Nonsmooth Differential-Algebraic Equations

Paul I. Barton

Process Systems Engineering Laboratory
Massachusetts Institute of Technology
Dynamic Modeling Frameworks in PSE

- Trade-off: applicability vs. ease of modeling & solving

Smooth Models
- Limited applicability
- Easy to model and solve
- Derivative information
- Strong existence & uniqueness theory

Nonsmooth Models
- Broad applicability
- Easy to model and solve
- Generalized derivative information
- (recently) Strong existence & uniqueness theory

Hybrid (Discrete/Continuous) Models
- (near) Universal applicability
- Often challenging to model and solve
- Limited derivative information
- Pathological behaviors (hard to exclude a priori)

Smooth approximation models
- Complementarity systems
Hybrid Automaton Framework

- Simple const. $P$ flash process:

\[ \dot{H}(t) = U(T_{out} - T(t)) \]

\[ M = M_L(t) + M_V(t) \]

\[ H(t) = Mh_V(t) - M_L(t)\Delta h_{vap}(T(t)) \]

\[ h_V(t) = Cp(T(t) - T_0) \]

\[ \log(P_{sat}(t)) = A - B/(T(t) + C) \]

- DAE embedded in hybrid automaton

**Mode 1**

$M_V = 0, M_L > 0$

$P \geq P_{sat}(T)$

**Mode 2**

$M_V, M_L > 0$

$P = P_{sat}(T)$

**Mode 3**

$M_V > 0, M_L = 0$

$P \leq P_{sat}(T)$
Hybrid vs. Nonsmooth

- Hybrid automaton formulation

```
\begin{align*}
\text{Mode 1} & : M_V = 0, \quad M_L > 0 \\
& : P \geq P^{\text{sat}}(T) \\
\text{Mode 2} & : M_V, \quad M_L > 0 \\
& : P = P^{\text{sat}}(T) \\
\text{Mode 3} & : M_V > 0, \quad M_L = 0 \\
& : P \leq P^{\text{sat}}(T)
\end{align*}
```

- "Continuous" disjunction:

```
\begin{bmatrix}
M_V(t) = 0 \\
M_L(t) > 0 \\
P \geq P^{\text{sat}}(T(t))
\end{bmatrix} \times 
\begin{bmatrix}
M_V(t) > 0 \\
M_L(t) > 0 \\
P = P^{\text{sat}}(T(t))
\end{bmatrix} \times 
\begin{bmatrix}
M_V(t) > 0 \\
M_L(t) = 0 \\
P \leq P^{\text{sat}}(T(t))
\end{bmatrix}
```
Hybrid vs. Nonsmooth

- "Continuous" disjunction:

\[
\begin{bmatrix}
M_V(t) = 0 \\
M_L(t) > 0 \\
P \geq P_{\text{sat}}(T(t))
\end{bmatrix} \otimes \begin{bmatrix}
M_V(t) > 0 \\
M_L(t) > 0 \\
P = P_{\text{sat}}(T(t))
\end{bmatrix} \otimes \begin{bmatrix}
M_V(t) > 0 \\
M_L(t) = 0 \\
P \leq P_{\text{sat}}(T(t))
\end{bmatrix}
\]

- Nonsmooth equation:

\[
0 = \text{mid}\left( M_V(t), P - P_{\text{sat}}(T(t)), -M_L(t) \right)
\]
Nonsmooth Models: Applications

- Intensive properties with flow reversals
- Flow transitions (laminar, turbulent, choked)
- Thermodynamic phase changes
- Crystallization kinetics: growth vs. dissolution
- Flow control devices, diodes
- Irregularities in vessel geometry
- Dynamic flux balance analysis (DFBA) systems
  - e.g., aerobic to anaerobic switch
- Various “elements” of controllers
- Protecting domains of functions (abs)
- Piecewise properties
- etc., etc.
Flow Reversal: Intensive Properties

\[ \frac{dM_i^A}{dt}(t) = -\left( \max\left(F_{\text{out}}^A(t),0\right)x_i^A + \min\left(F_{\text{in}}^B(t),0\right)x_i^B \right) \]

\[ \frac{dM_i^B}{dt}(t) = -\frac{dM_i^A(t)}{dt}; \quad F_{\text{out}}^A(t) = F_{\text{in}}^B(t) \]

\[ F_{\text{out}}^A(t) = c \frac{h^A(t) - h^B(t) + \Delta h(t)}{\sqrt{|h^A(t) - h^B(t) + \Delta h(t)| + \varepsilon}} \]

\[ \begin{align*}
x_1^A, x_1^B & \quad 0 \quad 0.1 \quad 0.2 \quad 0.3 \quad 0.4 \quad 0.5 \quad 0.6 \quad 0.7 \quad 0.8 \quad 0.9 \quad 1 \\
M^A, M^B & \quad 0 \quad 0.5 \quad 1 \quad 1.5 \quad 2 \quad 2.5 \\
t [\text{sec}] & \\
\end{align*} \]
Multi-component Dynamic VLE

Mass and energy balances:

\[
\frac{dM_i}{dt}(t) = F_{in} z_i(t) - F_L(t) x_i(t) - F_V(t) y_i(t)
\]

\[
\frac{dU}{dt}(t) = F_{in} h_m(t) - F_L(t) h_L(t) - F_V(t) h_V(t) + Q(t)
\]

\[
M_i(t) = M_L(t) x_i(t) + M_V(t) y_i(t)
\]

\[
\sum_{i=1}^{n_c} M_i(t) = M_L(t) + M_V(t)
\]

\[
H(t) = M_L(t) h_L(t) + M_V(t) h_V(t)
\]

\[
H(t) = U(t) + P(t)V
\]

Thermodynamic phase equilibrium:

\[
y_i(t) = k_i(t) x_i(t)
\]

\[
0 = \text{mid} \left( \frac{M_V(t)}{M_V(t) + M_L(t)} \sum_{i=1}^{n_c} x_i(t) - \sum_{i=1}^{n_c} y_i(t), \frac{M_V(t)}{M_V(t) + M_L(t)} - 1 \right)
\]

Flow control:

\[
F_V(t) = c_v \min(V_{V}^{\min}, V_V(t)) \max \left( 0, \frac{P(t) - P_0}{\sqrt{|P(t) - P_0| + \varepsilon}} \right)
\]

\[
F_L(t) = c_i \min(V_{L}^{\min}, V_L(t)) \max \left( 0, \frac{K_L(t)}{\sqrt{|K_L(t) + \varepsilon|}} \right)
\]

\[
K_L(t) = g \frac{V_L(t)}{A} + \frac{P(t) - P_0}{\rho_L(t)}
\]
Multi-component Phase Change

\[ M_L = 0 \]
\[ \sum_{i=1}^{n_c} x_i(t) \leq 1 \]

\[ M_V = 0 \]
\[ \sum_{i=1}^{n_c} y_i(t) \leq 1 \]

Crystallization Kinetics

- With the development of continuous crystallization processes, dissolution has to be considered in dynamic models of crystal size distribution:

\[
\frac{\partial (V_n)}{\partial t}(t,z) + K(t) \frac{\partial (V_n)}{\partial z}(t,z) = Q_{in}(t)n_{in}(t,z) - Q_{out}(t)n(t,z)
\]

\[S(t) = \left(x_i(t) - x_i^{sat}\right) / x_i^{sat}, \quad K(t) = \min\left(k_D S(t)\left|S(t)\right|^{n_D - 1}, k_G \left|S(t)\right|^{n_G}\right)\]

- With finite volume discretization of the size coordinate:

\[
\frac{dN_j}{dt}(t) + \frac{1}{\Delta z} \left(G(t)\left(N_j(t) - N_{j-1}(t)\right) + D(t)\left(N_{j+1}(t) - N_j(t)\right)\right) = Q_{in}(t)n_{in,j}(t) - Q_{out}(t)n_j(t), \quad j = 2, \ldots, m - 1
\]

\[
G(t) = \max\left(0, k_G S(t)\left|S(t)\right|^{n_G - 1}\right) \quad \quad N_j \equiv V n_j,
\]

\[
D(t) = \min\left(k_D S(t)\left|S(t)\right|^{n_D - 1}, 0\right) \quad \quad n_j - \text{density of crystals of size } (j - 1)\Delta L < L < j\Delta L
\]
Crystallization Kinetics

- Switching between regimes of positive and negative super-saturation:

\[ S(t) - \text{super-saturation} \]
\[ \varepsilon(t) - \text{volume fraction of solid} \]
\[ T(t) = \text{mid}(30, 0, -40 + |80 + 8t|) \]
Regularization of Nonsmooth DAEs

- Nonsmooth DAEs:
  \[
  \begin{align*}
  \dot{x}(t,p) &= f(t,p,x(t,p),y(t,p)) \\
  0 &= g(t,p,x(t,p),y(t,p)) \\
  x(t_0,p) &= f_0(p)
  \end{align*}
  \]

- \( f \) is piecewise continuous w.r.t. \( t \) and continuous w.r.t. \( p, x, y \)
- \( g \) is locally Lipschitz continuous
- “Index 1” Nonsmooth DAEs: generalized differentiation index one
- Existence, uniqueness, continuous/Lipschitz dependence on parameters, etc.

Given locally Lipschitz continuous \( f : \mathbb{R}^n \to \mathbb{R}^m \):

- Clarke’s Generalized Jacobian

\[
\partial f(x) := \text{conv} \left\{ H : Jf(x_j) \to H, x_j \to x, x_j \in X \setminus Z_f \right\}
\]

Example:

\[
\begin{align*}
\partial f(x) &= \{-1\} \\
\partial f(x) &= \{1\} \\
\partial f(0) &= [-1,1]
\end{align*}
\]

“Index-1” Nonsmooth DAE:
No singular matrix in the set \( \{M : \exists[N M] \in \partial g(t,p,x(t,p),y(t,p))\} \)

- If \( g \) is \( C^1 \):

\[
\left\{ \frac{\partial g}{\partial y} (t,p,x(t,p),y(t,p)) \right\}
\]

Clarke, Optimization and Nonsmooth Analysis, SIAM, 1990.
The “Red-line”
Novartis-MIT Center

Dynamic Optimization in PSE

- **Campaign continuous manufacturing:**
  - Maximize production, minimize off-spec.
  - “Discrete” phenomena: start-up/shut-down, phase changes, crystal growth/dissolution, etc., etc.

Dynamic Optimization of DAEs

In the smooth case:

- Semi-explicit index-1 DAE IVP

\[
\dot{x}(t, p) = f(t, p, x(t, p), y(t, p)) \\
0 = g(t, p, x(t, p), y(t, p)) \\
x(t_0, p) = f_0(p)
\]

- Sensitivity DAEs

\[
\frac{\partial \dot{x}}{\partial p} = \frac{\partial f}{\partial p} + \frac{\partial f}{\partial x} \frac{\partial x}{\partial p} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial p} \\
0 = \frac{\partial g}{\partial p} + \frac{\partial g}{\partial x} \frac{\partial x}{\partial p} + \frac{\partial g}{\partial y} \frac{\partial y}{\partial p} \\
\frac{\partial x}{\partial p}(t_0) = Jf_0(p_0)
\]

Update \( p \) via optimization
In the nonsmooth case:

- Semi-explicit "index-1" DAE IVP
  \[
  \dot{x}(t,p) = f(t,p,x(t,p),y(t,p)) \\
  0 = g(t,p,x(t,p),y(t,p)) \\
  x(t_0,p) = f_0(p)
  \]

- Nonsmooth sensitivity DAEs

Update $\mathbf{p}$ via optimization

### Dynamic Optimization of DAEs
\[ \dot{x}^{(k)} = f^{(k)}(t, p, x^{(k)}), \quad S^{(k)} = \frac{\partial x^{(k)}}{\partial p} \]

\[ S^{(k+1)} - S^{(k)} = - \left[ f^{(k+1)} - f^{(k)} \right] \frac{\partial t}{\partial p} \]

Transversality: \( \frac{\partial g^{(k)}}{\partial x^{(k)}} \dot{x}^{(k)} \neq 0 \)

Nonsmooth DAE Sensitivities: Generalized Derivatives

- Want generalized derivative elements \( \frac{\partial x_{t_f}(p_0)}{\partial p}, \frac{\partial y_{t_f}(p_0)}{\partial p} \)

  - Nonsmooth analog of \( \frac{\partial x}{\partial p}(t_f,p_0), \frac{\partial y}{\partial p}(t_f,p_0) \)
  
  - Difficult to evaluate in general (lack of sharp calculus rules, etc.)

- New tool: lexicographic directional (LD-)derivatives

  - Nonsmooth analog to classical directional derivative

  - Applicable to a wide class of functions (\( C^1, PC^1, \) convex, arbitrary compositions of such, etc.)

  - Satisfies strict calculus rules (e.g. chain rule)

  - Accurate, automatable and computationally cheap method

Lexicographic Differentiation

- $f : X \in \mathbb{R}^n \rightarrow \mathbb{R}^m$ is L-smooth at $x \in X$ if it is locally Lipschitz continuous and directionally differentiable, and if, for any $M := [m_{(1)} \cdots m_{(p)}] \in \mathbb{R}^{n \times p}$, the following functions exist:

  $$f_{x,M}^{(0)} : d \mapsto f'(x;d)$$

  $$f_{x,M}^{(1)} : d \mapsto [f_{x,M}^{(0)}]'(m_{(1)};d)$$

  $$\vdots$$

  $$f_{x,M}^{(p)} : d \mapsto [f_{x,M}^{(p-1)}]'(m_{(p)};d)$$

- If the columns of $M$ span $\mathbb{R}^n$, then $f_{x,M}^{(p)}$ is linear, $L$-derivative:

  $$J_L f(x;M) := Jf_{x,M}^{(p)}(0)$$

- Lexicographic subdifferential:

  $$\partial_L f(x) = \{ J_L f(x;M) : M \in \mathbb{R}^{n \times n}, \det M \neq 0 \}$$

Lexicographic Differentiation

- Systematically probes local derivative information

\[ f(x_1, x_2) = \max(0, \min(x_1, x_2)) \]

\[ f_{0,1}^{(2)}(d) = d_2 \]

\[ J_L f(0, 1) = J f_{0,1}^{(2)}(0) \]

\[ f_{0,1}^{(0)}(d) = \max(0, \min(d_1, d_2)) \]

\[ f_{0,1}^{(1)}(d) = \max(0, d_2) \]
LD-Derivatives

- Given L-smooth $f$ and directions matrix $M := \begin{bmatrix} m_{(1)} & \cdots & m_{(k)} \end{bmatrix}$

$$f'(x; M) := \begin{bmatrix} f_{x, M}^{(0)}(m_{(1)}) & f_{x, M}^{(1)}(m_{(2)}) & \cdots & f_{x, M}^{(k-1)}(m_{(k)}) \end{bmatrix}$$

- If $M$ is square and nonsingular:
  $$f'(x; M) = J_L f(x; M) M$$

- If $f$ is $C^1$ at $x$:
  $$f'(x; M) = Jf(x) M$$

- Sharp LD-derivative chain rule:
  $$\left[ f \circ g \right]'(x; M) = f'(g(x); g'(x; M))$$

Generalized Derivatives Landscape

- **Given** \( f : \mathbb{R}^n \to \mathbb{R}^m \)
  - If \( m = 1 \) (e.g., objective function):
  - If \( f \) is \( \text{PC}^1 \):
    \[ \partial_L f(x) = \partial_B f(x) = \partial f(x) = Jf(x) \]
  - If \( f \) is \( \text{C}^1 \):
    \[ \partial_L f(x) = \partial_B f(x) = \partial f(x) = Jf(x) \]
- LD-derivatives furnish gen. deriv. elements (green dots) in tractable way

In the nonsmooth case:

- Semi-explicit "index-1" DAE IVP
  \[ \dot{x}(t, p) = f(t, p, x(t, p), y(t, p)) \]
  \[ 0 = g(t, p, x(t, p), y(t, p)) \]
  \[ x(t_0, p) = f_0(p) \]

- Nonsmooth sensitivities DAEs
  \[ \dot{X}(t) = [f_t](p_0, z(t, p_0); (M, Z(t))) \]
  \[ 0 = [g_t](p_0, z(t, p_0); (M, Z(t))) \]
  \[ X(t_0) = [f_0](p_0; M) \]

where

\[ z = (x, y), Z = (X, Y) = [z_t](p_0; M) \]

Update \( p \) via optimization

\[ x(t, p_0), \quad J_L x_t(p_0; M), \quad y(t, p_0) \]
\[ J_L y_t(p_0; M) \]

Simple Flash Process: Mode Sequence

- Mode sequence varies under parametric perturbations

\[
\begin{align*}
\dot{H}(t) &= U(T_{out} - T(t)) \\
M &= M_L(t) + M_V(t) \\
H(t) &= Mh_V(t) - M_L(t) \Delta h_{vap}(T(t)) \\
h_v(t) &= C_p(T(t) - T_0) \\
\log(P^{sat}(t)) &= A - B / (T(t) + C) \\
+ & \text{ hybrid automaton}
\end{align*}
\]
Simple Flash Process: Sensitivities

- Nonsmooth sensitivities:

\[
\begin{align*}
\dot{S}_H(t) &= U(1 - S_T(t)) \\
S_H(t) &= MCpS_T(t) - \Delta h_{\text{vap}}'(T(t))S_T(t) \\
0 &= \text{mid}'(M_V(t), P - P_{\text{sat}}(T(t)), -M_L(t); (S_V(t), -P_{\text{sat}}'(T(t))S_T(t), -S_L(t))) \\
S_V(t) &= -S_L(t)
\end{align*}
\]

No Notion of Mode Sequence Needed
The "Red-line" Process Flow Diagram

---

The “Red-line”
Nonsmooth Process Model

No hold-up Liquid-Liquid-Equilibrium:

\[ \text{mid} \left( \frac{F\text{raf}}{F\text{raf} + F\text{pur}} \right) - 1, \sum w_i\text{pur} - \sum w_i\text{raf}, \frac{F\text{raf}^1}{F\text{raf} + F\text{pur}^1} \right) = 0 \]

Flow set point for dilution:

\[ F_S = \max \left( \frac{F\text{raf} w\text{raf}}{w^{s,p}} - F\text{raf}, 0 \right) \]

Crystallization kinetics:

\[ G(t) = \max \left( 0, k_G S(t) |S(t)|^{n_G-1} \right) \]
\[ D(t) = \min \left( k_D S(t) |S(t)|^{n_D-1}, 0 \right) \]

Valves (weirs):

\[ F_{out} = \max \left( 0, u(t) \frac{V\text{Cr}1 - V_{\text{min}}}{\sqrt{V\text{Cr}1 - V_{\text{min}}}} + \epsilon \right) \]

Dynamic Liquid-Liquid-Equilibrium with hold-ups:

\[ \text{mid} \left( \frac{M\text{raf}}{M\text{raf} + M\text{pur}}, -1, \sum w_i\text{pur} - \sum w_i\text{raf}, \frac{M\text{raf}^2}{M\text{raf}^2 + M\text{pur}^3} \right) = 0 \]
The “Red-line”
Nonsmooth Process Simulation
Dynamic Optimization: Problem formulation

Our approach, inspired by the ‘turnpike theory’:

\[
\begin{align*}
\max_{u, t^{on}, t^{off}} \int_{t^{on}}^{t^{off}} J = \text{objective function} \\
\text{s.t.} \quad \dot{x} &= f(x, y, u, t), \quad x(0) = x_0, \quad t \in [0, t^{final}] \\
0 &= g(x, y, u, t) \\
\text{safety constraints} \quad t \in [0, t^{final}] \\
\text{quality constraints} \quad t \in [t^{on}, t^{off}] \\
\text{final time constraints} \quad t = t^{final}
\end{align*}
\]

Comparison to the classical approach:

\[
\dot{x} = f(x, y, u, t) = 0, \quad \forall t \in [t^{on}, t^{off}]
\]

A. M. Sahlodin and P. I. Barton, Optimal Campaign Continuous Manufacturing
Case study: Multi Objective Optimization

\[
\max \quad \text{Yield} = \frac{MW^C_1}{MW^P} \int_{t_{on}}^{t_{off}} F^P_{on-spec} \bigg/ \int_{0}^{t_f} F_{C_1}
\]

\[
\max \quad \text{Productivity} = \int_{t_{on}}^{t_{off}} F^P_{on-spec}
\]
Case study: Pareto Curves

\[
\begin{align*}
\text{max} & \quad \text{Yield} \\
\text{s.t} & \quad x_P > 0.26 \\
& \quad \sum x_i < 0.01 \\
\int_{t_{on}}^{t_{off}} F_{on-spec}^P = M^P & \quad \text{operational constraint}
\end{align*}
\]

Maximize yield

Maximize productivity
Nonsmooth DAEs

◆ Summary of Progress:
  - Possess a strong mathematical theory (recently)
    - Hence, formulate model this way if you can!
  - Easy-to-use and solve and do sensitivity analysis
  - Applicable to variety of operational problems:
  - Numerical toolkit: amenable to computationally tractable (e.g. automatic differentiation) methods
    - See Khan and Barton, *OM&S* 30 (2015)
    - LD-derivative rules for abs, min, max, mid, 2-norm, etc.

◆ Future Work:
  - Numerical implementations
  - “High-index” nonsmooth DAEs
  - Adjoint sensitivities
Acknowledgments

- Peter Stechlinski, Michael Shoham Patrascu, Harry Watson, Kamil Khan, Ali Sahlodin

Funding Sources:
- The Novartis-MIT Center for Continuous Manufacturing
- Natural Sciences and Engineering Research Council of Canada (NSERC)
Flow Transitions

- Transitioning between flow regimes can be modeled using one nonsmooth equation

\[
\text{Nu}(t) = \max(\text{Nu}_l(t), \text{Nu}_r(t))
\]

\[
\text{Nu}_l(t) = 14.5
\]

\[
\text{Nu}_r(t) = \left(\text{Nu}_i \exp \left(\left(\text{Re}(t) - \text{Re}_c \right) / b\right) + \text{Nu}_i^c\right)^c
\]

\[
\text{Nu}_l(t) = 0.023 \text{Re}(t)^{0.8} \text{Pr}^{1/3}
\]

Sensitivities of Nonsmooth DAEs

- DAE Smooth vs. Nonsmooth:
  - Nonsmooth sensitivities:
    \[
    \begin{align*}
    \dot{X}(t) &= [f_t]'(p_0, x(t,p_0), y(t,p_0); (M,X(t), Y(t)) \\
    0 &= [g_t]'(p_0, x(t,p_0), y(t,p_0); (M,X(t), Y(t)) \\
    X(t_0) &= [f_0]'(p_0; M)
    \end{align*}
    \]
  - Non-smooth and nonlinear DAE system
  - Unique solution and unique initialization
  - \( X \) continuous, \( Y \) discontinuous
  - Once solved, obtain generalized derivative elements (sensitivities) via linear equation solve

- Smooth sensitivities:
  \[
  \begin{align*}
  \frac{\partial x}{\partial p}(t_0) &= Jf_0(p_0) \\
  \frac{\partial x}{\partial p}(\cdot, p_0), \frac{\partial y}{\partial p}(\cdot, p_0) &\text{ continuous}
  \end{align*}
  \]
  Linear DAE system
  Unique soln. & init.