Second-order Analysis for Optimal Control of the Schrödinger Equation

Maria Soledad Aronna[1], Frédéric Bonnans and **Axel Kröner**[2]

[1] Ecole de Mathematica/FGV

[2],[3]INRIA Saclay and CMAP, École Polytechnique

Partial differential equations, optimal design and numerics
Benasque, 01.09.2015

Content

Content

Semigroup setting

Framework: Hilbert space H .

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 there exists $M \geq 1$, $\omega \geq 0$ such that

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

Infinitesimal generator of a C_0 semigroup

(Unbounded) linear operator $\mathcal{A}$ in H such that

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

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

Characterization of C_0 semigroups

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

Theorem

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

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

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 $-\mathcal{A}$ instead of $\mathcal{A}$).

Differential equations

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

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

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

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

The mild, or semigroup solution is by the definition

$$y(t) = e^{-t\mathcal{A}}y_0 + \int_0^t e^{-(t-s)\mathcal{A}} f(s)$$

Nonlinear differential equations

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

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

by

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

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

Dual semigroup

If $\mathcal{A}$ (unbounded) linear operator in H with domain $D(\mathcal{A})$: its dual $\mathcal{A}^*$ is the linear operator over H^* with domain

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

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

Theorem

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

Adjoint equation

Consider the direct and adjoint differential equation, where $a \in L(H)$, $f \in L^1(0,T;H)$, $g \in L^1(0,T;H)$:

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

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

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

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

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

Integration by parts (IBP)

We have that

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

Application to optimal control:

y 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 in $W^{1,1}(0,T)$. Then

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

Content

The optimal control problem

For Hilbert space H, let $\mathcal{B}_1 \in H$, $\mathcal{B}_2 \in L(H)$.

Bilinear state equation

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

Cost function

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

Costate equation

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

Control space (scalar)

$$\mathcal{U} := L^1(0,T)$$

Control constraints

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

Existence

Reduced cost $F(u) := J(u, \Psi[u])$.

The optimal control problem is

$$\text{Min } F(u); \quad u \in \mathcal{U}_{ad}. \tag{P}$$

The *compactness hypothesis* is, for some $s \in [1, \infty]$:

$$\begin{cases} \text{For given } y_0 \in H, \text{ the mapping } f \mapsto y(y_0, f) \\ \text{is compact from } L^s(0,T;H) \text{ to } L^2(0,T;H). \end{cases}$$

Lemma

Let compactness hypothesis hold. Then $u \mapsto \Psi[u]$ is continuous from $\mathcal{U}_\infty$, endowed with the weak$*$ topology, to $C(0,T;H)$ endowed with the weak topology.

Theorem

Let compactness hypothesis hold. Then the optimal control problem has a nonempty set of solutions.

First order optimality conditions

Solution of state equation $\Psi[u]$

Reduced gradient (based on IBP)

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

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

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

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.
- Goldberg, Tröltzsch SICON 1993.

Content

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)

$$Q(v) := \int_0^T \big(a_1\|z(t)\|^2 + v(t)\langle \hat{p}(t), \mathcal{B}_2 z(t)\rangle\big)\,\mathrm{d}t + a_2\|z(T)\|^2,$$

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 (Second Order Necessary Condition)

If $\hat{u}$ local solution then $Q(v) \geq 0$ for any $v \in \mathcal{U}$.

Absence of information in the Legendre-Clebsch condition

Remember that

$$Q(v) := \int_0^T \big(a_1\|z(t)\|^2 + v(t)\langle \hat{p}(t), \mathcal{B}_2 z(t)\rangle\big)\, \mathsf{d}t + a_2\|z(T)\|^2.$$

No quadratic term in the control.

Idea: transformation of the quadratic form: Goh (1966).

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 $[\mathcal{A}, \mathcal{B}_2] := \mathcal{A}\mathcal{B}_2 - \mathcal{B}_2\mathcal{A}$:

$$\dot{\xi} + \mathcal{A}\xi = \hat{u}\mathcal{B}_2\xi - wb_z^1;$$

where 'formally'

$$b_z^1 = -\mathcal{B}_2 f - [\mathcal{A}, \mathcal{B}_2]\hat{\Psi} - \mathcal{A}\mathcal{B}_1$$

Note that v does not appear in the dynamics for ξ !
We assume: $[\mathcal{A}, \mathcal{B}_2]\hat{\Psi} \in L^\infty(0, T; H)$, $\mathcal{A}\mathcal{B}_1 \in H$.

Then we can take ξ as semigroup solution of the above equation.

A general IBP formula

Let $B \in L(\mathcal{H})$ and set $\Phi(t) := By(t)$.

Lemma

Let $[\mathcal{A}, B]y \in L^1(0,T;H)$ *and* $B^*\phi \in D(\mathcal{A}^*)$ *when* $\phi \in D(\mathcal{A}^*)$. *Then*

$$\dot{\Phi} = aBy + [B,a]y + Bb + [\mathcal{A}, B]y$$

in the semigroup sense, and therefore, if $\varphi \in W^{1,1}(0,T)$*:*

$$\int_0^T \dot{\varphi}(t)\langle p(t), \Phi(t)\rangle dt = \Big[\varphi(t)\langle p(t), \Phi(t)\rangle\Big]_0^T - \int_0^T \varphi(t)\Big(\langle p(t), aBy(t) + [B,a]y(t) + Bb + [\mathcal{A}, B]y(t)\rangle - \langle g(t), \Phi(t)\rangle\Big) dt.$$

Hypotheses for the Goh transform

Set $u = \hat{u} + v$ and

$$\Psi = \Psi[u]; \ \ \delta\Psi := \Psi - \hat{\Psi}; \ \ \eta := \delta\Psi - z.$$

Hypotheses:

$$\begin{cases} \mathcal{A}\mathcal{B}_1, \ \ \mathcal{A}\mathcal{B}_2\mathcal{B}_1, \ \ \mathcal{A}\mathcal{B}_2^2\mathcal{B}_1 \ \text{ belong to } \ H, \\ [\mathcal{A}, \mathcal{B}_2]\,, \ \ [\mathcal{A}, \mathcal{B}_2^2], \ \ [\mathcal{A}, \mathcal{B}_2^3] \ \text{ applied to any } \Psi, \, \mathcal{B}_2\hat{\Psi} \text{ belong to } \ L^2(0,T;H). \end{cases}$$

Quadratic form: first step

$Q(v) = Q_1(v) + Q_2(v)$ with setting $\mathcal{B}(t) := \mathcal{B}_1 + \mathcal{B}_2\hat{\Psi}(t)$:

$$\begin{aligned} Q_1(v) &:= a_1 \int_0^T \|z(t)\|^2 \mathrm{d}t + a_2\|z(T)\|^2 \\ &= a_1 \int_0^T \|\xi(t) + w(t)\mathcal{B}(t)\|^2 \mathrm{d}t + a_2\|\xi(T) + w(T)\mathcal{B}(T)\|^2 \end{aligned}$$

$$\begin{aligned} Q_2(v) &:= \int_0^T v(t)\langle \hat{p}(t), \mathcal{B}_2 z(t)\rangle \mathrm{d}t \\ &= \int_0^T v(t)\langle \hat{p}(t), \mathcal{B}_2\xi(t)\rangle \mathrm{d}t + \int_0^T v(t)w(t)\langle \hat{p}(t), \mathcal{B}_2\mathcal{B}(t)\rangle \mathrm{d}t \end{aligned}$$

Quadratic form: first step

We need to use the IBP formula one time for each several time, e.g.:

$$
\begin{aligned}
l_2(w) &= \int_0^T v(t)\langle \hat{p}(t), \mathcal{B}_2\xi(t)\rangle \mathrm{d}t \\
&= w_T(\hat{p}_T, \mathcal{B}_2\xi_T)_H + a_1 \int_0^T w(t)(\hat{\Psi}(t) - \Psi_d(t), \mathcal{B}_2\xi(t))_H \mathrm{d}t \\
&\quad - \int_0^T w(t)^2(\hat{p}(t), \mathcal{B}_2 b_z^1(t))_H \mathrm{d}t - \int_0^T w(t)(\hat{p}(t), [\mathcal{A}, \mathcal{B}_2]\xi(t))_H \mathrm{d}t.
\end{aligned}
$$

Goh transform in the second variation

We have that $Q(v) = \Omega(w,h)$, where $h = w(T)$:

$$\Omega = \Omega_T + \Omega_a + \Omega_b,$$

where

$$\Omega_T := a_2 \|\xi(T) + h\mathcal{B}(T)\|_H^2 + h^2(\hat{p}_T, \mathcal{B}_2\mathcal{B}_1 + \mathcal{B}_2^2\hat{\Psi}_T)_H + h(\hat{p}_T, \mathcal{B}_2\xi_T)_H,$$

$$\Omega_a := \int_0^T \Big(a_1\|\xi\|_H^2 + 2a_1 w(\xi,\mathcal{B})_H + 2a_1 w(\hat{\Psi} - \Psi_d, \mathcal{B}_2\xi)_H - 2w(\hat{p}, [\mathcal{A},\mathcal{B}_2]\xi)_H\Big)\mathrm{d}t,$$

$$\Omega_b := \int_0^T w^2(t)R(t)\mathrm{d}t,$$

with $R \in L^\infty(0,T;H)$ given by

$$R(t) := a_1(\|\mathcal{B}\|_H^2 + (\hat{\Psi} - \Psi_d, \mathcal{B}_2\mathcal{B})_H) + (\hat{p}, r(t))_H,$$

where $r(t) := \mathcal{B}_2^2 f(t) - \mathcal{A}\mathcal{B}_2\mathcal{B}_1 + 2\mathcal{B}_2\mathcal{A}\mathcal{B}_1 - \big[[\mathcal{A},\mathcal{B}_2],\mathcal{B}_2\big]\hat{\Psi}$.

Schrödinger application: $f = 0$, $\mathcal{B}_1 = 0$, then

$$R(t) := a_1(\|\mathcal{B}\|_H^2 + (\hat{\Psi} - \Psi_d, \mathcal{B}_2\mathcal{B})_H) - (\hat{p}, \big[[\mathcal{A},\mathcal{B}_2],\mathcal{B}_2\big]\hat{\Psi})_H$$

Second order optimality conditions III

Corollary

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

Goh-Legendre condition

Lemma

Let $w \mapsto \xi$ *be compact* $L^2(0,T) \to L^2(0,T;H)$. *Then*

$$R(t) \geq 0 \quad \text{a.e.}$$

Taylor expansion of cost function using w

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

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

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

$$\Omega(w,h) \geq 2\alpha\left(h^2 + \|w\|_2^2\right), \quad \text{for all } h \in \mathbb{R},\ w \in L^2(0,T). \quad (SOSC)$$

Theorem

If $(SOSC)$ holds, and $R(t)$ is continuous, then $\hat{u}$ satisfies the weak quadratic growth condition

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

Content

Heat equation

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

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

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

Cost function:

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

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

Schrödinger equation

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

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

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 $b\Delta\Psi$)

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

Regularity result: if $\Psi_0 \in H_0^1(\Omega) \cap H^2(\Omega)$ and $\hat{u} \in L^\infty(0,T)$ then

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

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

Singular arc in the Schrödinger equation

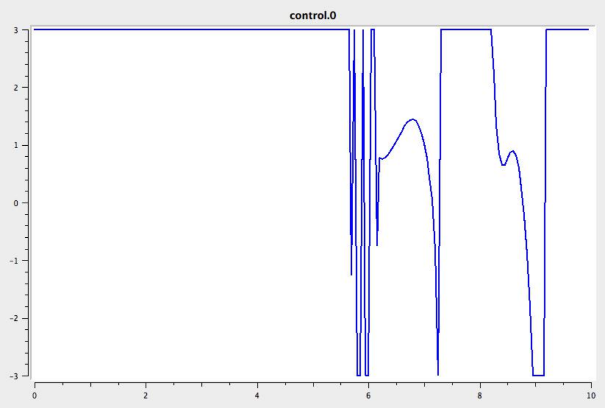

Figure : Presence of singular arcs, Schrödinger equation

Summary

Optimal control of a semigroup.

Control enters cost functional and equation linearly.

Goh-transformation.

Second-order sufficient optimality condition.

Thank you for your attention.