Second Order Analysis for Optimal Control Problems with Singular Arcs

J. Frédéric Bonnans
INRIA Saclay and CMAP, Ecole Polytechnique
Joint work with M.S. Aronna (Fond. G. Vargas, Rio de Janeiro)
and A. Kroener (Inria and CMAP, Ecole Polytechnique)

Terry Fest, May 17-21, 2015

Semigroup setting

Framework: reflexive Banach space H (later: Hilbert space).

C_0 **(or strongly continous) semigroup**: Family $T(t)$, for $t \geq 0$, of bounded linear operators such that $T(0) = I$ and

$$T(s+t) = T(s)T(t), \quad s, t \geq 0$$

$$x = \lim_{t \downarrow 0} T(t)x, \quad \text{for all } x \in H.$$

Then (easy) there exists $M \geq 1$, $\omega \geq 0$ such that

$$\|T(t)\| \leq Me^{\omega t} \quad \text{for all } t \geq 0.$$

Infinitesimal generator of a of C_0 semigroup

(Unbounded) linear operator A in H such that

$$Ax = \lim_{t \downarrow 0} \frac{T(t)x - x}{t}$$

with domain the set of x such that the above limit exists.

Characterization of C_0 semigroups

If $\lambda I + A$ is onto with a bounded inverse, we say that λ belongs to the **resolvent set** $\rho(A)$ and denote by $R_\lambda(A) := (\lambda I + A)^{-1}$ the **resolvent**.

Theorem 1. *A linear operator A is the infinitesimal generator of a C_0 semigroup $T(t)$ such that $\|T(t)\| \leq Me^{\omega t}$, iff A is closed with dense domain, and for all $\lambda > \omega$, $\lambda \in \rho(A)$ and*

$$\|R_\lambda(A)^{-n}\| \leq M/(\lambda - \omega)^n, \quad n = 1, 2, \ldots$$

If $M = 1$, $\omega = 0$ we have a contraction semigroup: $\|T(t)\| \leq 1$.

Ref: A. Pazy, Semigroups of linear operators and applications to partial differential equations, Springer, 1983 (with convention $-A$ instead of A).

Differential equations

In the sequel, $T(t)$ denoted by e^{-tA}. If $A \in L(H)$ then

$$e^{-tA} = I - tA + \frac{1}{2}t^2A^2 + \cdots$$

For $f \in L^1(0,T;H)$ consider the differential equation over $(0,T)$:

$$\dot{y} + Ay = f; \quad y(0) = y_0.$$

The mild, or semigroup solution is by the definition

$$y(t) = e^{-tA}y_0 + \int_0^t e^{-(t-s)A}f(s)\mathrm{d}s.$$

Nonlinear differential equations

If $F : H \to H$ we define the solution of

$$\dot{y}(t) + Ay(t) = F(y(t)) + f(t);\ t \in (0,T); \quad y(0) = y_0.$$

by

$$y(t) = e^{-tA}y_0 + \int_0^t e^{-(t-s)A}(F(y(t)) + f(s))\mathrm{d}s$$

whenever is fixed-point equation is well-defined (as is e.g. if F is Lipschitz).

Dual semigroup

If A linear operator in H with domain $D(A)$: its **adjoint** A^* is the linear operator over H^* with domain

$$\{x^* \in H;\ \ \exists y^* \in H^*;\ \langle x^*, Ax\rangle = \langle y^*, x\rangle, \quad \text{for all } x \in D(A)\ \}.$$

If $\lambda \in \rho(A)$ then $R_\lambda(A)^* = R_\lambda(A^*)$.

Theorem 2. *Let A be the infinitesimal generator of a C_0 semigroup e^{-tA}. Then the semigroup $(e^{-tA})^*$ over H^* is C_0 and its generator is A^*.*

This theorem does not hold if H is not reflexive.

Adjoint equation

Consider the direct and adjoint differential equation, where $a \in L(H)$, $f \in L(0,T;H)$, $g \in L(0,T;H^*)$:

$$\dot{z}(t) + Az(t) = az(t) + f(t);\ t \in (0,T); \quad z(0) = z_0.$$

$$-\dot{p}(t) + A^*p(t) = a^*p(t) + g(t);\ t \in (0,T); \quad p(T) = p_T.$$

The semigroup solutions in $C(0,T;H)$ and $C(0,T;H^*)$ are

$$z(t) = e^{-tA} z_T + \int_0^t e^{-(t-s)A}(a^* z(s) + f(s))\mathrm{d}s$$

$$p(t) = e^{-(t-T)A^*} p_T + \int_t^T e^{-(t-s)A^*}(a^* p(s) + g(s))\mathrm{d}s$$

Integration by parts (IBP)

We have that

$$\langle p(T), z(T)\rangle + \int_0^T \langle g(t), z(t)\rangle \mathrm{d}t = \langle p(0), z(0)\rangle + \int_0^T \langle p(t), b(t)\rangle \mathrm{d}t.$$

Application to optimal control:

z solution of linearized state equation

p costate

LHS $=$ directional derivative of cost

RHS $=$ expression of reduced gradient

Another integration by parts formula

Let w be the primitive of $v \in L^1(0,T)$. Then

$$\int_0^T \dot{w}(t)\langle p(t), z(t)\rangle \mathrm{d}t = \big[w(t)\langle p(t), z(t)\rangle\big]_0^T - \int_0^T w(t)\Big(\langle p(t), b(t)\rangle - \langle g(t), z(t)\rangle\Big)\mathrm{d}t$$

The optimal control problem

Here H Hilbert space, $\mathcal{B}_1 \in H$, $\mathcal{B}_2 \in L(H)$. Bilinear state equation

$$\dot{\Psi} + A\Psi = f + u(\mathcal{B}_1 + \mathcal{B}_2\Psi); \quad \Psi(0) = \Psi_0. \tag{1}$$

Cost function

$$J(u, \Psi) := \alpha \int_0^T u(t)\mathrm{d}t + \frac{a_1}{2} \int_0^T \|\Psi(t) - \Psi_d(t)\|_{\mathcal{H}}^2 \mathrm{d}t + \frac{a_2}{2} \|\Psi(T) - \Psi_{dT}\|_{\mathcal{H}}^2; \tag{2}$$

Costate equation

$$-\dot{p} + \mathcal{A}^* p = a_1(\Psi - \Psi_d) + u\mathcal{B}_2^* p; \quad p(T) = a_2(\Psi(T) - \Psi_{dT}(T)). \tag{3}$$

Control set

Control space (scalar)

$$\mathcal{U} := L^2(0,T)$$

Control constraints

$$u_m \le u(t) \le u_M.$$

First order optimality conditions

Solution of state equation $\Psi[u]$

Reduced cost $F(u) := J(u, \Psi[u])$; Reduced gradient (based on IBP)

$$DF(u)v = \int_0^T \langle p(t), \mathcal{B}_1 + \mathcal{B}_2 \Psi(t) \rangle v(t) \mathrm{d}t$$

Assume (for ease of exposition) solution $\hat{u}$ unconstrained, associated state $\hat{\Psi}$ and costate $\hat{p}$: then

$$\langle p(t), \mathcal{B}_1 + \mathcal{B}_2 \hat{\Psi}(t) \rangle = 0 \quad \text{a.e. on } (0, T).$$

Refs on semigroup approach to optimal control

- X. Li, Y. Yao, in LNCIS 75, 1985.
- X. Li, J. Yong, SICOPT 1991, Birkhäuser, 1995.
- H.O. Fattorini in LNCIS 75, 1985, AMO 1987, Elsevier 2005.
- H.O. Fattorini, H. Frankowska MCSS 1991.

Second order optimality conditions

Lagrangian (formally)

$$J(u, \Psi) + \int_0^T \langle p(t), f(t) + u(t)(\mathcal{B}_1 + \mathcal{B}_2\Psi(t)) - \dot{\Psi}(t) - \mathcal{A}\Psi(t)\rangle \mathrm{d}t$$

Second order optimality conditions II

Hessian of reduced cost (formally) assuming $a_1 = 1$ and no final cost

$$Q(v) := \int_0^T \left(\|z(t)\|^2 + v(t)\langle \hat{p}(t), \mathcal{B}_2 z(t)\rangle \right) \mathrm{d}t$$

where $z = z[v]$ solution of linearized equation (formally)

$$\dot{z} + \mathcal{A}z = \hat{u}\mathcal{B}_2 z + v(\mathcal{B}_1 + \mathcal{B}_2\hat{\Psi}); \quad z(0) = 0.$$

Theorem 3. *If $\hat{u}$ local solution then $Q(v) \geq 0$ for any $v \in L^1(0,T)$.*

Goh transform for the linearized system

Set

$$\xi := z - w(\mathcal{B}_1 + \mathcal{B}_2 \hat{\Psi}); \quad w(t) := \int_0^t v(s) \mathrm{d}s$$

Then $\xi(0) = 0$ and formally, with $[\mathcal{A}, \mathcal{B}_2] := \mathcal{A}\mathcal{B}_2 - \mathcal{B}_2\mathcal{A}$:

$$\dot{\xi} + \mathcal{A}\xi = \hat{u}\mathcal{B}_2\xi - w([\mathcal{A}, \mathcal{B}_2]\hat{\Psi} + \mathcal{A}\mathcal{B}_1).$$

Note that v does not appear here !

We assume that $\hat{\Psi} \in \mathrm{dom}([\mathcal{A}, \mathcal{B}_2])$ and that $[\mathcal{A}, \mathcal{B}_2]\hat{\Psi} \in L^\infty(0, T; H)$. Then we can take ξ as semigroup solution of the above equation.

Goh transform in the second variation

We have that $Q(v) = \Omega(w,h)$, where $h = w(T)$ and setting $\mathcal{B} := \mathcal{B}_1 + \mathcal{B}_2\hat{\Psi}$:

$$\Omega := \Omega_t + \Omega_b, \tag{4}$$

$$\Omega_t(w,h) := \int_0^T \|\xi + w\mathcal{B}\|^2_{\mathcal{H}}\, \mathrm{d}t, \tag{5}$$

$$\Omega_b(w,h) := \Omega_b^1(w,h) + \frac{1}{2}\Omega_b^2(w,h) \tag{6}$$

for

$$\Omega_b^1(w,h) := h(\hat{p}_T, \mathcal{B}_2\xi_T)_{\mathcal{H}} + \int_0^T w(t)(\hat{\Psi}(t) - \Psi_d(t), \mathcal{B}_2\xi(t))_{\mathcal{H}}\mathrm{d}t$$
$$- \int_0^T w(t)(\hat{p}(t), [\mathcal{A}, \mathcal{B}_2]\xi(t))_{\mathcal{H}}\mathrm{d}t,$$
$$\Omega_b^2(w,h) := h^2(\hat{p}_T, \mathcal{B}_2^2\hat{\Psi}_T + \mathcal{B}_{2,T}\mathcal{B}_{1,T})_{\mathcal{H}} + \int_0^T w(t)^2(\hat{\Psi}(t) - \Psi_d(t), \mathcal{B}_2^2\hat{\Psi}(t))_{\mathcal{H}}\mathrm{d}t$$
$$+ \int_0^T w(t)^2(\hat{p}(t), [\mathcal{A}, \mathcal{B}_2^2]\hat{\Psi}(t) - [\mathcal{A}, \mathcal{B}_2]\mathcal{B}_1)_{\mathcal{H}}\mathrm{d}t$$
$$- \int_0^T w(t)^2(\hat{p}(t), \mathcal{B}_2^2 f(t) + \mathcal{B}_2\mathcal{A}\mathcal{B}_1 + \hat{u}\mathcal{B}_2\mathcal{B}_1 - \mathcal{B}_2 b_z(t))_{\mathcal{H}}\mathrm{d}t. \quad (7)$$

Need for regularity

For the above expressions to be well-defined we need that

$$[\mathcal{A}, \mathcal{B}_2]\xi \in L^\infty(0, T; H) \quad \text{with} \quad \xi = z - w(\mathcal{B}_1 + \mathcal{B}_2\hat{\Psi})$$

Critical step

$$[\mathcal{A}, \mathcal{B}_2]\xi \in L^\infty(0, T; H)$$

Remember that $\xi(0) = 0$ and

$$\dot{\xi} + \mathcal{A}\xi = \hat{u}\mathcal{B}_2\xi - w([\mathcal{A}, \mathcal{B}_2]\hat{\Psi} + \mathcal{A}\mathcal{B}_1).$$

Again if $[\mathcal{A}, \mathcal{B}_2]\hat{\Psi} \in L^\infty(0, T; H)$ this will follow from specific regularity results.

Second order optimality conditions III

Corollary 1. *If* $\hat{u}$ *local solution then*

$$\Omega(w, h) \geq 0, \quad \textit{for any } (w, h) \in L^2(0, T) \times \mathbb{R}.$$

Proof based on
- continuity of Ω in the $L^2(0, T) \times \mathbb{R}$ topology
- In the limit, w and h independent

Taylor expansion of cost function using w

We have the Taylor expansion where $w(t) := \int_0^t v(s)\mathrm{d}s$:

$$F(\hat{u}+v) = F(\hat{u}) + DF(\hat{u})v + \frac{1}{2}\Omega(w) + o(\|w\|_1^2)$$

Second order sufficient condition: for some $\alpha > 0$:

$$\Omega(w) \geq 2\alpha\|w\|_2^2, \quad \text{for all } w \in L^2(0,T). \qquad (SOSC)$$

Theorem 4. *If $(SOSC)$ holds, then $\hat{u}$ satisfies the weak quadratic growth condition*

$$F(\hat{u}+v) \geq F(\hat{u}) + DF(\hat{u})v + \frac{1}{2}\alpha\|w\|_2^2$$

Heat equation

I: Setting: $\Omega \subset \mathbb{R}^3$ open, bounded, smooth boundary

Heat equation: $b \in H_0^1(\Omega) \cap W^{2,\infty}(\Omega)$, $y_0 \in C(\bar{\Omega}) \cap H_0^1(\Omega)$, $y = y(x,t)$

$$\left\{ \begin{array}{l} \dot{y} - \Delta y = u(t)b(x)y \quad \text{in } Q := \Omega \times [0,T] \\ y = 0 \text{ on } \partial\Omega \times [0,T]; \qquad y(\cdot,0) = y_0. \end{array} \right. \tag{8}$$

Cost function

$$J(u) = \tfrac{1}{2} \int_Q (y(x,t) - y_d(x,t))^2 \mathrm{d}x\mathrm{d}t \tag{9}$$

Semigroup property

We need to study for $\lambda \geq 0$

$$\lambda y - \Delta y = f \in L^2(\Omega).$$

Then integrating by parts (Dirichlet boundary conditions)

$$\lambda \|y\|_2^2 + \int_\Omega |\nabla y(x)|^2 \mathrm{d}x = \int_\Omega y(x) f(x) \mathrm{d}x \leq \|y\|_2 \|f\|_2$$

implying that the heat equation corresponds to a contraction semigroup.

Well-posedness of ξ equation

Here $\mathcal{A} = -\Delta$ with domain $D(\mathcal{A}) := H_0^1(\Omega) \cap H^2(\Omega)$.

We have to compute (cancellation of $b\Delta y$)

$$[-\Delta, b]y = (-\Delta b)y + 2\nabla b \cdot \nabla y.$$

Known regularity result: if $y_0 \in H_0^1(\Omega)$ and $\hat{u} \in L^2(0, T)$ then

$$y \in C(0, T; H_0^1(\Omega)) \quad \Rightarrow \quad [-\Delta, b]y \in C(0, T; L^2(\Omega)).$$

Same analysis gives $[-\Delta, b]\xi \in C(0, T; L^2(\Omega))$.

Schroedinger equation

Here Ω as before and $\Psi(x,t) \in \mathbb{C}$:

$$\dot{\Psi} - i\Delta\Psi = f$$

Semigroup property: consider

$$\lambda\Psi - i\Delta\Psi = f$$

Multiply by $\hat{\Psi}$ (conjugate), integrate over Ω:

$$\lambda\|\Psi\|_2^2 + i\int_\Omega |\nabla\Psi|^2 \mathrm{d}x = \int_\Omega f(x)\Psi(x)\mathrm{d}x$$

Use Cauchy-Schwarz and take real parts: obtain contraction semigroup

Here $\mathcal{A} = -i\Delta$ with domain (complex spaces) $D(\mathcal{A}) := H_0^1(\Omega) \cap H^2(\Omega)$.

We have to compute (cancellation of $b\Delta y$)

$$[-i\Delta, b]y = (-i\Delta b)y + 2i\nabla b \cdot \nabla y.$$

Regularity result: if $y_0 \in H_0^1(\Omega) \cap H^2(\Omega)$ and $\hat{u} \in L^\infty(0,T)$ then

$$y \in C(0,T; H_0^1(\Omega)) \quad \Rightarrow \quad [-i\Delta, b]y \in C(0,T; L^2(\Omega)).$$

Same analysis gives $[-i\Delta, b]\xi \in C(0,T; L^2(\Omega))$.

Numerical experiment I

Do such singular arcs really occur in practice ?

Or is the solution bang-bang ?

Numerical experiment support the existence of singular arcs !

Computations based on the (free software) optimal toolbox

http://bocop.org

Numerical experiment I: Optimal control by the Neumann BC at $x = 0$
$nx = 50$, $nt = 200$, implicit Euler scheme, $y_0 = 1$, $y_d = 0$, $\alpha = 0$.

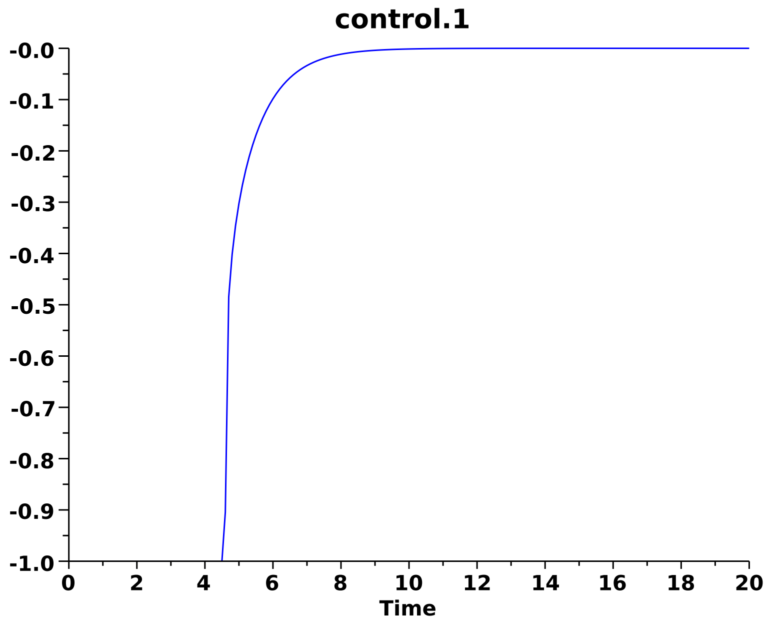

Figure 1: Optimal control: $u \in [-1, 1]$.

A Fuller type phenomenon

- Ad constraint $\dot{u} \in [-1, 1]$

- Infinite dimensional extension of the classical Fuller problem

- The Goh transform should be performed twice, see (5).

- Known chattering phenomenon in Fuller's problem: infinite sequence of bang arcs before entering the singular arc.

- Similar behavior for the control of the heat equation ?

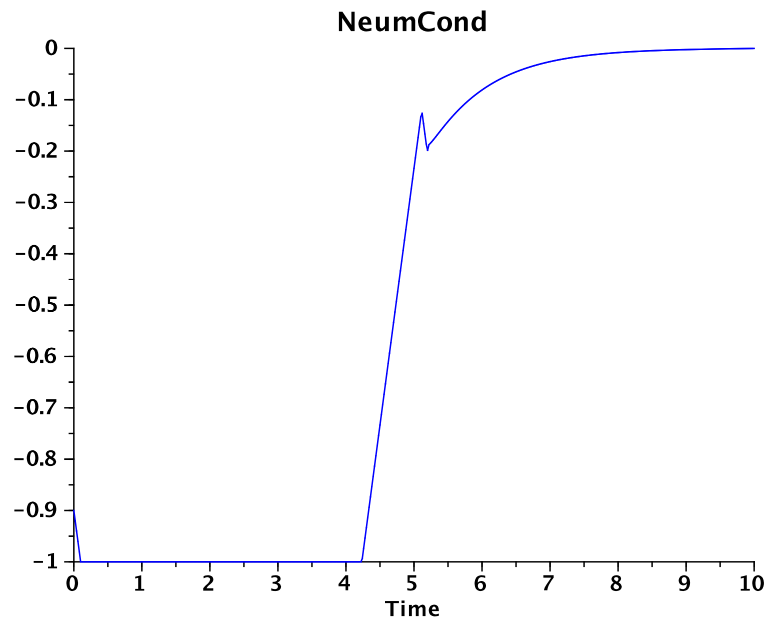

Figure 2: Neumann condition $u \in [-1, 1]$ with bounded derivative.

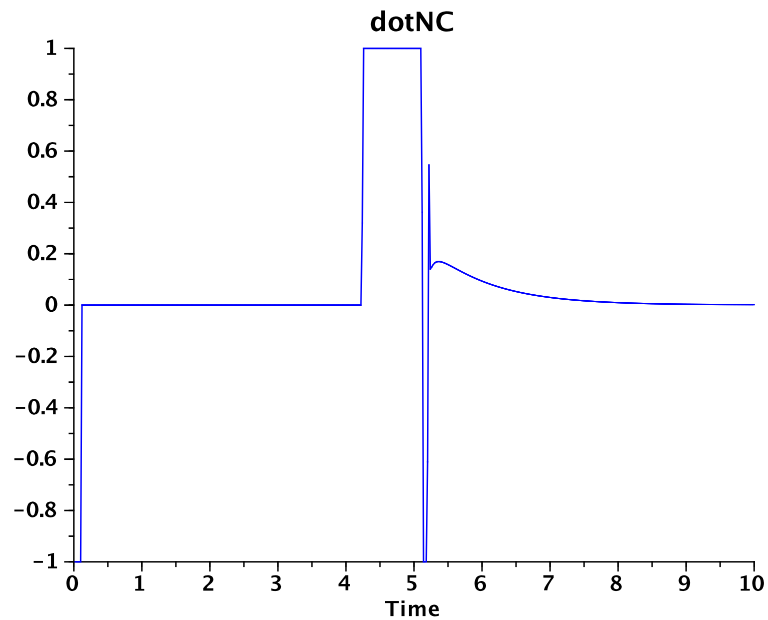

Figure 3: Derivative of the Neumann condition, restricted to $[-1, 1]$.

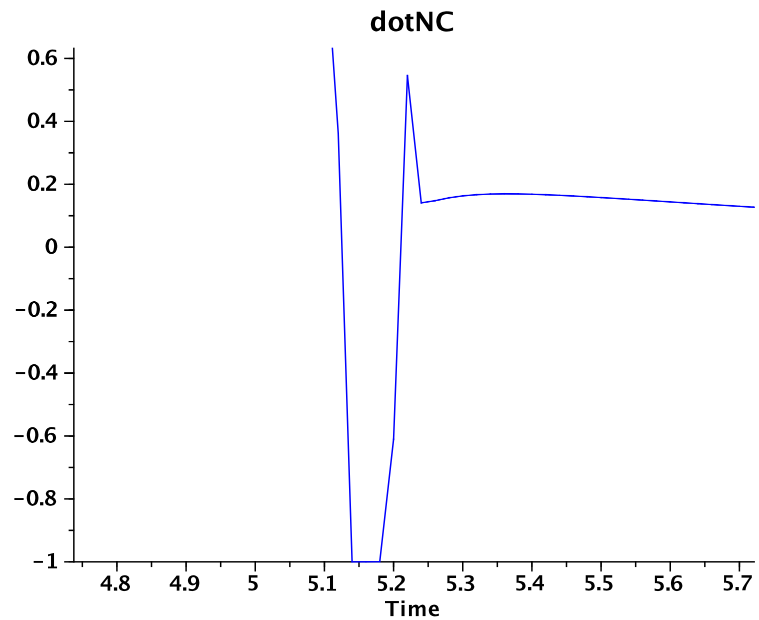

Figure 4: Zoom on the derivative of the Neumann condition.

Singular arc in the Schrödinger equation

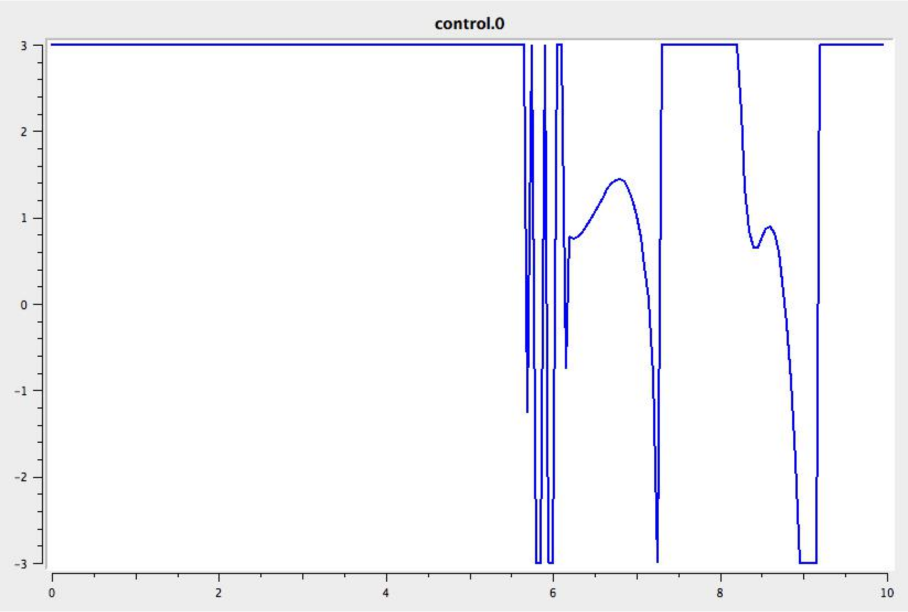

Figure 5: Presence of singular arcs, Schrödinger equation

Questions and comments

- Partial extension of optimality conditions in (1).
  Elliptic case: link with recent work by Casas (4).
- Coefficients of u functions of y ?
- Final constraints, sensitivity analysis ?
- Related article (3).
- Link with the shooting algorithm (2)

References

[1] M.S. Aronna, J.F. Bonnans, A.V. Dmitruk, and P.A. Lotito. Quadratic order conditions for bang-singular extremals. *Numerical Algebra, Control and Optimization, AIMS Journal*, 2(3):511–546, 2012.

[2] J.F. Bonnans. Singular arcs in the optimal control of a parabolic equation. pages 281–292, 2013. Proc 11th IFAC Workshop on Adaptation and Learning in Control and Signal Processing (ALCOSP), Caen, F. Giri ed., July 3-5, 2013.

[3] J.F. Bonnans. Optimal control of a semilinear parabolic equation with singular arcs. *Optim. Methods and Soft.*, 29(2):964–978, 2014.

[4] E. Casas. Second order analysis for bang-bang control problems of PDE. *SIAM J. Control Optim.*, 50(4):2355–2372, 2012.

[5] B.S. Goh. Necessary conditions for singular extremals involving multiple control variables. *SIAM J. Control*, 4:716–731, 1966.