Time Optimal and Minimum Norm Control and Semismooth Newton Method

Kazufumi Ito

Center for Research in Scientific Computation
North Carolina State University
Raleigh, North Carolina

CONTENTS:

- Time Optimal Control
- Semismooth Newton
- Minimum norm solution
- Mixed problem
- Examples– Bilinear Control, Spin-half system, Heat equation

— Karl Kunisch, Qin Zhang

Time Optimal control $f = Ax + F(x) + Bu$

$$\min \tau \quad \text{s.j.} \ \frac{d}{dt}x(t) = f(x,u), \ |u(t)| \le 1, \quad x(\tau) = x_1, \ x(0) = x_0 \in X$$

Assume x_1 is reachable. For $\epsilon > 0$ consider the

$$\min \quad \int_0^\tau (1 + \epsilon\,|u(s)|^2)\,ds$$

Necessary optimality condition

$$\frac{d}{dt}x(t) = f(x,u), \quad u(t) = -\mathrm{sign}_\epsilon(f_u^*\lambda(t)). \quad x(0) = x_0, \ \ x(\tau) = x_1$$
$$-\frac{d}{dt}\lambda(t) = f_x^*\lambda(t)$$
$$1 + \lambda(\tau)\cdot f(x(\tau), u(\tau)) = 0$$

$$\mathrm{sign}_\epsilon(s) = \begin{cases} 1 & s \ge \epsilon \\ \frac{s}{\epsilon} & |s| \le \epsilon \\ -1 & s \le -\epsilon \end{cases}$$

With coordinate change $t = \tau s$, $F(z) = 0$ for $z = (x, \lambda, u, \tau) \in H^1(0,1;R^n) \times H^1(0,1,R^n) \times L^2(0,1,R) \times R$:

$$\frac{d}{dt}x(t) = \tau\, f(x,u)$$

$$u(t) = -\mathsf{sign}_\epsilon(f_u^* \lambda(t),)$$

$$x(0) = x_0, \;\; x(1) = x_1,$$

$$-\frac{d}{dt}\lambda(t) = \tau\, f_x^* \lambda(t),$$

$$1 + \lambda(1) \cdot f(x(1), u(1)) = 0.$$

Semismmoth Newton algorithm

$$z_{k+1} = z_k - J^{-1}F(z_k)$$

where $J\Delta z + F = 0$ is written as

$$\frac{d}{dt}\Delta x = \tau(A\Delta x + B\Delta u) + \Delta\tau(Ax + Bu) + f = 0 \quad \text{in } L^2(0,1,R^n)$$

$$-\frac{d}{dt}\Delta\lambda = \tau A^*\Delta\lambda + \Delta\tau A^*\lambda + g = 0 \quad \text{in } L^2(0,1,R^n)$$

$$\Delta u + G_\epsilon(B^*\lambda)B^*\Delta\lambda + e = 0 \quad \text{in } L^2(0,1,R^n)$$

$$\Delta x(0) = 0, \quad \Delta x(1) + \Delta_1 = 0 \quad \text{in } R^n$$

$$\Delta\lambda(1)\cdot(Ax(1) + Bu(1)) + \lambda(1)\cdot(A\Delta x + B\Delta u) + \Delta_2 = 0 \text{in } R.$$

$$\text{with } G_\epsilon(s) = \begin{cases} 0 & |s| \geq \epsilon \\ \frac{1}{\epsilon} & |s| < \epsilon \end{cases}$$

Lagrange Multiplier Approach to Variational Problems, SIAM book, 08

Semismooth Function: $F : D \subset X \to Z$ is called N-differentiable, if there exists a family of mappings $G : U \to \mathcal{L}(X, Z)$ such that

$$\lim_{|h| \to 0} \frac{|F(x+h) - F(x) - G(x+h)h|_Z}{|h|_X} = 0$$

for all $x \in U$. Moreover, F is semismooth at x if

$$\lim_{t \to 0^+} G(x + t\,h)h \quad \text{exists unifomly in } |h| = 1.$$

Theorem (Local) Suppose F is semismooth at x^* and $|F'(x,h)| \geq \beta\,|h|$ for $\beta > 0$ and all $h \in X$. and assume for a N-derivative G in a neighborhood of x $|G(y)^{-1}| \leq 2\beta$ for all $y \in N(x^*)$. Then the Newton iterates:

$$x^{k+1} = x^k - G(x^k)^{-1}F(x^k)$$

are well-defined and converges to x^* superlineraly in a neighborhood $N(x^*)$.

$$|x^{k+1} - x^*| \leq |G(x^k)^{-1}(F(x^k) - F(x^*) - G(x^k)(x^k - x^*)| \leq o(|x^k - x^*|).$$

Solvability (Linear Case): Let $(\Delta\lambda(1), \Delta\tau)$ be unknown.

$$\Delta\lambda(t) = e^{\tau A^*(1-t)}\Delta\lambda(1) + \int_t^1 e^{\tau A^*(s-t)}(A^*\lambda(s)\Delta\tau + g(s))\,ds$$

$$\Delta x(t) = \int_0^t e^{\tau A(t-s)}(BG_\epsilon(b^*\lambda(s))B^*\Delta\lambda(s) + (Ax + Bu)\Delta\tau + be(s) + g(s))\,ds$$

Thus, we obtain the system of linear equations for $(\Delta\lambda(1), \Delta\tau)$:

$$\begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix} \begin{pmatrix} \Delta\lambda(1) \\ \Delta\tau \end{pmatrix} + q = 0 \qquad (1)$$

where $q \in R^{n+1,1}$ depends on $F = (f, g, e, \Delta_1, \Delta_2)$ linearly. That is, if the matrix $\Phi \in R^{(n+1)\times(n+1)}$ is invertible , then J is bounded invertible.

Let $\{t : |B^*\lambda(t)| < \epsilon\} = \cup_i (t_i, t_{i+1})$ then

$$\Phi_{11} = \frac{1}{\epsilon} \sum_i e^{\tau A(1-t_i)} \mathcal{G}(\Delta t_i) e^{\tau A^*(1-t_i)}$$

$$\mathcal{G}(\Delta) = \int_0^\Delta e^{\tau A s} B B^* e^{\tau A^* s} \, ds$$

Reduced Equations and Backward Shooting method

Linear Case: $\lambda(t)$ is linear equation (independent of x) and u can eliminated by $u = -\text{sign}_\epsilon(B^*\lambda)$. Thus, $x(\cdot)$ is a function of $(\lambda(1), \tau)$ by solving the state equation with $x(0) = x_0$. One can define adjoint-forward shooting method: Find $(\lambda(1), \tau)$ that satisfies

$$\Phi(\lambda(1), \tau) = (x(1) - x_1, 1 + \lambda(1) \cdot (Ax(1) + Bu(1)) = 0$$

Backward Shooting Method We use the scaled transvesariity condition

$$1 + \tau p(t) \cdot (Ax(1) + Bu(1)) = 0$$

In this way one can eliminate $\tau > 0$ by this. So, we may employ the backward shooting method: find $p(1)$ such that $\Phi(p(1)) = x(0) = x_0 = 0$ in which we solve the coupled system with the terminal condition $(x(1), p(1))$ is given.

Theorem u_ϵ converges to u^* with minimum norm. For $0 \leq \epsilon \leq \hat{\epsilon}$

$$\tau_\epsilon \leq \tau_{\hat{\epsilon}}$$

$$\tau_\epsilon \int_0^1 |u_\epsilon|^2 \, dt \geq \tau_{\hat{\epsilon}} \int_0^1 |u_{\hat{\epsilon}}|^2 \, dt$$

$$\tau_0 \leq \tau_\epsilon \leq \tau_0 (1 + \frac{\epsilon}{2})$$

Necessary Optimality

With coordinate change $t = \tau s$,

$$\min \quad J_\epsilon(u,\tau) = \tau \int_0^1 (1 + \frac{\epsilon}{2}|u(t)|^2)\, dt$$

subject to $\frac{d}{dt} = \tau\, f(x,u), \;\; x(0) = x_0$ and

$$u \in U_{ad} = \{u \in L^2(0,1;R^m) : |u(t)| \le 1, \text{ a.e. } \}.$$

Let $g_1(u,\tau) = Cx(1) - y = 0$ (Finite Rank), $g_2(u,\tau) = |x(1) - \bar{x}|^2 \le \delta$

$$\min \quad J_\epsilon(u,\tau) \quad \text{subject to } g_1(u,\tau) = 0,\; g_2(u,\tau) \le 0 \text{ and } u \in U_{ad}.$$

Regular Point Condition: $0 \in \text{int}\{g_u(U_{ad} - u_\epsilon) + g_\tau(R^+ - \tau_\epsilon)\} \quad (2)$

$$\int_0^1 (\epsilon\, u_\epsilon, u - u_\epsilon)\, dt + (g_u(u - u_\epsilon), \mu_\epsilon) \ge 0, \;\; (g_\tau, \mu_\epsilon) = 0 \text{ for all } u \in U_{ad}.$$

$$AdjointEquation: \quad -\frac{d}{dt}p_\epsilon(t) = f_x(x_\epsilon, u_\epsilon)^* p_\epsilon$$

$$(g_\tau, \mu_\epsilon) = \int_0^1 (f(x_\epsilon, u_\epsilon), p_\epsilon)\, dt.$$

$$(g_u(v), \mu_\epsilon) = \tau \int_0^1 (f_u^* p_\epsilon, v)\, ds$$

Thus, the necessary optimality is written as

$$\int_0^1 (\epsilon u_\epsilon + f_u(x_\epsilon, u_\epsilon)^* p_\epsilon, u - u_\epsilon)\, dt \geq 0.$$

for all $u \in U_{ad}$ and

$$\int_0^1 (1 + \frac{\epsilon}{2}|u_\epsilon|^2 + (p_\epsilon, f(x_\epsilon, u_\epsilon)))\, dt = 0$$

Standard Minimum Norm Problems (unconstrained):

$$\min \quad \int_0^\tau |u|^2 \, dt$$

subject to $x(\tau) = x_1$.

$$u^*(t) = B^* e^{A^*(\tau - t)} \mu, \ \ \mu = G(\tau)^{-1}(x_1 - e^{A\tau} x_0).$$

L^∞ Minimum Norm Problem minimize $\sum_{i=1}^m \ \gamma_i^2$ subject to

$$|u_i(t)| \le \gamma_i \text{ on } [0, \tau], \quad \text{and} \quad x(\tau) = x_1.$$

Necessary Optimality:

$$u^* = -\text{sign}_\epsilon (f_u^* \lambda(t)), \quad \text{and} \quad \gamma_i - \int_0^\tau |(f_u^* \lambda(t))_i| \, dt = 0$$

Mixed Problem Min $\tau + \frac{1}{2}|\gamma|^2$

Quantum Control System

$$\frac{d}{dt}\Psi(t) = -iH(t)\Psi(t)$$

with controlled Hamiltonian: $H(t) = H_0 + \sum_{i=1}^{m} \epsilon_i(t)\, H_i.$

Spin-$\frac{1}{2}$ system

$$\frac{d}{dt}\begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} = \begin{pmatrix} 0 & I_z \\ -I_z & 0 \end{pmatrix}\begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} + \epsilon(t)\begin{pmatrix} 0 & I_x \\ -I_x & 0 \end{pmatrix}\begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix}$$

where the Pauli spin matrices are given by

$$I_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad I_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$\Psi(0) = \frac{1}{\sqrt{2}}(1\ 0\ 0\ 1), \quad \frac{1}{2}|\Psi(\tau) - \mathcal{O}|^2 \leq \delta, \ \mathcal{O} = (1\ 0\ 0\ 0).$$

Mixed Problem Min $\tau + |\gamma|^2$

Heat Equation Control

$$\frac{\partial}{\partial t}y(t) = \Delta y(t), \quad \frac{\partial}{\partial \nu}y(t) = Bu(t)$$

Find

$$u \in U_{ad} = \{|u_i(t)| \leq \gamma_i\}$$

that minimizes $|\gamma|^2$ subject to

$$Cy = 0 \quad \text{and} \quad |y|_{L^2} \leq \delta.$$

- Constrained Approximate Controllability?
- Finite rank constrains (for Low frequency Components)
- Nonlinear System – Navier Stokes?

Numerical Method:

$$\frac{x_{k+1}-x_k}{\Delta t} = \tau(A+Bu^k)\frac{x_{k+1}+x_k}{2}$$

$$-\frac{p_{k+1}-p_k}{\Delta t} = \tau(A^*+Bu^k)\frac{p_{k+1}+p_k}{2}$$

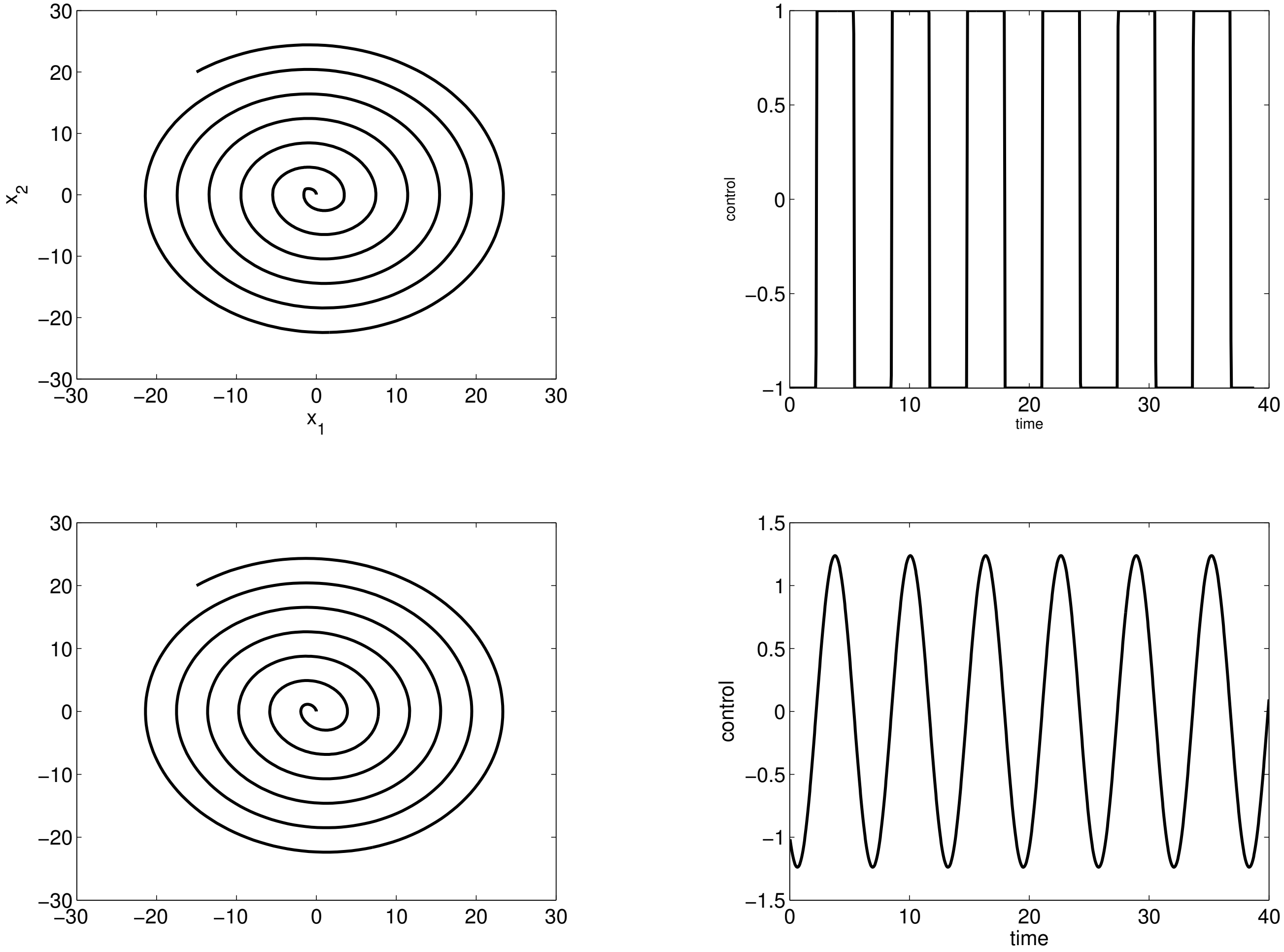
$$u^k = -\mathsf{sign}_\epsilon((B\frac{x_{k+1}+x_k}{2})^*\frac{p_{k+1}+p_k}{2})$$

- Shooting Method with Initialization: Give a guess for τ solve the standard minimum norm problem to obtain an initialization for $p(1)=\mu$

- Continuation Method: Solve the minimum time for x_0. Then, we solve the problem for the neighborhood of x_0 with the initialization of $p(1)$ via the solution for x_0

- γ problem; We use the continuation via τ.

Harmonic Oscillator: $x'' + x = u$, $x_0 = [-15, 20]$ to $x_1 = [0, 0]$



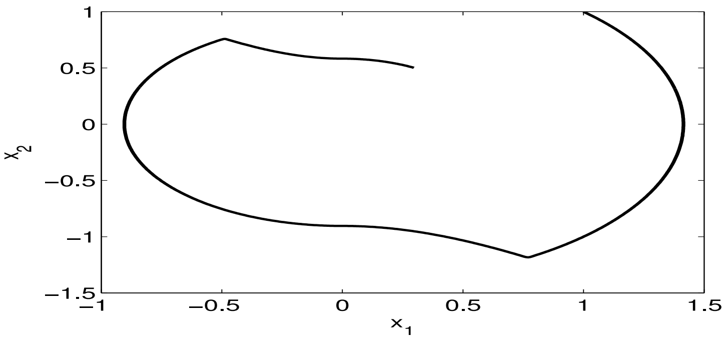
Bilinear Control:

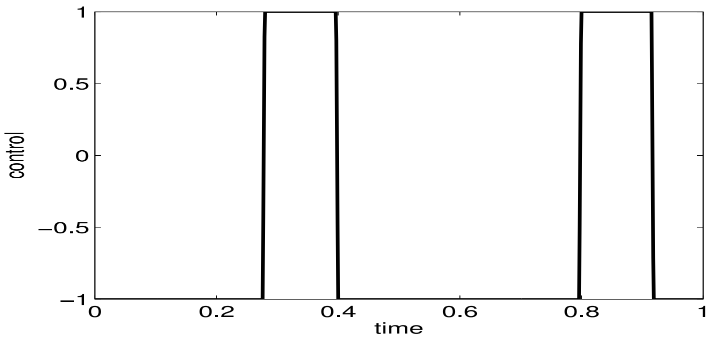
$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{(Hyperbola and Circle)}$$

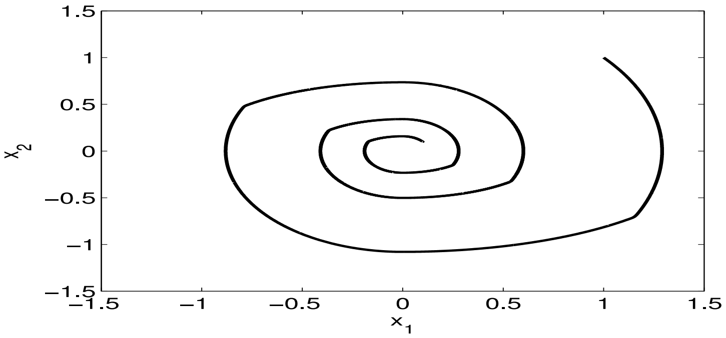
$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 \\ \frac{1}{2} & 0 \end{pmatrix} \quad \text{(Ellipses)}$$

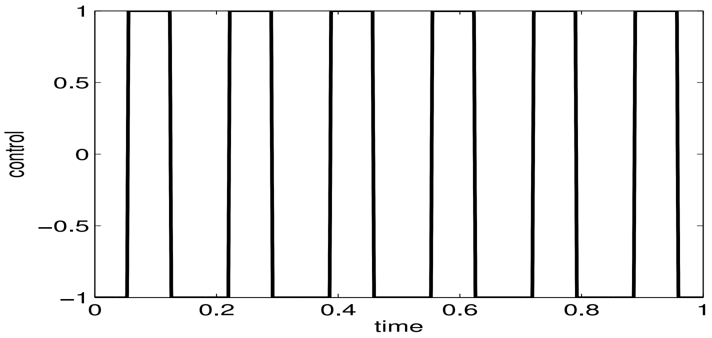
$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{(Sprials)}$$

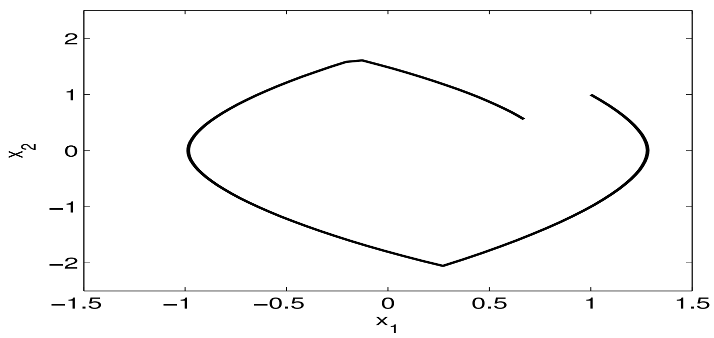
Target	initial guess of τ	optimal τ	adjoint $p(1)$
$(0.5; -0.5)$	1	2.2639	(0.5524; -0.3310)
$(0; -0.5)$	1	3.0178	(0.5524; -0.3310)
$(0; 0.5)$	4	5.9885	(-0.3340; 0.2325)
$(0.7; 0.5)$		3.8114	
$(2; 4)$	4	6.5845	(-0.0290; -0.0180)
$(2; -4)$	2	3.0438	(0.0328; 0.0988)

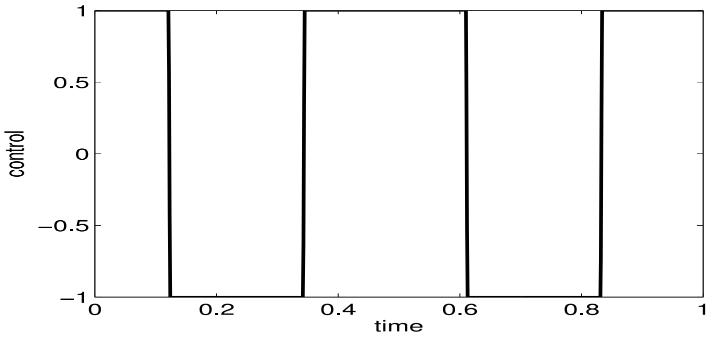

x_2
x_1
1
0.5
0
−0.5
−1
−1.5
−1
−0.5
0
0.5
1
1.5


control
time
1
0.5
0
−0.5
−1
0
0.2
0.4
0.6
0.8
1


x_2
x_1
1.5
1
0.5
0
−0.5
−1
−1.5
−1.5
−1
−0.5
0
0.5
1
1.5


control
time
1
0.5
0
−0.5
−1
0
0.2
0.4
0.6
0.8
1

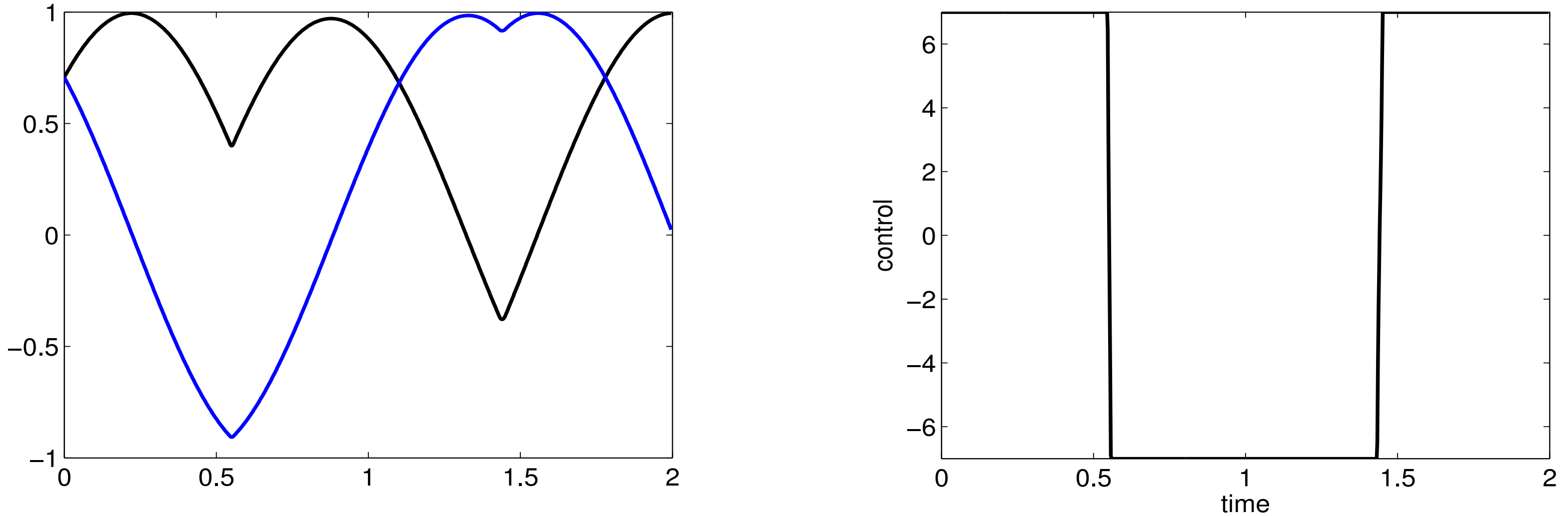

x_2
x_1
2
1
0
−1
−2
−1.5
−1
−0.5
0
0.5
1
1.5


control
time
1
0.5
0
−0.5
−1
0
0.2
0.4
0.6
0.8
1

Spin Half Control

$$p(1) = \mu\mathcal{O}, \quad (\Psi(1), \mathcal{O}) = 1 - \delta.$$

γ	δ	optimal τ	number of switch
15	$0.01/2$	0.9520	0
9	$0.01/2$	1.5465	0
7	$0.01/2$	1.9961	2



1D-Heat Equation Control:

$$p(\tau) = \mu y(\tau), \;\; |y(\tau)|^2_{L^2} - \delta^2 = 0, \;\; \gamma - \int_0^\tau |B^* p| dt = 0$$

Discretization (Fourth Order)

$$Q \frac{d}{dt} y(t) = H y(t) + B u(t)$$

with $B = (0, \cdots, 1) \in R^m$ and

$$Q = \begin{pmatrix} \frac{10}{12} & \frac{1}{12} & & \\ \frac{1}{12} & \frac{10}{12} & \ddots & \\ & \ddots & \ddots & \frac{1}{12} \\ & & \frac{1}{12} & \frac{10}{24} \end{pmatrix}_{m \times m}, \quad H = m^2 \begin{pmatrix} 2 & -1 & & \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}_{m \times m}$$

time horizon τ	time step dt	δ	optimal γ	number of switch
0.5	10^{-4}	0.05	0.4711	0
0.4	10^{-4}	0.05	0.4928	1
0.3	10^{-4}	0.05	0.8730	2
0.2	10^{-4}	0.05	1.8924	3
0.1	10^{-4}	0.05	9.1700	4
0.1	$5 * 10^{-5}$	0.025	13.4338	4