OPTCOL).

Optimization Results

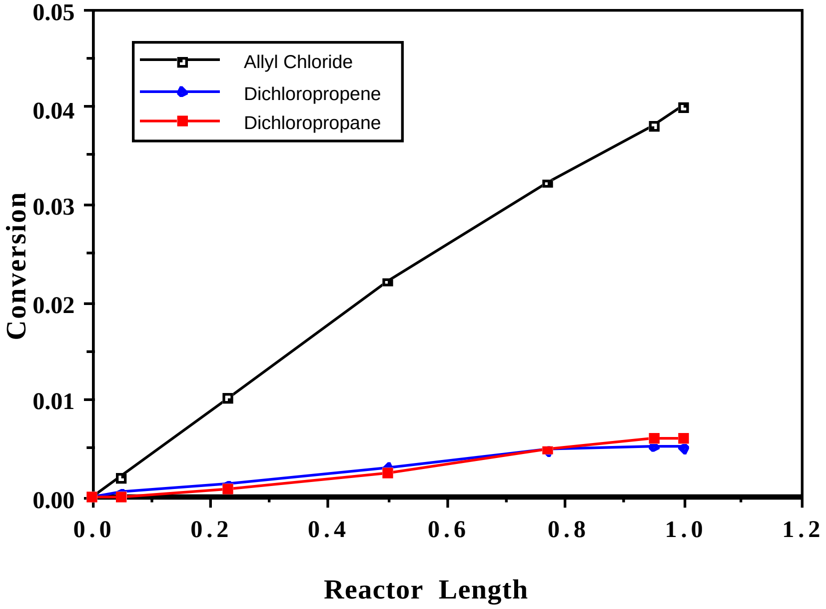
Variables (Conditions)	OPTFIX	OPTODE	OPTCOL
P_c (psia)	207.60	233.06	277.31
F_{lk}	90.14	97.98	95.23
F_{hK}	92.46	99.0*	95.37
R_f	1.989	1.996	1.59
n	6.999	4.50	4.56
T (^{o}F)	842.70	848.23	850.0*
L (feet)	-----	66.34	66.63
f_v	0.0887	0.0856	0.0928
ϕ ($/hr)	403.50	748.65	748.65
# iter	4	11	9
CPU sec $^+$	560.32	1485.98	945.85
CPU s/iter $^+$	140.08	135.09	105.09
CPUs /QP +	0.67	1.8	11.5

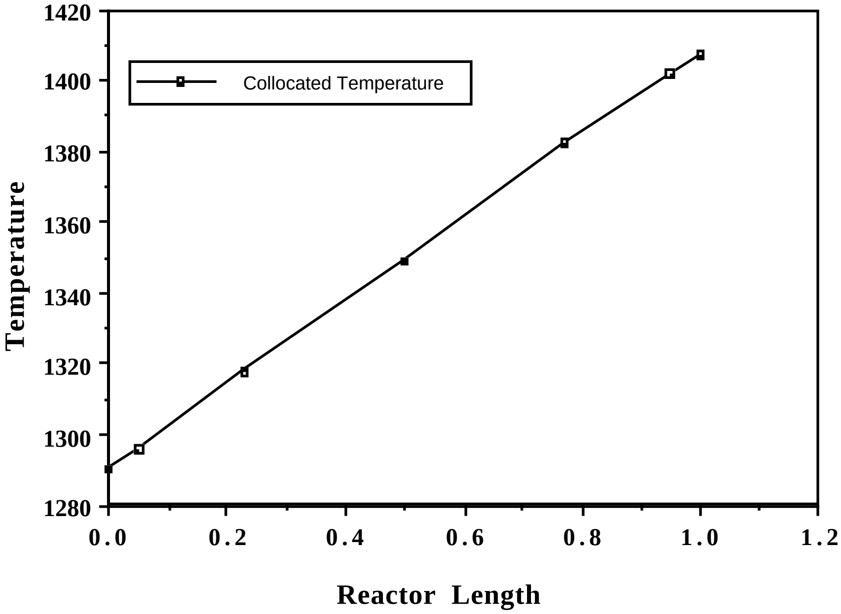
* variable at its bound

$^+$ Micro VAX II

Collocation vs. Integrated Reactor Models

- Rigorous enthalpy information used
- 5 sec. vs. 35 sec. for evaluation
- Exact Agreement





OPTIMAL CONTROL PROBLEMS
Characteristics of Control Profiles, u(t)

Parameterized Control Profiles

- Variable length, piecewise polynomial, etc.
- Only option in sequential approach
- Same characteristics as DAE model (e.g., index 1)

Control & States Approximated Together

- Solve controls with (similar) accuracy as states
- Formulate "implicit" DAE system.

$$\begin{aligned} &\nabla L\ (z,u) = 0 \\ &z' = f\ (z, u, t) \\ &g\ (z,u) \le 0 \\ &u^T g\ (z,u) = 0 \end{aligned}$$

ODE's plus algebraic optimality conditions

DAE Systems

Let $y = [z^T, u^T]^T$

$F(y, y', t) = 0$	General form
$Ay' + f(y,t) = 0$	Linear, Implicit
$z' = f(z,u,t)$ $g(z,u) = 0$	Semi-Implicit

Software Available

- LSODI (BDF, ODEPACK)
- DASSL (BDF, Petzold et al.)
- DASPK (BDF, Petzold et al.)
- LIMEX (Deuflhard et al.)
- RADAU5 (Hairer and Wanner)

Degree of Difficulty for DAE System ==>Index:

Min number of times algebraic variables need to be differentiated to yield ODE system

Index of DAE Problem can be Arbitrarily High

Staged Process

$$\dot{x}_1 = f(x_1, x_2)$$
$$\dot{x}_2 = f(x_2, x_3)$$
$$----------------$$
$$\dot{x}_n = f(x_n, u)$$
$$g(x_1) = 0$$

Index of DAE system is n+1

Treatment of High Index DAE's

1) Reformulate to lower index system, apply index 1 DAE (or ODE) solver.

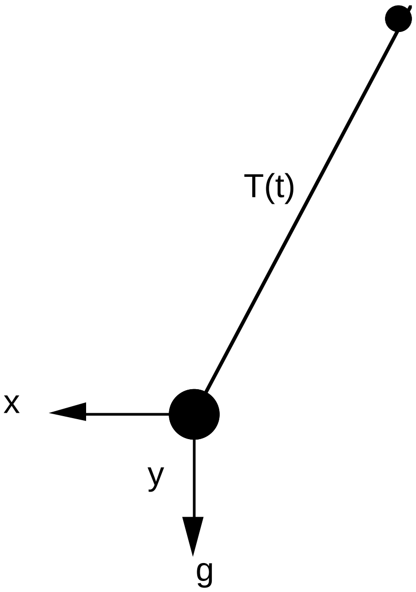
   - need consistent initial conditions
   - errors in ODE's may violate equations

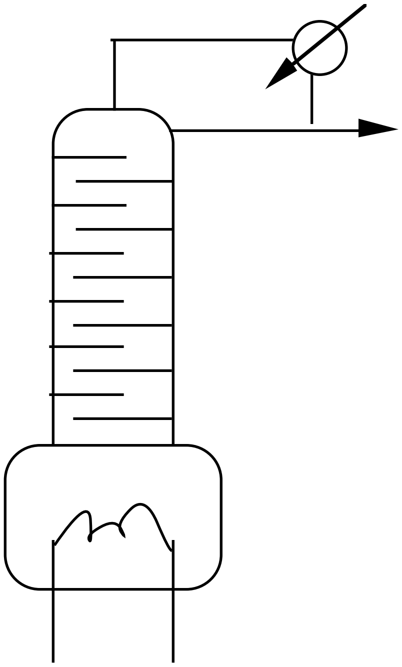
2) Apply higher index DAE solver (Gear et al.)

   - difficult to find for index ≥ 3
   - require well implemented BDF or R-K method.

Example: Pendulum Problem (index 3)

$x' = u$
$y' = v$
$u' = Tx$
$v' = g-Ty$
$x^2 + y^2 = 1$
$2xu + 2Yv = 0$
$2(-T + yg + u^2 + v^2) = 0$
$T' = 3\,vg - 2Tux - 2Tvy$





Process Example: Batch Distillation (index ?)

$x' = f(x,R)$
$x_{N+1} = 0.99$
find R.

Why are High Index DAE's Difficult to Solve?

- Possible order reduction in approximation method for higher index variables
  - need correct adjustment by DAE solver
  - can have O(1/h) approximation
- Can lead to unstable error propagation
  - different criteria for R-K or BDF methods.
- Implementation details
  - variables inconsistently initialized (failure)
  - singularities in Jacobian for Index ≥ 2, for $h=0$.

High index DAEs can be reformulated to index 1 or 2

- over the entire time domain.

Optimal Control Problems as DAE's

Index - 1^+ Nonsingular, nonstate constrained optimal control.

$$\partial H/\partial u = \Omega(u) = 0$$
$$\dot{z} = f(z,u)$$

Index - 2^+ State variable constrained problem.

$$\dot{z} = f(z,u)$$
$$g(z) \leq 0 \quad \left\{ \frac{\partial g}{\partial z} = \vartheta(u) = 0 \right\}$$

Index - 3^+ Singular optimal control problem

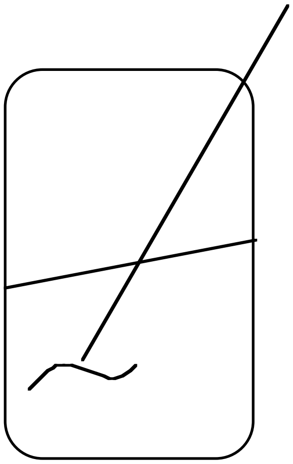
$$\frac{\partial H}{\partial u} \neq \vartheta(u) = 0; \; \frac{\partial \dot{H}}{\partial u} = 0 \neq \vartheta(u)$$
$$\frac{\partial \ddot{H}}{\partial u} = \vartheta(u)$$
$$\dot{z} = f(z,u)$$

Index can change over different parts of profile and this is not known before the optimal solution is known.

- need DAE discretizations that can deal with increase in DAE Index. . .
- for orthogonal collocation:
  - order reduction for R-K methods
  - stability problems for orthogonal collocation (L-stable method ==> Radau collocation)

Index 1

Batch Reactor Optimization



A --> B --> C

$$\max_{T} \quad c_2(1.0)$$

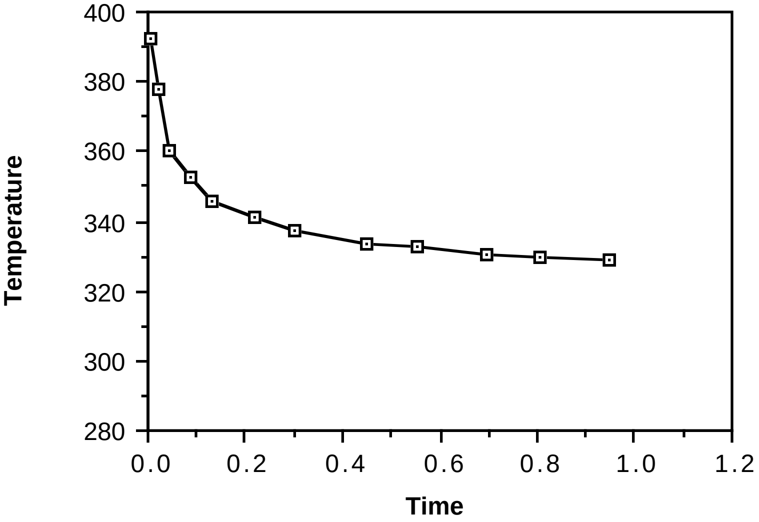
$$\text{s.t.} \quad \frac{dc_1}{dt} = -k_1(T)\,c_1^2$$

$$\frac{dc_2}{dt} = -k_1(T)\,c_1^2 - k_2(T)\,c_2$$

$$k_i(T) = A_{i0} \exp\{-E_i/RT\} \qquad i=1,2$$

$$c_1(0) = 1.0,\ c_2(0) = 0$$

$$298 \le T \le 398$$

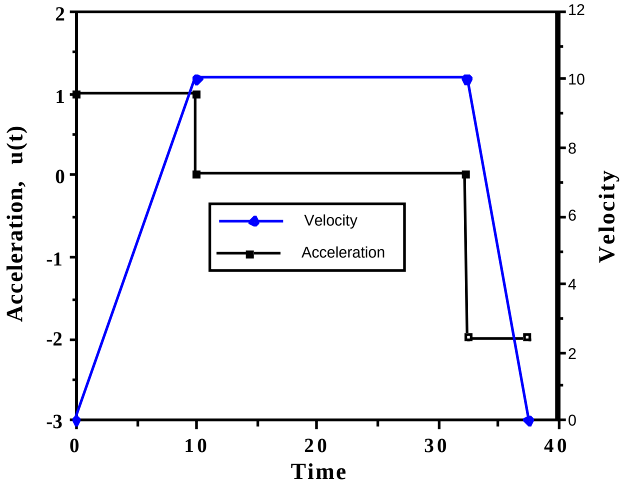


Index 2

Car Problem [minimum time with speed limit]

$$
\begin{aligned}
\text{Min} \quad & t_f \\
\text{s.t.} \quad & z_1' = z_2 \\
& z_2' = u \\
& z_2 \le z_{max} \\
& -2 \le u \le 1
\end{aligned}
$$

Data from "Car Problem"

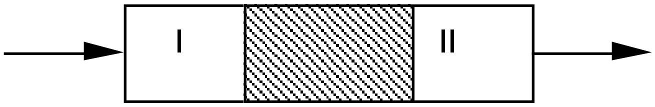


Identical to Analytical Results

- 2-point collocation
- 3 elements
- Linear profiles
- Tight error control required

Index 3

Singular Control Problem



A <==> B --> C
I II

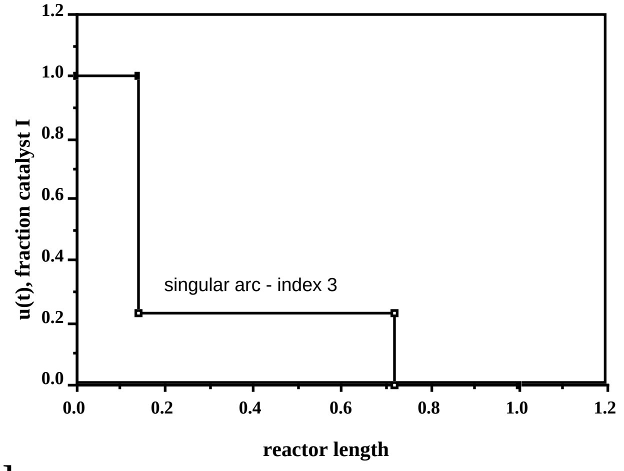
$$\text{Max} \quad 1 - z_a(t_f) - z_b(t_f)$$

$$\text{s.t.} \quad z_a' = u\,(k_{-I}\, z_b - k_I\, z_a)$$

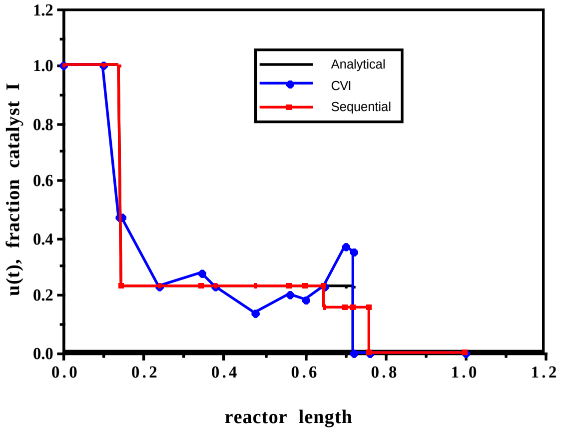
$$z_b' = u\,(k_I\, z_a - k_{-I}\, z_b) - (1-u)\, k_{II}\, z_b$$

Simultaneous Solution

- Radau Collocation on Moving Finite Elements
- Matches Analytical Solution



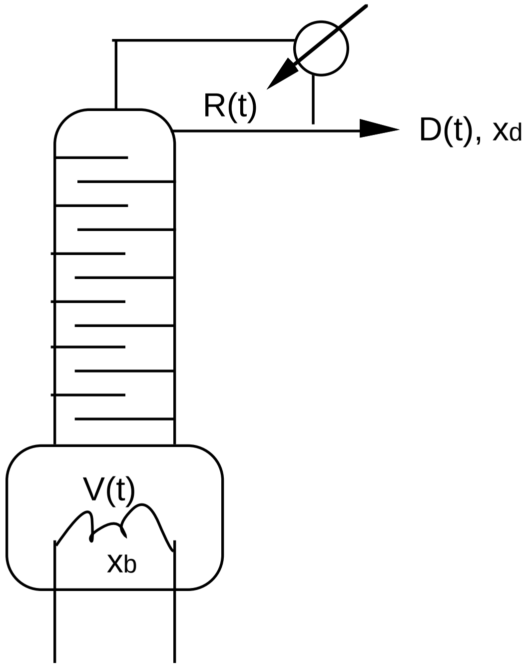
Other Methods



SIMULTANEOUS DAE OPTIMIZATION
Case Studies

- Reactor - Based Flowsheets
- Fed-Batch Penicillin Fermenter
- Temperature Profiles for Batch Reactors
- Parameter Estimation of Batch Data
- Synthesis of Reactor Networks
- Batch Crystallization Temperature Profiles
- Grade Transition for LDPE Process
- Ramping for Continuous Columns
- Reflux Profiles for Batch Distillation and Column Design
- Batch Process Integration

Some Interesting Optimal Profiles
Batch Distillation Column



Optimal Reflux with Holdup

$$\max_{\substack{R(t),V(t)\\ 0 \le t \le t_f}} \quad D(t_f)$$

Mass Balances

$$\frac{dS}{dt} = -\frac{V}{R+1}$$

$$\frac{dx_b^j}{dt} = -\frac{V}{R+1}\frac{\left[x_b^j - x_d^j\right]}{S}$$

S = overall mass

j = 1,2,...number of components

$$\frac{dx_{1,N+1}}{dt} = \frac{V}{H_{N+1}}\left[y_{1,N} - x_{1,N+1}\right]$$

$$\frac{dx_{1,p}}{dt} = \frac{V}{H_p}\left[y_{1,p-1} - y_{1,p} + \tfrac{R}{R+1}\left(x_{1,p+1} - x_{1,p}\right)\right], p = 1,\ldots,N$$

$$\frac{dx_{1,0}}{dt} = \frac{V}{S}\left[x_{1,0} - y_{1,0} + \tfrac{R}{R+1}\left(x_{1,1} - x_{1,0}\right)\right]$$

$$\frac{dD}{dt} = \frac{V}{R+1}$$

Equilibrium Equations

$$\sum_{1}^{C} x_{i,p} = 1.0 \qquad \sum_{1}^{C} y_{i,p} = 1.0 \qquad y_{i,p} = K_{i,p}\ x_{i,p}$$

Initial Conditions

$$S^0 x^0_{i,0} = \left(S^0 - \sum_{p=1}^{N+1} H_p\right) x_{i,0} + \sum_{p=1}^{N+1} H_p\, x_{i,p} \qquad i = 1,2,\ldots,C$$

Binary Column (55/45, Cyclohexane, Toluene)
$S^0 = 100$, $H_p = 1$, $N = 10$, ~8000 variables, < 2 CPU hrs. (Vaxstation 3200)

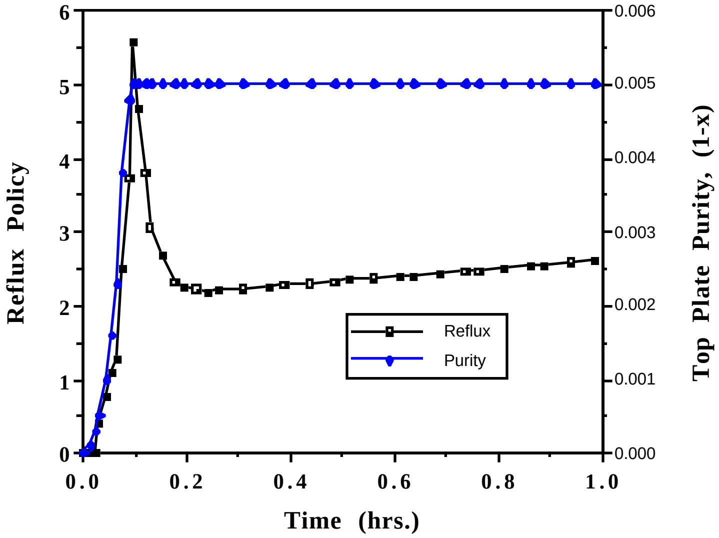
Cases Considered

- Path Purity Constraint
- Integral Purity Constraint at t_f
- Optimal Constant Reflux
- Optimal Piecewise Constant Reflux

Optimal Reflux Policies

Path Purity Constraint $(x_1(t) \geq 0.995)$

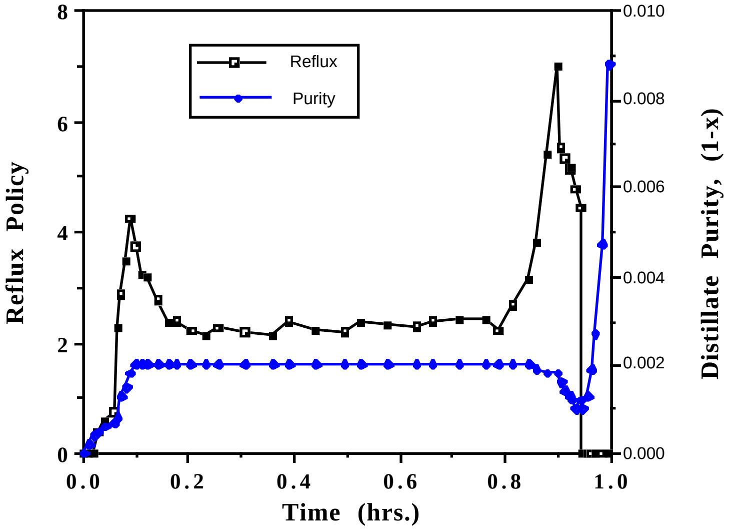
$D^*(t_f) = 38.61$



Integral Purity Constraint

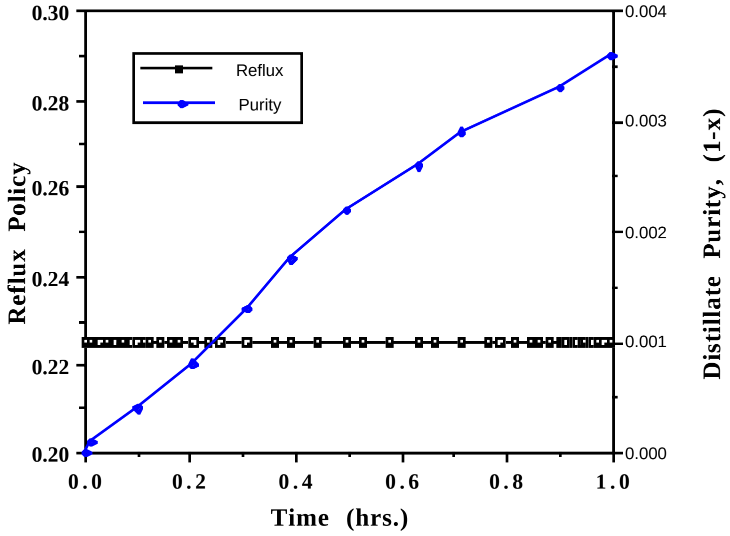
$(\int x_d(t)\, V/(R+1)\, dt)\, /D(t_f) \geq 0.998$

$D^*(t_f) = 42.34$



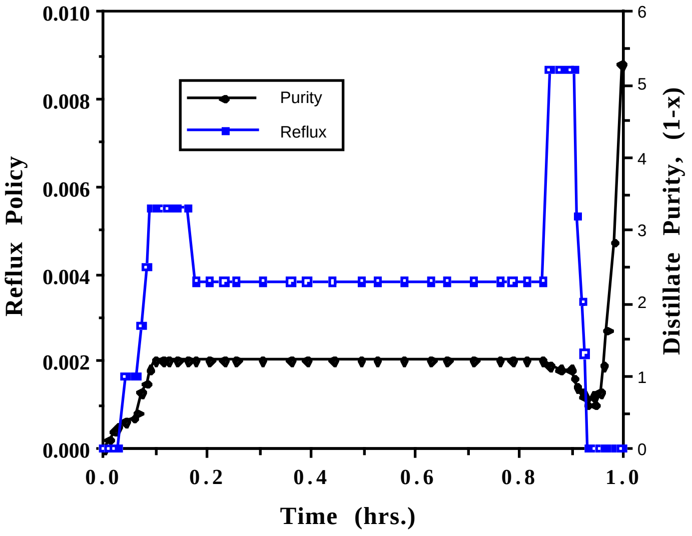
Optimal Constant Reflux (Integral Purity Constraint)

$D^*(t_f) = 38.9$



Piecewise Constant Policy (Integral Purity Constraint)

$D^*(t_f) = 42.26$



Mimics Optimal Continuous Policy

Dynamic Optimization as a Framework for Integration

- Naturally and directly handles interactions and multiple conditions
- Trade-offs handled in an unambiguous, quantitative manner
- Larger problems to solve

Research Questions

1. Which large-scale algorithms, software and hardware are available?

2. What are the best problem formulations?

3. Is further algorithmic development needed?

Dynamic Optimization Studies for Process Integration
Some Recent Literature Results

- Design and Dynamic Performance
  (Logsdon and Biegler, 1991)

- Design, Scheduling and Dynamic Performance
  (Bhatia and Biegler, 1996)

- Scheduling and Dynamic Performance
  (Mujtaba and Macchieto, 1992)

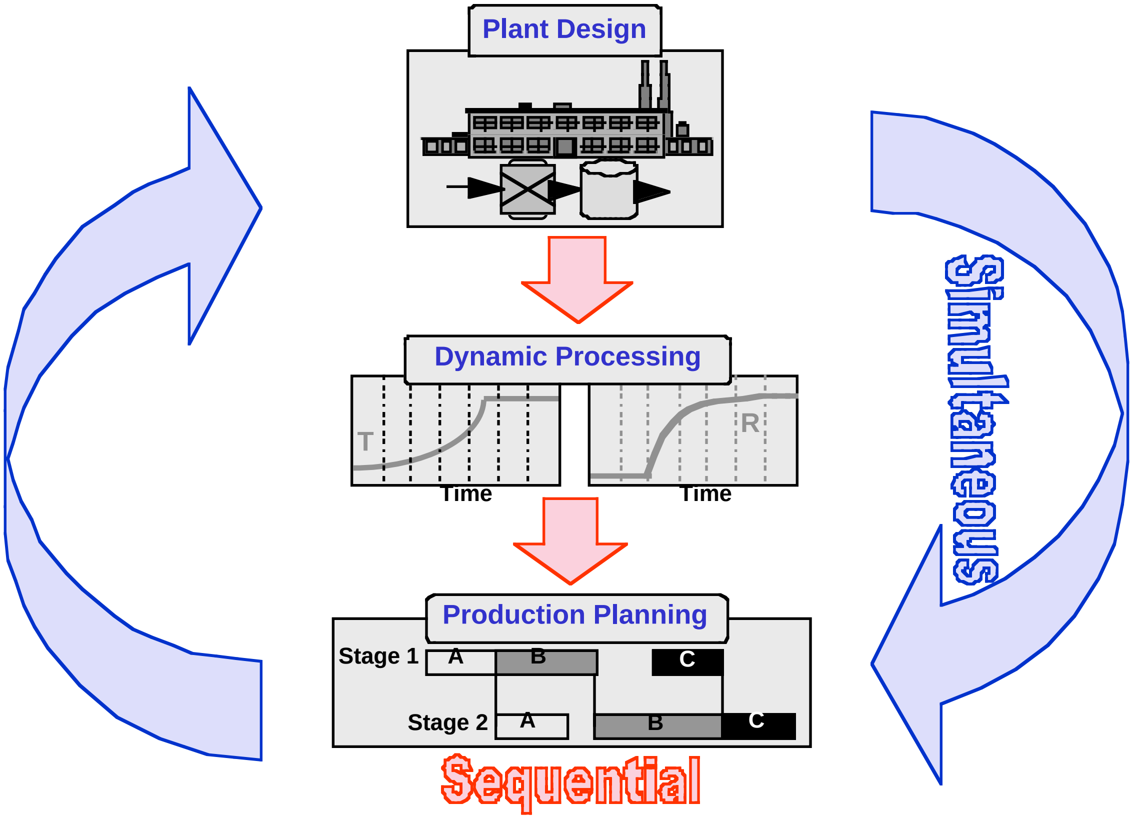
- Interactions of Process Subsystems
  (Balakrishna and Biegler, 1993; Lakshmanan and Biegler, 1996 -)

- Interactions of Control and Design
  (Perkins and coworkers, 1985 - ; Perkins and Walsh, 1988; Mohideen et al., 1997)

- Interactions of Maintenance and Design
  (Pistikopoulos et al., 1995)

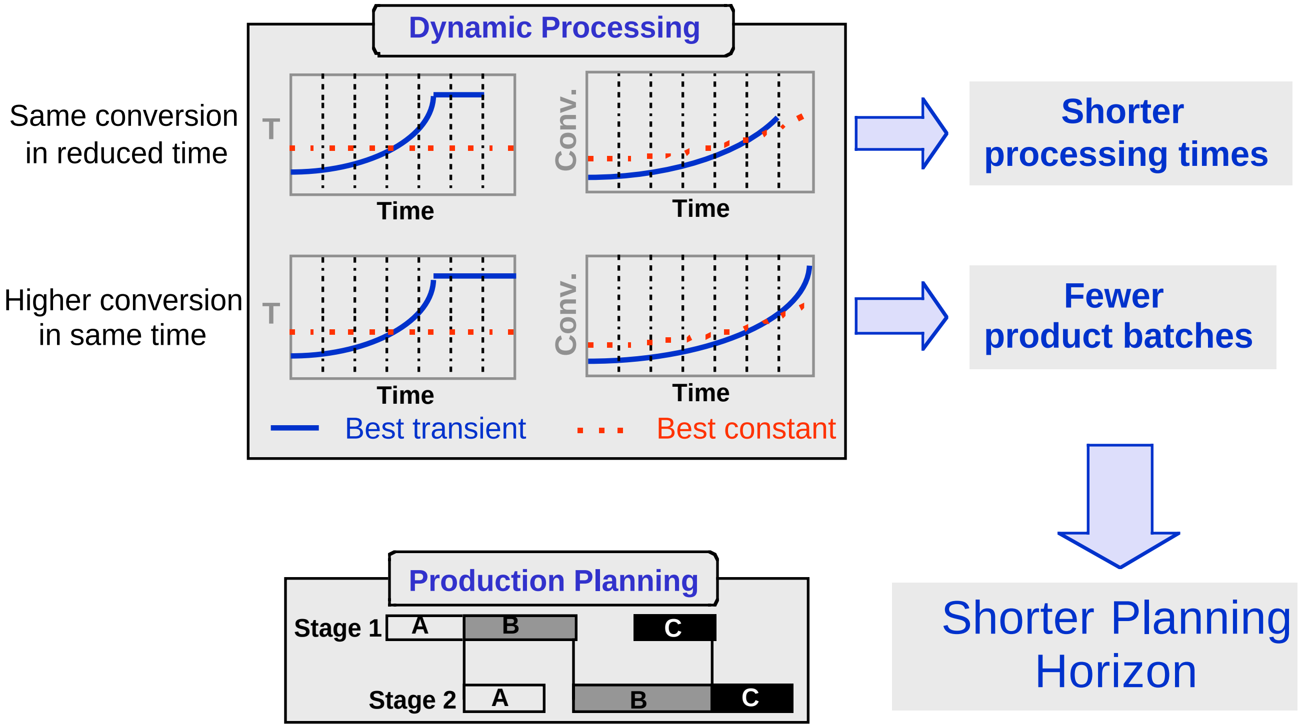
- Safety, Design and Performance
  (Abel, Helbig and Marquardt, 1997)

A Simple Case Study



- What are the Interactions between Design, Dynamics and Planning?
- What are the Differences between Sequential and Simultaneous Strategies?
- Especially Important in Batch Systems

Some Issues -



Same conversion in reduced time

Higher conversion in same time

Shorter processing times

Fewer product batches

Shorter Planning Horizon

1. Simultaneous Dynamic Optimization

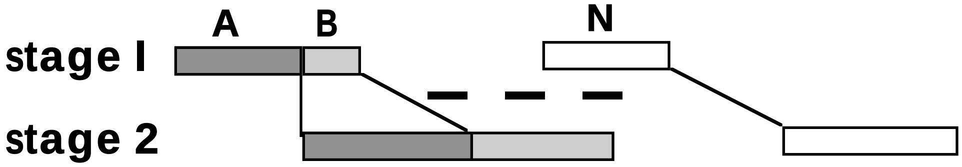
- discretize differential equations (DAEs) and state and control profiles
- large-scale optimization problem
- handles profile constraints directly
- incorporates equipment design variables directly
- DAE model solved only once
- converges for unstable systems

2. Scheduling Formulation

- combinatorial sequencing of tasks, products and equipment
- expensive discrete optimization, often solved by heuristics
- will consider ideal transfer policies ==> leads to closed form relations

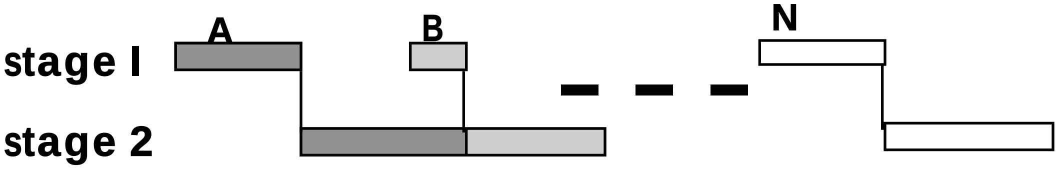
Unlimited Intermediate Storage (UIS)

- Short production cycle
- Cycle time independent of product sequence



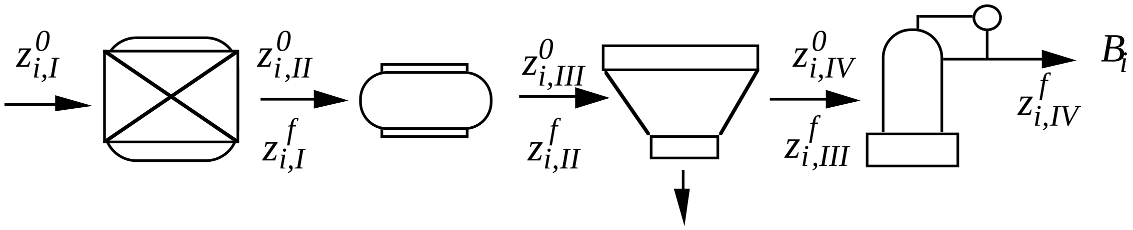
Zero Wait (ZW)

- Immediate transfer required
- Slack times dependent on product pair
- Longer production cycle required



Process Example

$$\boxed{\begin{array}{l} A+B\rightarrow C \\ C+B\rightarrow P+E \\ P+C\rightarrow G \end{array}}$$



- 4 stages
- 3 products (P) of different (high) purity
- dynamic reactor - temperature profile
- dynamic distillation column - reflux profile

Profit Maximization Cases

- SQ - Sequential Design - Scheduling - Dynamics
- SM - Simultaneous Design and Scheduling Dynamics with fixed endpoints
- SM* - Simultaneous Design, Scheduling and Dynamics

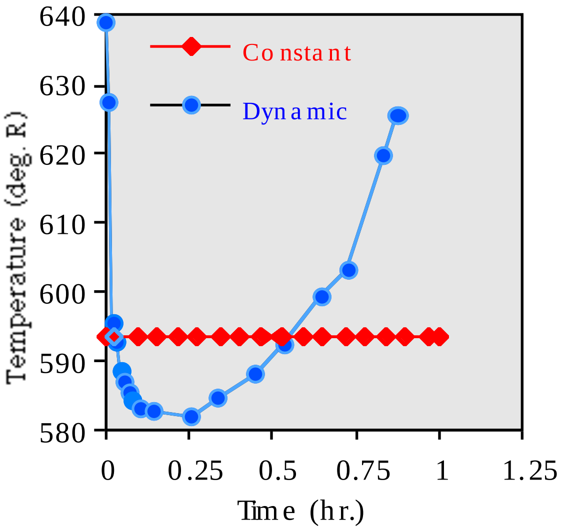
Comparison of Dynamic vs. Best Constant Profiles

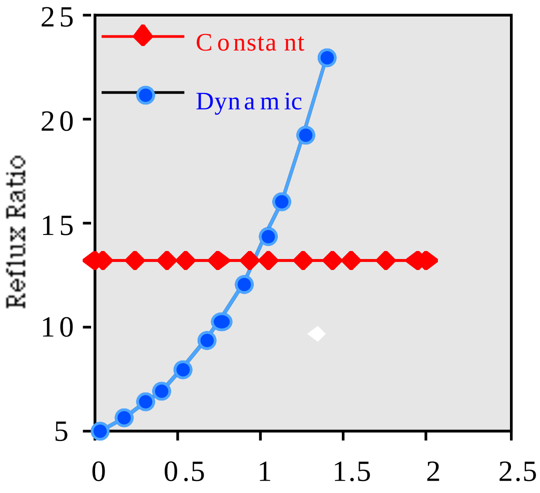
R0 - best constant temperature profile

R1 - optimal temperature policy

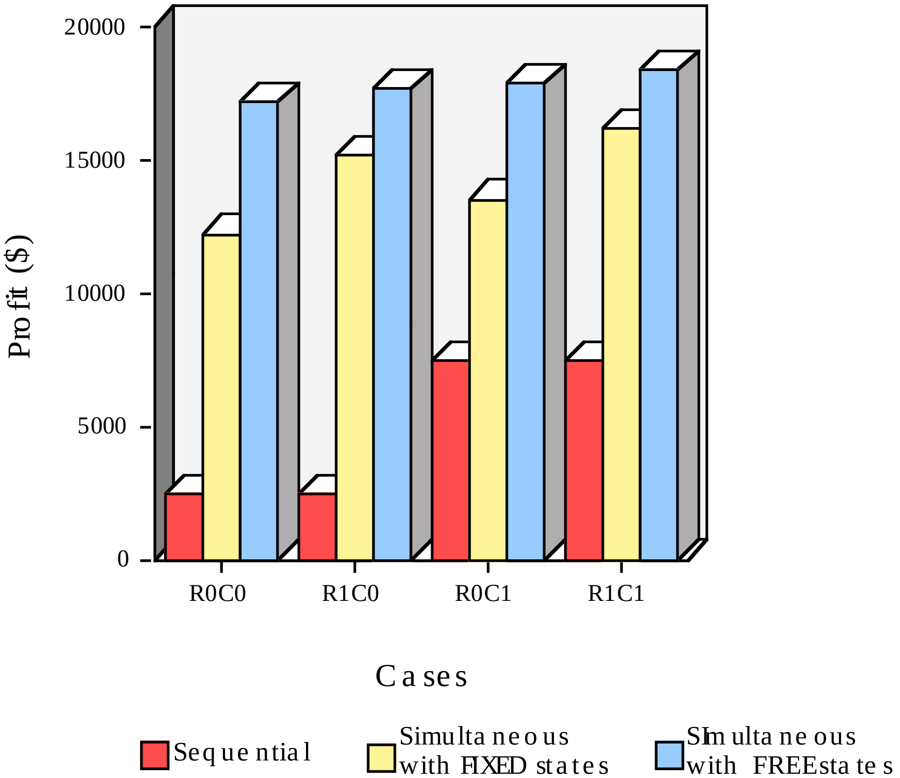
C0 - best constant reflux ratio

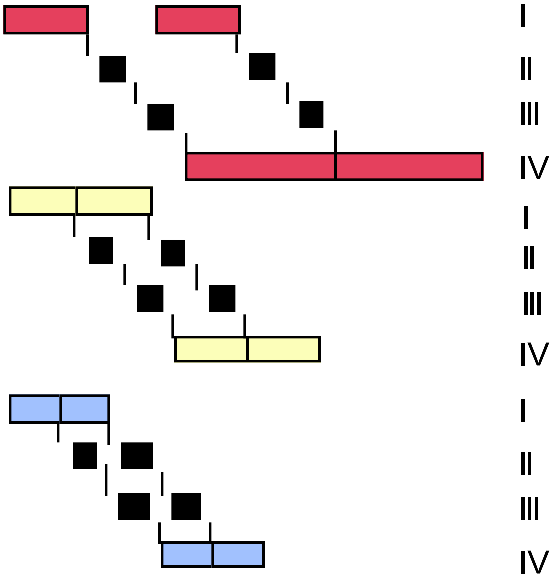
C1 - optimal reflux ratio





Results for Simultaneous Formulation-ZW Case





- ZW schedule becomes tighter
- less dependent on product sequences.

Summary

Sequential Approaches

- Parameter Optimization
    - Gradients by:
      Adjoints
      Sensitivity Equations
- Optimal Control (Profile Optimization)
    - Variational Methods
    - NLP-Based Methods
- Require Repeated Solution of Model
- State Constraints are Difficult to Handle

Simultaneous Approach

- Discretize ODE's using orthogonal collocation on finite elements (solve larger optimization problem)
- Accurate approximation of states, location of control discontinuities through element (α) placement.
- Straightforward addition of state constraints.
- Deals with unstable systems

Simultaneous Strategies are Effective

- Directly enforce constraints
- Solve model only once
- Avoid difficulties at intermediate points

Special Formulations

- Simultaneous error control
- Treatment of DAE's

Large-Scale Extensions

- SQP decomposition method
- Exploit structure of DAE discretization

Reference Texts

Bryson, A.E. and Y.C. Ho, Applied Optimal Control, Ginn/Blaisdell, (1968).

Himmelblau, D.M. and T.F. Edgar, Optimization of Chemical Processes, McGraw-Hill, (1987).

Ray. W.H., Advanced Process Control, McGraw-Hill, (1981).

Software

- Dynamic Optimization Codes

  DynoPC - simultaneous optimization code (CMU)
  COOPT - sequential optimization code (Petzold)
  gOPT - sequential code integrated into gProms (PSE)
  MUSCOD - multiple shooting optimization (Bock)
  NOVA - SQP and collocation code (DOT Products)

- Sensitivity Codes for DAEs

  DASOLV - staggered direct method (PSE)
  DASPK 3.0 - various methods (Petzold)
  SDASAC - staggered direct method (sparse)
  DDASAC - staggered direct method (dense)
  DSL48s - staggered corrector method (Barton)

Further References

Albuquerque, J. and L.T. Biegler, "Decomposition Algorithms for On-line Estimation with Nonlinear DAE Models," *Computers and Chemical Engineering*, 21, 3, p. 283 (1997)

Albuquerque, J., V. Gopal, G. Staus, L. T. Biegler, and B. E. Ydstie, "Interior Point SQP Strategies For Structured Process Optimization Problems" *Computers and Chemical Engineering*, **21**, p. S853 (1997)

Ascher U.M. and R.J. Spiteri. Collocation Software for Boundary Value Differential-Algebraic Equations. *SIAM J. Sci. Comput*. 1994, 15 (4), 939-952.

Ascher U.M., R.M. Mattheij and R.D. Russell. *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*; Prentice-Hall: Englewood Cliffs, NJ, 1988.

Betts, J.T.; Huffman, W.P.; "Application Of Sparse Nonlinear Programming To Trajectory Optimization," J. Guid. Control Dyn., Vol.15, No.1; Jan.-Feb. 1992; pp. 198-206

Betts, J. T. and P. D. Frank, "A sparse nonlinear optimization algorithm," *J. Opt. Theo. Applics.*, 82 (1994), p. 519

Bhatia, T., and L. T. Biegler, "Dynamic Optimization in the Design and Scheduling of Multiproduct Batch Plants," *I & EC Research*, **35**, 7, p. 2234 (1996)

Biegler, L. T., "Advances in Nonlinear Programming Concepts for Process Control," *Rroc. ADCHEM '97*, p. 543 (1997)

Biegler, L. T., J. Nocedal and C. Schmid, "A Reduced Hessian Method for Large Scale Constrained Optimization," *SIAM J. Optimization*, **5**, 2, p. 314 (1995)

Bock, H. G., Recent Advances in Parameter Identification Techniques for ODEs, *Numerical Treatment of Inverse Problems in Differential and Integral Equations*, P. Deuflhard and E. Hairer (eds.), Birkhauser, Heidelberg (1983)

Cervantes, A. and L. T. Biegler, "Large-Scale Dynamic Optimization Using Simultaneous Nonlinear Programming Formulations," submitted for publication (1997)

Chen Q. et al.. The Equation-Based SPYRO Model for Olefin-Porduction Applications. *AIChE Spring National Meeting*. 1996.

De Hoog R. and M. Mattheij. On Dichotomy and Well Conditioning in BVP. *SIAM J. Numer. Anal.* 1987, 24(1), 89-105.

Mohideen, J., J. Perkins and S. Pistikopoulos., "Design of Dynamic Systems Under Uncertainty," Paper 190c, Annual AIChE Meeting, Miami Beach, 1995

Santos, L.O., N. deOliveria and L.T. Biegler, "Reliable and Efficient Optimization Strategies for Nonlinear Model Predictive Control," *Proc. Fourth IFAC Symposium on Dynamics and Control of Chemical Reactors, Distillation Columns and Batch Processes (DYCORD '95), p. 33 * (1995).

Tanartkit, P. and L. T. Biegler, "Stable Decomposition for Dynamic Optimization," *I & EC Research*, **34**, p. 1253 (1995)

Tanartkit, P., and L. T. Biegler, "Reformulating Ill-Conditioned DAE Optimization Problems," *I & EC Research*, **34**, p. 1253 (1995)

Tjoa, I-B and L.T. Biegler, "Simultaneous Solution and Optimization Strategies for Parameter Estimation of Differential-Algebraic Equation Systems," I&EC Research, 30, p. 376 (1991).