An introduction to optimal control of partial differential equations, Part II

Fredi Tröltzsch

Summer School on Applied Analysis 2016

Chemnitz, 19-23 September 2016

Control of parabolic PDEs

Outline

1. Control of the Schlögl model
   - Traveling waves
   - Optimal control problem
   - An example
   - Necessary optimality conditions
   - Sparse control of the Schlögl model

2. Sparse control of spiral waves – FitzHugh Nagumo equations
   - Analysis of the FitzHugh-Nagumo system
   - Sparse optimal control of the FitzHugh-Nagumo system

A nonlinear heat equation

Now, the time t comes into play. Let us consider a standard uncontrolled 1D semilinear parabolic equation for the temperature $y = y(x,t)$ with initial and boundary conditions.

1D semilinear parabolic problem

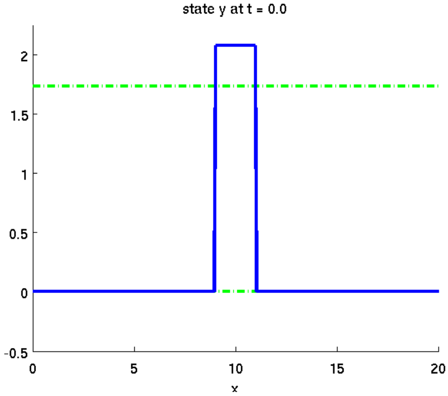
$$\begin{aligned}
&\frac{\partial y}{\partial t} - \frac{\partial^2 y}{\partial x^2} + \frac{1}{3} y^3 = 0 && (x,t) \in (0,L) \times (0,T) \\
&y(x,0) = y_0(x), && x \in (0,L) \\
&\frac{\partial y}{\partial x}(0,t) = 0, \quad \frac{\partial y}{\partial x}(L,t) = 0, && t \in (0,T).
\end{aligned}$$

For $L = 20$, we test the following initial function y_0:

$$y_0(x) = \begin{cases} 1.2\sqrt{3}, & x \in [9,11] \\ 0, & \text{else.} \end{cases}$$

We shall later write $\dfrac{\partial y}{\partial t} =: y_t$

Distribution of heat



The temperature decreases to a small constant value. We do not observe any wave type behavior.

Because the equation is parabolic?

Video: Uncontrolled state function y

Notice: Throughout the handout of the slides, the videos cannot be played, you see only their first snapshot.

A Schlögl model (Nagumo equation)

Consider now a slightly changed semilinear heat equation with the same initial and boundary conditions as above. The term $-y$ essentially changes the behavior!

A special 1D Schlögl model

$$\frac{\partial y}{\partial t} - \frac{\partial^2 y}{\partial x^2} + \frac{1}{3} y^3 - y = 0, \qquad (x,t) \in (0,L) \times (0,T).$$

F. Schlögl,
A characteristic critical quantity in nonequilibrium phase transitions
Z. Phys. B – Condensed Matter (1983).

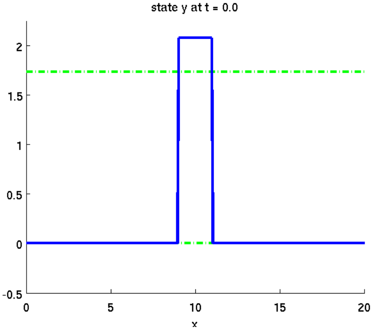
The associated elliptic equation $0 = -y_{xx} + \frac{1}{3} y^3 - y$ has 3 constant solutions (fixed points)

$$y(x) \equiv -\sqrt{3},\ 0,\ \sqrt{3}.$$

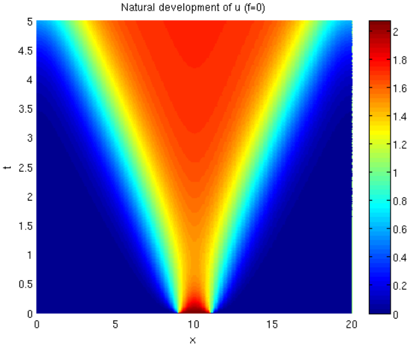
Notice that the nonlinearity $y \mapsto y^3 - y$ is not a monotone function.

Propagating wave fronts

We consider the same initial function y_0 as above for the Schlögl model.



Video: Propagating fronts



A different visualization

The N-dimensional Schlögl model

We consider the equation in $Q := \Omega \times (0, T)$, where $\Omega \subset \mathbb{R}^N$, $N \geq 1$, is a bounded Lipschitz domain with boundary Γ; n is the outward unit normal vector.

$$\begin{aligned} y_t - \Delta y + R(y) &= u && \text{in } Q \\ \partial_n y &= 0 && \text{in } \Gamma \times (0, T) \\ y(x,0) &= y_0(x) && \text{in } \Omega \end{aligned}$$

with reaction term

$$R = \rho\,(y - y_1)(y - y_2)(y - y_3), \quad \rho > 0, \quad y_i \in \mathbb{R}.$$

In the next numerical examples we have $N = 1$, $\Omega = (0, L)$, and $R = \frac{1}{3} y^3 - y = \frac{1}{3}(y + \sqrt{3})\, y\, (y - \sqrt{3})$.

Important property

There is some $R_0 \in \mathbb{R}$ such that

$$R'(y) \geq R_0 \quad \forall y \in \mathbb{R}.$$

Weak solutions of the Schlögl model

$\rightarrow W_2^{0,1}(Q),\ W_2^{1,1}(Q)$

Definition (Weak solution)

A function $y \in W_2^{0,1}(Q) \cap L^\infty(Q)$ is said to be a weak solution of the Schlögl model above, if

$$-\int_Q y\, v_t\, dxdt + \int_Q \nabla_x y \cdot \nabla_x v\, dxdt + \int_Q R(y) v\, dxdt = \int_\Omega y_0\, v(0)\, dx + \int_Q u\, v\, dxdt$$

holds for all $v \in W_2^{1,1}(Q)$ with $v(T) = 0$.

The existence and uniqueness of a unique weak solution can be shown.

However, this concept does not yet fit to the needs of optimal control. Here, the test function must belong to $W_2^{1,1}(Q)$. Later, an adjoint state must be inserted as test function. And this adjoint state only has the same regularity as y.

Fortunately enough, one can prove that y and p belong to $W(0,T)$.

Well-posedness of the Schlögl model

Definition ($W(0,T)$)

$$W(0,T) = \{y \in L^2(0,T;H^1(\Omega)) : \frac{\partial y}{\partial t} \in L^2(0,T;H^1(\Omega)')\},$$

where $\partial y/\partial t$ is defined in the sense of vector-valued distributions.

Theorem (Existence and uniqueness)

- *To each control $u \in L^p(Q)$, $p > N/2+1$, there exists a unique weak solution $y_u \in W(0,T) \cap L^\infty(Q)$ that obeys $y_u \in C(\bar{\Omega} \times (0,T])$.*
- *If y_0 is continuous in $\bar{\Omega}$, then the solution y_u is continuous in $\overline{Q}$.*
- *The mapping $G : u \mapsto y_u$ is of class C^∞.*

Main idea of the proof

It holds

$$R'(y) \geq R_0.$$

We take $\eta > |R_0|$, perform the well known transformation $y(x,t) = e^{\eta t} v(x,t)$

$$\Rightarrow \qquad \frac{\partial}{\partial t}(e^{\eta t} v(x,t)) = \eta\, e^{\eta t} v(x,t) + e^{\eta t} \frac{\partial}{\partial t} v(x,t)$$

and we get the equation

$$v_t - \Delta v + \underbrace{e^{-\eta t} R(e^{\eta t} v) + \eta\, v}_{monotone} = e^{-\eta t} u.$$

This is an equation with a monotone nonlinearity. Now we follow E. Casas, SICON 1998, or J.P. Raymond and H. Zidani, SICON 1999; or my AMS-book, Thm. 5.5

Outline

Optimal control problem

$$\min J(y,u) := \frac{1}{2}\int_Q (y(x,t) - y_Q(x,t))^2\,dxdt + \frac{\nu}{2}\int_Q u^2(x,t)\,dxdt$$

($\nu > 0$ fixed) subject to

$$\begin{aligned} y_t - \Delta y + R(y) &= u \\ \partial_n y &= 0 \\ y(\cdot,0) &= y_0, \end{aligned}$$

$$u \in U_{ad} := \{u \in L^2(Q) : \alpha \le u(x,t) \le \beta \quad \text{for a.a. } (x,t) \in Q\}.$$

This problem has at least one optimal solution.

Notice

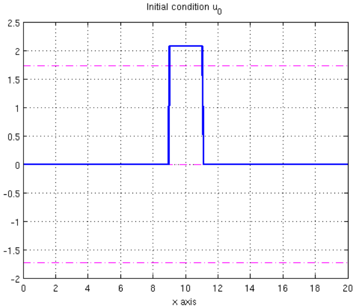
Without control constraints, for $\nu = 0$ the existence of an optimal solution is not guaranteed. If $\nu > 0$, this does not happen provided that $N \le 2$. Otherwise we need control bounds. In computations, we take $\nu \sim 10^{-5} \ldots 10^{-8}$.

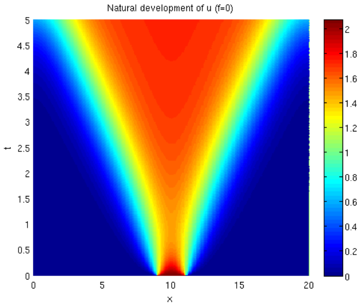
Outline

1. Control of the Schlögl model
   - Traveling waves
   - Optimal control problem
   - An example
   - Necessary optimality conditions
   - Sparse control of the Schlögl model
2. Sparse control of spiral waves – FitzHugh Nagumo equations
   - Analysis of the FitzHugh-Nagumo system
   - Sparse optimal control of the FitzHugh-Nagumo system

Uncontrolled, i.e. for $u = 0$, the nonnegative initial function y_0 below generates the "natural" propagating front y_{nat} shown in the right-hand side.

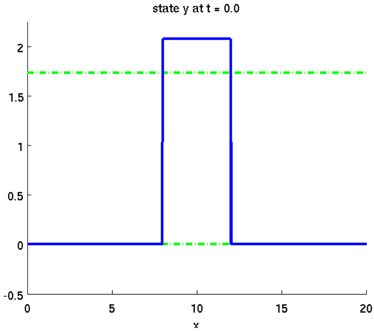
$$y_0(x) = \begin{cases} 1.2\sqrt{3}, & x \in [9, 11] \\ 0, & \text{else.} \end{cases}$$



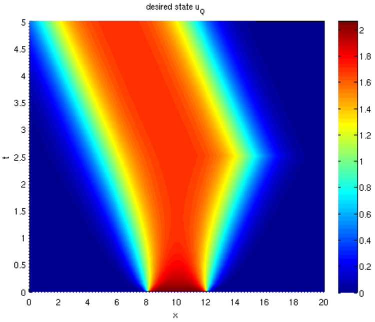


Re-routing

Our goal is to re-route this expanding wave front.



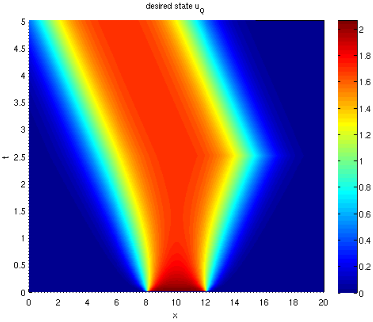
Video: Desired front y_Q



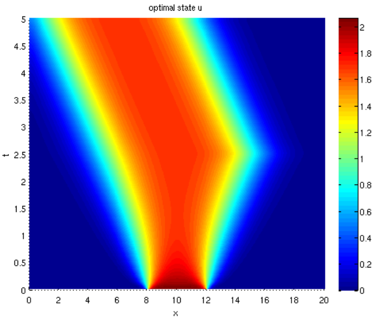
Different visualization

Re-routing of a propagating front

We applied the nonlinear cg method by Hestenes and Stiefel with the Hager-Zhang step-size rule.



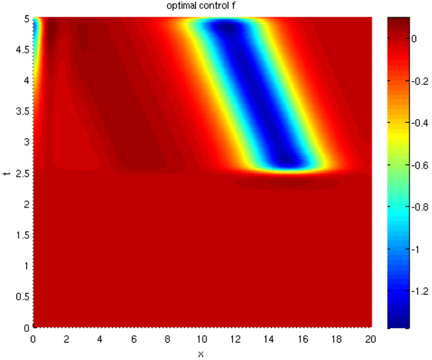
Desired state y_Q



Optimal state $\bar{y}$

Associated optimal control

The control is acting just in the places, where the traveling wave front has to be pushed. This somehow confirms our intuition.



Optimal control $\bar{u}$

To implement the nonlinear cg method, we needed the reduced gradient, i.e. the gradient of the reduced functional. Let us discuss now the theory for this.

Outline

1. Control of the Schlögl model
   - Traveling waves
   - Optimal control problem
   - An example
   - Necessary optimality conditions
   - Sparse control of the Schlögl model
2. Sparse control of spiral waves – FitzHugh Nagumo equations
   - Analysis of the FitzHugh-Nagumo system
   - Sparse optimal control of the FitzHugh-Nagumo system

The Lagrangian function

Now, the equation is parabolic. It is not obvious, how the adjoint equation should look like. Let us employ our formal Lagrangian technique to derive it.

Definition (Lagrangian function)

$$\begin{aligned}\mathcal{L}(y,u,p) \quad := \quad & J(y,u) - \int_Q (y_t - \Delta y + R(y) - u)\, p \, dxdt \\ & - \int_\Sigma \partial_n y \, p_s \, dsdt - \int_\Omega (y(0) - y_0) p_0 \, dx\end{aligned}$$

with multiplier functions p, p_s, p_0.

The adjoint equation is obtained by $\partial_y \mathcal{L}(y,u,p) = 0$, i.e. by

$$\partial_y \mathcal{L}(y,u,p)\, v = 0 \quad {}''\forall'' \; v.$$

Here, we vary with respect to all "sufficiently smooth" v.

Derivation of the adjoint equation

$$
\begin{aligned}
0 = \partial_y \mathcal{L}(y,u,p)\, v \;\; &= \;\; \int_Q (y - y_Q)\, v\, dx - \int_Q (v_t - \Delta v + R'(y)\, v)\, p\, dxdt \\
&\quad - \int_\Sigma \partial_n v\, p_s\, dsdt - \int_\Omega v(0)\, p_0\, dx.
\end{aligned}
$$

First, we perform an integration by parts w.r. to t and x

$$
\begin{aligned}
0 \;\; &= \;\; \int_Q (y - y_Q)\, v\, dx - \int_\Omega v(T)p(T) - v(0)\, p(0)\, dx + \int_\Sigma (p\, \partial_n v - v\, \partial_n p) dsdx \\
&\quad - \int_Q v\, (-p_t - \Delta p + R'(y)p)\, dxdt - \int_\Omega v(0) p_0\, dx - \int_\Sigma \partial_n v\, p_s\, dsdt.
\end{aligned}
$$

Now, we vary w.r. to all v with $\partial_n v = v = 0$ on Σ and $v(0) = v(T) = 0$ in Ω,

$$
0 = \int_Q \left[y - y_Q - (-p_t - \Delta p + R'(y)p) \right] v\, dxdt \quad \forall\, v,
$$

hence

$$
-p_t - \Delta p + R'(y)p = y - y_Q \;\; \text{in } Q.
$$

This is the adjoint partial differential equation.

Derivation of the adjoint equation

Now, we allow also v(T) to vary freely, while $\partial_n v = v = 0$ on Σ and $v(0) = 0$ are still required. This gives

$$p(T) = 0.$$

Next, also $v(0)$ can vary freely,

$$0 = \int_\Omega v(0)(p_0 - p(0))\,dx \Rightarrow p_0 = p(0).$$

Now, also v is allowed to be arbitrary on Σ, hence

$$\partial_n p = 0 \ \text{ on } \Sigma.$$

Finally, $\partial_n v$ is not required to vanish. We find

$$p_s = p|_\Sigma.$$

Adjoint equation and reduced gradient

Summarizing, we have obtained the

Adjoint equation

$$\begin{aligned} -p_t - \Delta\, p + R'(y)\, p &= y - y_Q \\ \partial_n p &= 0 \\ p(\cdot, T) &= 0. \end{aligned}$$

This is a well-posed backward parabolic equation (it can be transformed to a standard forward equation by the transformation $\tilde{t} := T - t$).

Analogously to the elliptic control problem we set:

Definition (Reduced objective functional)

$$f(u) := \frac{1}{2} \int_Q |y_u - y_Q|^2 + \frac{\nu}{2} u^2 \, dxdt.$$

The mapping $G : u \mapsto y_u$ is Fréchet differentiable in $L^\infty(Q)$, the same holds true for f. What is the expression for the derivative?

The reduced gradient

Again, completely analogous to the elliptic case, one can show that

$$f'(u)v = \int_Q (y_u - y_Q)\, z\, dxdt + \int_Q u\, v\, dxdt,$$

where z is the solution of the linearized state equation

$$\begin{aligned} z_t - \Delta z + R'(y)\, z &= v \\ \partial_n z &= 0 \\ z(\cdot, 0) &= 0. \end{aligned}$$

This is a quite implicit representation of $f'(u)$. To tickle out the increment v in the first term, we need the adjoint state p. After some work, we find

$$f'(u)v = \int_Q (p(x,t) + \nu u(x,t))\, v(x,t)\, dxdt.$$

Reduced gradient

It turns out that the linear functional $f'(u)$ can be extended continuously from $L^\infty(Q)$ to the Hilbert space $L^2(Q)$. Thanks to the Riesz representation theorem, we know that $f'(u)$ can be identified with a function of $L^2(Q)$. This function is the reduced gradient. The representation from the last slide shows

Reduced gradient

The reduced gradient $f'(u)$ is given by the function $d \in L^2(Q)$,

$$d(x,t) := p(x,t) + \nu u(x,t).$$

In the space $L^\infty(Q)$, located in the point u, the reduced gradient points in the direction of steepest ascent of f. Hence $-d$ points in the direction of steepest descent.

The gradient method

We briefly sketch the gradient method for the unconstrained case, i.e. for

$$U_{ad} = L^\infty(Q).$$

Continuous version of the gradient method:

1. Set $k = 0$; fix $\varepsilon > 0$, fix $u_0 \in L^\infty(Q)$.
2. Solve the Schlögl equation to find y_k
3. Insert $y := y_k$ in the adjoint equation and compute p_k
4. Set $d_k := -(p_k + \nu\, u_k)$
5. Find a suitable stepsize $s_k > 0$
6. New iterate

$$u_{k+1}(x,t) = u_k(x,t) - s_k\,[p_k(x,t) + \nu\, u_k(x,t)].$$

7. If $\|d_k\|_{L^2(Q)} < \varepsilon$, then STOP
8. $k := k + 1$, goto (1)

Necessary optimality conditions for $\bar{u}$

In principle, the structure of the necessary optimality conditions is the same as for the case of the semilinear elliptic control problem. But the associated analysis is more demanding. We just state them without proof.

Theorem

Let $\bar{u}$ be optimal for the control problem above. Then there exists an adjoint state $\bar{p} \in W(0,T) \cap L^\infty(Q)$ (solving the adjoint equation above with $\bar{y} = y_{\bar{u}}$ inserted) such that the variational inequality

$$\int_Q (\bar{p}(x,t) + \nu\,\bar{u}(x,t))(u(x,t) - \bar{u}(x,t))\,dxdt \geq 0 \quad \forall u \in U_{ad}$$

is satisfied.

Outline

1. Control of the Schlögl model
   - Traveling waves
   - Optimal control problem
   - An example
   - Necessary optimality conditions
   - Sparse control of the Schlögl model

2. Sparse control of spiral waves – FitzHugh Nagumo equations
   - Analysis of the FitzHugh-Nagumo system
   - Sparse optimal control of the FitzHugh-Nagumo system

Recall the Schlögl model

We consider again the

State equation

$$
\begin{aligned}
y_t - \Delta y + R(y) &= u && \text{in } Q \\
\partial_n y &= 0 && \text{in } \Gamma \times (0, T) \\
y(x, 0) &= y_0(x) && \text{in } \Omega.
\end{aligned}
$$

with

$$R(y) = \rho\,(y - y_1)\,(y - y_2)\,(y - y_3).$$

From now on, we have $\Omega \subset \mathbb{R}^N$, $N \leq 3$.

Optimal control problem

$$\min_{u \in U_{ad}} \quad J(u) := I(u) + \mu\, j(u)$$

where

$I(u) :=$ quadratic tracking type functional (next slide)

$j(u) := \int_Q |u(x,t)|\, dxdt = \|u\|_{L^1(Q)}, \qquad \mu \geq 0,$ notice: nondifferentiable!

$U_{ad} := \{u \in L^\infty(Q) \mid u(x,t) \in [\alpha, \beta] \text{ for a.a. } (x,t) \in Q\}$

Assume $\alpha < 0 < \beta$. Special case: $\alpha = -\beta$, i.e. $|u(x,t)| \leq \beta$.

Sparsity

By the term $\mu\, j(u)$, the optimal control becomes sparse. The larger μ is, the smaller is the support of the optimal control $\bar{u}$.

A standard tracking type functional

$$I(u) := \frac{1}{2}\int_Q c_Q(x,t)\,\left(y_u(x,t) - y_Q(x,t)\right)^2\,dxdt$$
$$+\frac{\nu}{2}\int_Q u^2(x,t)\,dxdt \qquad \text{with } \nu > 0$$

and a nonnegative bounded coefficient function c_Q.

They are chosen positive (say $= 1$) in the regions, where we are interested in approaching y_Q or y_T and zero in the other regions.

Some references on sparse controls

G. Stadler
Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices,
Computational Optimization and Applications (2009)

E. Casas, R. Herzog, G. Wachsmuth
Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional,
SIAM Journal on Optimization (2012)

R. Herzog, G. Stadler, G. Wachsmuth
Directional sparsity in optimal control of partial differential equations,
SIAM Journal on Control and Optimization (2012)

E. Casas, C. Clason, K. Kunisch
Approximation of elliptic control problems in measure spaces with sparse solutions,
SIAM Journal on Control and Optimization (2012)

E. Casas, R. Herzog, G. Wachsmuth
Analysis of sparse optimal control problems of semilinear parabolic equations,
In preparation

E. Casas, C. Clason, K. Kunisch
Parabolic control problems in measure spaces with sparse solutions,
SIAM Journal on Control and Optimization (2013)

Sparsity in a nutshell

Explanation for a very simple optimization problem in $\mathbb{R}$.

Optimization problem

$$\min \left\{ \frac{1}{2}(u - y_d)^2 + \frac{\nu}{2}\, u^2 + \mu\, |u| \right\} \quad \text{subject to} \; -1 \le u \le 1.$$

The problem has a unique solution $\bar{u}$. Assume that μ is large and $\bar{u} > 0$. Then

$$|\bar{u}| = \bar{u} \quad \Rightarrow \quad |\bar{u}|' = 1.$$

Variational inequality

$$[\bar{u} - y_d + \nu\, \bar{u} + \mu](u - \bar{u}) \ge 0 \quad \forall u \in [-1, 1],$$

hence

$$\bar{u} = -1, \quad \text{if} \quad \bar{u} - y_d + \nu\, \bar{u} + \mu > 0, \quad \text{i.e. if } \mu \text{ is large enough.}$$

Therefore $\bar{u} = -1$ contradicting $\bar{u} > 0$. Analogously, we cannot have $\bar{u} < 0$, hence $\bar{u} = 0$ follows for all sufficiently large μ.

The subdifferential

Assume that U is a real Banach space with dual space U' (space of all linear and continuous functionals on U), and let $\phi : U \to \overline{\mathbb{R}}$ is a proper convex functional.

Definition (Subdifferential)

Let $u \in U$ be fixed. The subdifferential $\partial\phi(u) \subset U'$ is the set

$$\partial\phi(u) = \{\lambda \in U' : \phi(v) \geq \phi(u) + \langle \lambda \, , \, v - u \rangle_{U',U} \qquad \forall v \in U.$$

Example 1:

$U = U' = \mathbb{R}$, $\phi : u \mapsto |u|$:

$$\partial\phi(u) = \begin{cases} \{1\}, & u > 0 \\ [-1, 1], & u = 0 \\ \{-1\}, & u < 0. \end{cases}$$

The subdifferential

Example 2:

$U = L^1(Q)$, $U' = L^\infty(Q)$, $\phi : u \mapsto \|u\|_{L^1(Q)}$:

$$\partial\phi(u) = \{\lambda \in L^\infty(Q) \text{ satisfying a.e. the conditions below}\}$$

$$\lambda(x,t) \in \begin{cases} \{1\}, & u(x,t) > 0 \\ [-1,1], & u(x,t) = 0 \\ \{-1\} & u(x,t) < 0. \end{cases}$$

Then

$$\int_Q |v(x,t)|\,dxdt \geq \int_Q |u(x,t)|\,dxdt + \int_Q \lambda(x,t)(v(x,t) - u(x,t))\,dxdt \quad \forall v \in L^1(Q).$$

Necessary optimality conditions

In what follows, let us denote the adjoint state by φ instead of p. This even improves the forme parallel use of the integrability index p and the adjoint state $p(\cdot)$.

Theorem (Necessary optimality conditions)

If $\bar{u}$ is a local solution to the optimal control problem, then there exists a unique adjoint state $\bar{\varphi} \in W(0,T)$ such that, with $\bar{\lambda} \in \partial j(\bar{u}) \subset L^\infty(Q)$,

$$\int_Q \big(\bar{\varphi}(x,t) + \nu\,\bar{u}(x,t) + \mu\,\bar{\lambda}(x,t)\big)\,(u(x,t) - \bar{u}(x,t))\,dxdt \geq 0 \qquad \forall u \in U_{ad}.$$

Sparsity

Theorem

For almost all $(x,t) \in Q$*,*

$$\begin{aligned} \bar{u}(x,t) &= 0 \;\; \textit{iff} \;\; |\bar{\varphi}(x,t)| \leq \mu, \\ \bar{\lambda}(x,t) &= \mathbb{P}_{[-1,1]}\left\{-\frac{1}{\mu}\bar{\varphi}(x,t)\right\}. \end{aligned}$$

Proof: Long and tricky.

- The first relation expresses the effects of sparsity.
- The second is used for updating the conjugate gradient in the nonlinear projected cg method. This is remarkable, since $\bar{\lambda}$ is uniquely determined here.

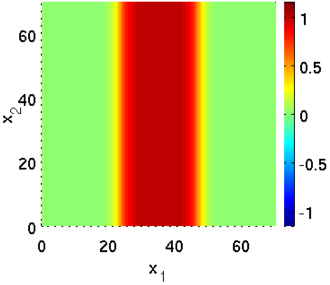
Example: Turning a wave front

- $R(y) = y(y - 0.25)(y + 1)$
- $\Omega = (0,70) \times (0,70)$
- 141×141 node points in Ω
- $y_0(x) := \left(1 + \exp\left(\frac{\frac{70}{3} - x_1}{\sqrt{2}}\right)\right)^{-1} + \left(1 + \exp\left(\frac{x_1 - \frac{140}{3}}{\sqrt{2}}\right)\right)^{-1} - 1$

Example and Computations:

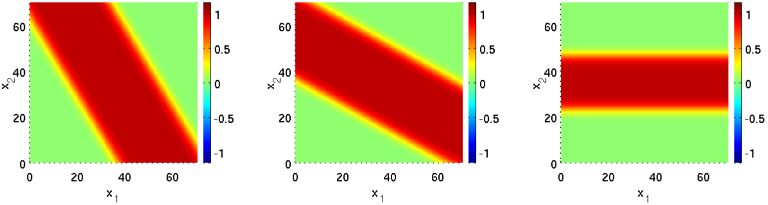
Christopher Ryll (TU Berlin)

Uncontrolled, the wave fronts expand in left and right x_1-direction and cover the whole spatial domain after $t \sim 65$.



Initial state y_0.

Desired trajectory y_Q at time instants

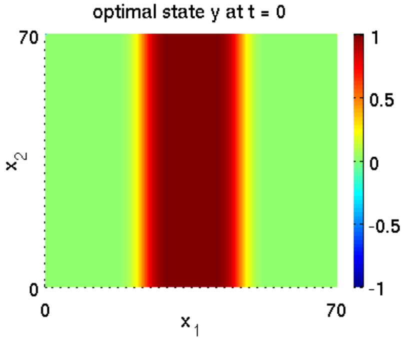


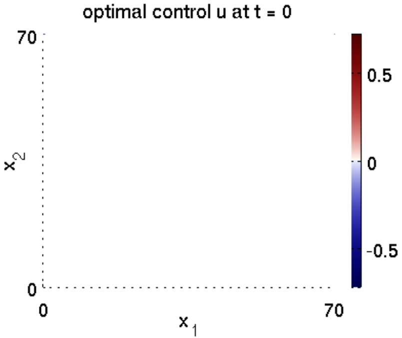
$t = 20$ $\quad$ $t = 40$ $\quad$ $t = 60$

This desired turning trajectory is implemented in the objective functional as y_Q, $c_Q = 1$.

Turning a wave, optimal control

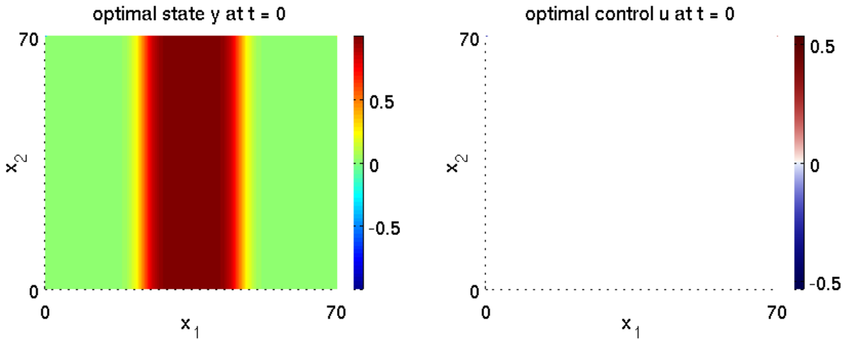
$\mu = 0$ video





Turning a wave, sparse optimal control

$\mu = 1$, video sparse control



Reference

E. Casas, C. Ryll, F. T.
Sparse optimal control of the Schlögl and FitzHugh-Nagumo systems
Computational Methods in Applied Mathematics 13 (2014), 415–442

Outline

1. Control of the Schlögl model
   - Traveling waves
   - Optimal control problem
   - An example
   - Necessary optimality conditions
   - Sparse control of the Schlögl model
2. Sparse control of spiral waves – FitzHugh Nagumo equations
   - Analysis of the FitzHugh-Nagumo system
   - Sparse optimal control of the FitzHugh-Nagumo system

The FitzHugh-Nagumo model

This model consists of two PDEs, the equation for the activator and the inhibitor. It plays an important role in neurobiology and is known to generate wave fronts (1D), spiral waves (2D), or scroll rings (3D).

FitzHugh-Nagumo equations

$$
\begin{aligned}
y_t(x,t) - \Delta y(x,t) + R\,(y(x,t)) + z(x,t) &= u(x,t) && \text{in } Q \\
\partial_n y(x,t) &= 0 && \text{in } \Sigma_T \\
y(x,0) &= y_0(x) && \text{in } \Omega \\
z_t(x,t) + \beta\, z(x,t) - \gamma\, y(x,t) + \delta &= 0 && \text{in } Q \\
z(x,0) &= z_0(x) && \text{in } \Omega.
\end{aligned}
$$

Nonlinearity

$$R(y) = \rho\,(y - y_1)(y - y_2)(y - y_3), \qquad \rho > 0$$

Some references

R. FitzHugh
Impulses and physiological states in theoretical models of nerve membrane,
Biophys. Journal (1961)

A. J. V. Brandão, E. Fernández-Cara, P. M. D. Paulo, M. A. Rojas-Medar
Theoretical analysis and control results for the FitzHugh-Nagumo equation,
Electron. J. Differential Equations (2008)

Kunisch, K., Wang, L.
Time optimal controls of the linear Fitzhugh-Nagumo equation with pointwise control constraints
J. Math. Anal. Appl. 395, 2012

E. Casas, C. Ryll, F. Tröltzsch
Sparse optimal control of the Schlögl and FitzHugh-Nagumo systems
Computational Methods in Applied Mathematics 13 (2014), 415–442

FitzHugh-Nagumo model

FitzHugh-Nagumo equations

$$
\begin{aligned}
y_t(x,t) - \Delta y(x,t) + R\,(y(x,t)) + z(x,t) &= u(x,t) && \text{in } Q\\
\partial_n y(x,t) &= 0 && \text{in } \Sigma_T\\
y(x,0) &= y_0(x) && \text{in } \Omega\\
z_t(x,t) + \beta\, z(x,t) - \gamma\, y(x,t) + \delta &= 0 && \text{in } Q\\
z(x,0) &= z_0(x) && \text{in } \Omega.
\end{aligned}
$$

$$R(y) = \rho\,(y - y_1)(y - y_2)(y - y_3).$$

Simplification

To simplify the presentation, assume $z_0 = 0,\ \delta = 0.$

- Assumption: $\Omega \subset \mathbb{R}^n$, $n \le 3$, bounded Lipschitz domain

Transformation to an integro-differential equation

$$
\begin{aligned}
z_t(x,t) + \beta\, z(x,t) - \; y(x,t) &= 0 \quad \text{in } Q \\
z(x,0) &= 0 \quad \text{in } \Omega.
\end{aligned}
$$

$$
\Longrightarrow \qquad z(x,t) = \int_0^t e^{-\beta(t-s)}\, y(x,s)ds = (K\,y)\,(x,t),
$$

with

$$
(K\,y)(x,t) = \int_0^t e^{-\beta(t-s)} y(x,s)ds.
$$

Transformation to an integro-differential equation

We insert the expression for z in the PDE for y,

$$y_t - \Delta y + R(y) + K\, y = u.$$

Since R is not monotone, we apply the same trick as for the Schlögl model,

$$y(x,t) := e^{\eta\, t}\, v(x,t)$$

with sufficiently large $\eta > 0$:

$$\frac{\partial}{\partial t} v - \Delta v + e^{-\eta\, t} R(e^{\eta\, t} v) + \eta v + K_\eta\, v = e^{-\eta\, t} u$$

$$(K_\eta\, v)\,(x,t) := \int_0^t e^{-(\beta+\eta)(t-s)}\, v(x,s) ds.$$

An important property

"K_η is small for large η". This means that the monotone term $\eta\, v$ dominates K_η for sufficiently large η.

A priori estimate for v

The parameter η is taken sufficiently large to

- make the operator K_η small and
- to get a monotone nonlinearity.

We write the parabolic PDE in the form

$$\frac{\partial}{\partial t} v - \Delta v + \underbrace{e^{-\eta\, t} R\left(e^{\eta\, t} v\right) + \frac{\eta}{3} v}_{R_\eta(t,v),\ \text{monotone}} + \underbrace{\frac{\eta}{3} v}_{\text{additional coercivity}} + \underbrace{\left(\frac{\eta}{3} v + K_\eta v\right)}_{\text{"positive"}} = e^{-\eta\, t} u.$$

Therefore, in energy estimates, this PDE behaves like the semilinear equation with monotone nonlinearity

$$\frac{\partial}{\partial t} v - \Delta v + R_\eta(t, v) + \frac{\eta}{3} v = e^{-\eta\, t} u.$$

Now there come L^2- and L^∞ a priori estimates to find out, in which ball we should find a solution.

Existence and uniqueness

Theorem

For all $\eta \geq \eta_0$, $u \in L^p(Q)$ with $p > 5/2$, and $y_0 \in L^\infty(\Omega)$, the integro-differential system has a unique solution $v \in W(0,T) \cap L^\infty(Q) \cap C\left(\overline{\Omega} \times (0,T]\right)$. There is a constant $C_\infty > 0$ such that

$$\|v\|_{L^\infty(Q)} + \|v\|_{W(0,T)} \leq C_\infty \left(\|u\|_{L^p(Q)} + \|y_0\|_{L^\infty(\Omega)} + |R(0)|\right).$$

Main idea of the proof – fixed point principle

- For given $w \in L^2(Q)$, we consider the semilinear equation

$$\frac{\partial}{\partial t} v - \Delta v + \underbrace{\hat{R}_\eta(t, v)}_{cut-off-fct.} + \frac{2}{3}\eta\, v = \underbrace{u - K_\eta w}_{\in L^2(Q)}$$

  subject to $v(\cdot, 0) = y_0$ and $\partial_\nu v = 0$ which has a unique solution $v \in W(0, T)$.

- Let

$$F : L^2(Q) \to L^2(Q), \qquad F : w \mapsto v.$$

  By our L^2-a-priori estimate, we can fix

$$M := C_2 \left(\|u\|_{L^2(Q)} + \|y_0\|_{L^2(\Omega)} + |R(0)| \right)$$

  and can find the solution in the L^2-ball with radius $2M$ centered at 0,

$$\|w\|_{L^2(Q)} \le 2\, M.$$

- We show $\quad F : B_{2M}(0) \to B_{2M}(0) \quad$ in $L^2(Q)$.

Steps of the proof

- By $W(0,T) \subset\subset L^2(Q)$, F is compact.
- Schauder's theorem: F has a fixed point $v \in B_{2M}(0)$.
- v solves the integro-differential system with the cutoff $\hat{R}_\eta$.
- $L^\infty-$a priori estimate $\Rightarrow \hat{R}_\eta(v) = R_\eta(v)$.
- Uniqueness is standard. $\square$

Existence and uniqueness

Theorem (Existence and uniqueness)

For all $\eta \geq \eta_0$, $u \in L^p(Q)$ *with* $p > 5/2$, *and* $y_0 \in L^\infty(\Omega)$, *the FitzHugh-Nagumo system has a unique solution* $(y_u, z_u) \in \left(W(0,T) \cap L^\infty(Q) \cap C\left(\overline{\Omega} \times (0,T]\right)\right)^2$. *There is a constant* $C_\infty > 0$ *such that*

$$\max\{\|y\|_{L^\infty(Q)}, \|y\|_{W(0,T)}, \|z\|_{L^\infty(Q)}, \|z\|_{W(0,T)}\} \leq C_\infty \left\{\|u\|_{L^p(Q)} + \|y_0\|_{L^\infty(\Omega)} + |R(0)|\right\}.$$

E. Casas, C. Ryll, F. Tröltzsch

Sparse optimal control of the Schlögl and FitzHugh-Nagumo systems

Computational Methods in Applied Mathematics 13 (2014), 415–442

Differentiability of the control-to-state mapping

Similarly, we prove the differentiability of the control-to-state mapping.

Theorem (Differentiability)

The solution mapping $G : u \mapsto (y_u, z_u)$ associated with the FitzHugh-Nagumo system is of class C^2 from $L^p(Q)$, $p > 5/2$, to $\big(W(0,T) \cap L^\infty(Q) \cap C(\overline{\Omega} \times (0,T])\big)^2$. The derivative $(y_h, z_h) := G'(u)h$ is equal to the pair (y, z) solving the system

$$
\begin{aligned}
\frac{\partial}{\partial t} y - \Delta y + R'(y_u) y + z &= h && \text{in } Q\\
\partial_n y &= 0 && \text{in } \Sigma_T\\
y(x,0) &= 0 && \text{in } \Omega\\
\frac{\partial}{\partial t} z + \beta z - \gamma y &= 0 && \text{in } Q\\
z(x,0) &= 0 && \text{in } \Omega.
\end{aligned}
$$

Adjoint system

For completeness, let us mention also the adjoint system for a pair of adjoint states $(\varphi_1, \varphi_2) \in W(0,T) \times W(0,T)$

Adjoint system

$$
\begin{aligned}
-\frac{\partial}{\partial t}\varphi_1 - \Delta\varphi_1 + R'(y)\,\varphi_1 - \gamma\,\varphi_2 &= c_Q^Y\,(y - y_Q) && \text{in } Q\\
\partial_n\varphi_1 &= 0 && \text{in } \Sigma_T\\
\varphi_1(x,T) &= c_T^Y(x)\,(y(x,T) - y_T(x)) && \text{in } \Omega\\
-\frac{\partial}{\partial t}\varphi_2 + \beta\,\varphi_2 + \varphi_1 &= c_Q^Z\,(z - z_Q) && \text{in } Q\\
\varphi_2(x,T) &= c_T^Z(x)\,(z(x,T) - z_T(x)) && \text{in } \Omega.
\end{aligned}
$$

Outline

1. Control of the Schlögl model
   - Traveling waves
   - Optimal control problem
   - An example
   - Necessary optimality conditions
   - Sparse control of the Schlögl model

2. Sparse control of spiral waves – FitzHugh Nagumo equations
   - Analysis of the FitzHugh-Nagumo system
   - Sparse optimal control of the FitzHugh-Nagumo system

Optimal control problem

$$\min_{u \in U_{ad}} \quad J(u) := I(u) + \mu\, j(u)$$

where

$$I(u) := \text{ quadratic tracking type functional (next slide)}$$

$$j(u) := \int_Q |u(x,t)|\, dxdt, \qquad \mu \geq 0$$

$$U_{ad} := \{u \in L^\infty(Q) \mid u(x,t) \in [\alpha, \beta] \text{ for a.a. } (x,t) \in Q\}$$

Assume $\alpha < 0 < \beta$.

Sparsity

By the term $\mu\, j(u)$, the optimal control becomes again sparse.

Quadratic functional

$$
\begin{aligned}
I(u) := \frac{1}{2}\int_Q c_Q^y(x,t)\,\left(y_u(x,t) - y_d(x,t)\right)^2\,dxdt \\
+ \frac{1}{2}\int_Q c_Q^z(x,t)\,\left(z_u(x,t) - z_d(x,t)\right)^2\,dxdt \\
+ \frac{\nu}{2}\int_Q u^2(x,t)\,dxdt \qquad \text{with } \nu > 0
\end{aligned}
$$

and nonnegative bounded and measurable coefficient functions c_Q^y, c_Q^z.

More general functionals can also be discussed.

Necessary optimality conditions

Lemma (Variational inequality)

If $(\bar{y}, \bar{z}, \bar{u})$ is a local solution to the optimal control problem, then there exists $\bar{\lambda} \in \partial j(\bar{u})$ such that

$$I'(\bar{u})(u - \bar{u}) + \int_Q \mu\, \bar{\lambda}(x,t)(u(x,t) - \bar{u}(x,t))\, dxdt \geq 0 \qquad \forall u \in U_{ad}.$$

Theorem (Necessary optimality conditions)

If $\bar{u}$ is a local solution to the optimal control problem, then there exists a unique pair of adjoint states $(\bar{\varphi}_1, \bar{\varphi}_2) \in W(0,T)^2$ such that, with $\bar{\lambda} \in \mu\, \partial j(\bar{u})$

$$\int_Q \left(\bar{\varphi}_1(x,t) + \nu\, \bar{u}(x,t) + \mu\, \bar{\lambda}(x,t)\right)(u(x,t) - \bar{u}(x,t))\ dxdt \geq 0 \qquad \forall u \in U_{ad}.$$

Sparsity

After a detailed pointwise discussion of the necessary optimality conditions, we find the following

Theorem

For almost all $(x, t) \in Q$*,*

$$\begin{aligned} \bar{u}(x,t) &= 0 \;\; \textit{iff} \;\; |\bar{\varphi}_1(x,t)| \le \mu, \\ \bar{\lambda}(x,t) &= \mathbb{P}_{[-1,1]}\left\{-\frac{1}{\mu}\bar{\varphi}_1(x,t)\right\}. \end{aligned}$$

- The first relation expresses the effects of sparsity.
- The second is used to set up the conjugate gradient in the nonlinear projected cg method for solving the optimal control problem.

Computational examples

Exciting a spiral wave

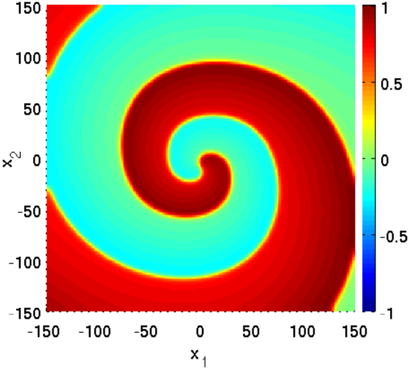
Let Ω be rectangular and $u = 1$ close to the bottom boundary of Ω in a certain short period of time and $u = 0$ elsewhere. As result, a traveling wave appears that propagates to the upper boundary of the spatial domain. After a short period of time, when the wave front is located between the upper and the bottom boundary, we set the state (y, z) equal to zero in the left half of Ω. Then the wave starts to curl up and forms a spiral pattern.

Example: Acceleration of a spiral wave

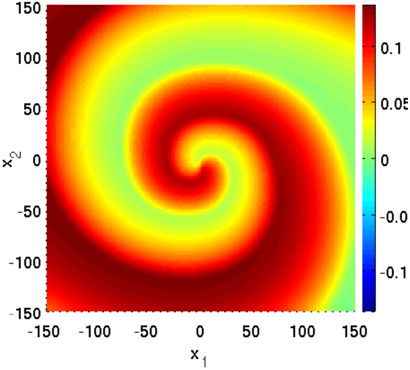
- $\Omega = (-150, 150)^2$, $T = 50$,
- $\gamma = 1/500$, $\delta = 0$, $\beta = 1/100$,
- Constraint $|u(x, t)| \leq 5$
- $R(y) = y(y - 1/20)(y - 1)$
- Initial state (y_0, z_0) as in the next figure

Initial states

The initial spirals were generated by the method explained at the last slide.



Initial state y_0



initial state z_0

Desired states

Desired states y_Q and z_Q

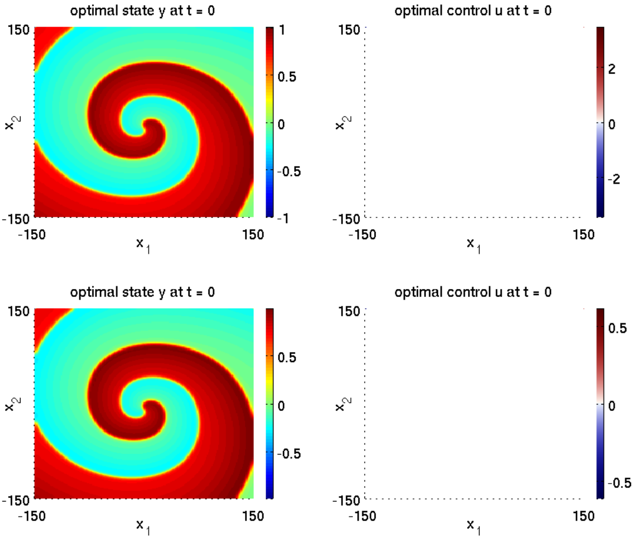
- $(y_{\text{nat}}, z_{\text{nat}}) :=$ Natural development of (y, z) for $u \equiv 0$, starting at (y_0, z_0).
- $y_Q(x,t) := y_{\text{nat}}\left(x, \frac{1}{5}t^2 + t\right), \quad z_Q(x,t) := z_{\text{nat}}\left(x, \frac{1}{5}t^2 + t\right).$

  The term $t^2/5$ accounts for the acceleration.

Accelerating spiral waves; videos

$\mu = 0$

$\mu = \frac{1}{3}$



Effects of sparsity

- A positive parameter $\mu > 0$ causes sparsity of the optimal control and accelerates the cg method considerably.
- Instead of 850 iterations in the case of $\mu = 0$, the CG-method stopped after only 59 iterations for $\mu = 1/3$.

Extinction of a spiral wave

Data

- $\Omega = (-120, 120)^2$, $T = 2500$
- $\gamma = \frac{3}{400}$, $y_2 = \frac{1}{200}$

Control bounds

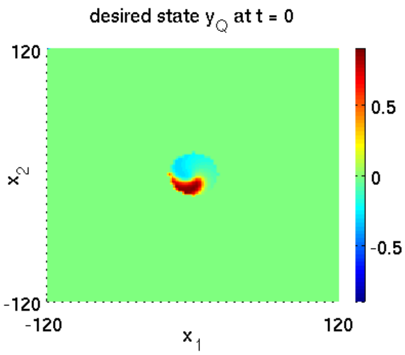
$$|u(x,t)| \leq 5$$

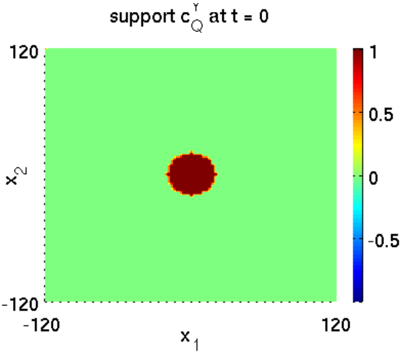
Moving domain of observation; videos

A hint from our physicists:

Control a spiral in its center! (Diploma thesis, Breuer 2006)

We take as observation domain a (moving) circle around the initial center point of the spiral and move this point to the boundary of Ω.

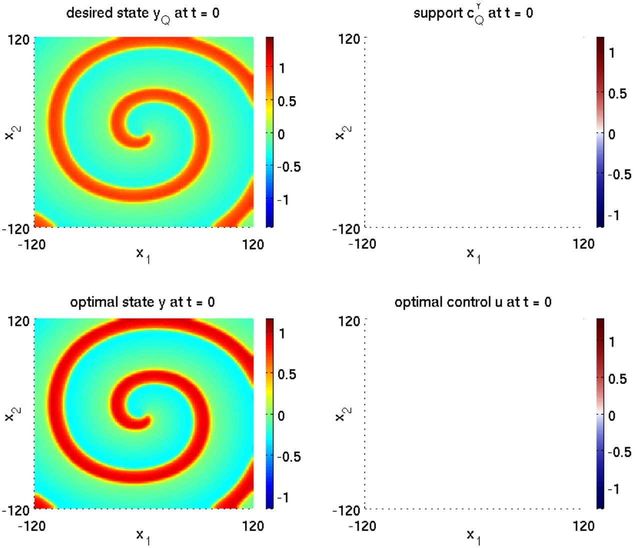




Extinction of a spiral wave; videos

$\mu = 0$

$\mu = 1$



Directional sparsity

References

R. Herzog, G. Stadler, G. Wachsmuth
Directional sparsity in optimal control of partial differential equations,
SIAM Journal on Control and Optimization (2012)

R. Herzog, J. Obermeier, G. Wachsmuth
Annular and Sectorial Sparsity in Optimal Control of Elliptic Equations
Computational Optimization and Applications 62, 2015

The optimal control problem

$$\min_{u \in U_{ad}} J(y,u) := \frac{1}{2}\|y - y_\Omega\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{1,2}$$

subject to the Poisson equation with right-hand side u,

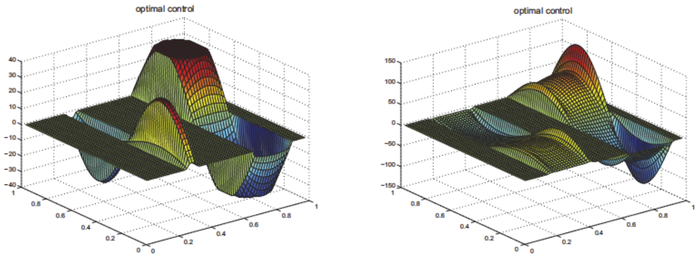
$$U_{ad} := \{u \in L^2(\Omega) : \, \alpha \le u(x) \le \beta \text{ a.e. in } \Omega\}.$$

and given $y_\Omega \in L^2(\Omega)$, $\nu > 0$, $\mu > 0$.

Definition

$$\|u\|_{1,2} = \int_{\Omega_1} \left(\int_{\Omega_2} u^2(x_1, x_2)\, dx_2 \right)^{\frac{1}{2}} dx_1 = \int_{\Omega_1} \|u(x_1, \cdot)\|_{L^2(\Omega_2)}\, dx_1$$

Here, the L^1-norm w.r.t. to x_1 causes sparsity in the L^2-norm w.r.t. to x_2. This means that, for certain subsets of Ω_1, the L^2-norm vanishes. The optimal control forms striped patterns.



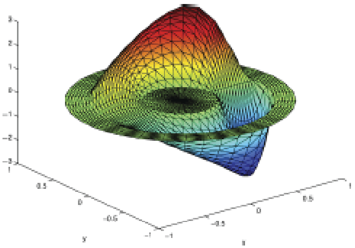
Examples of optimal controls with directional sparsity, copied out of

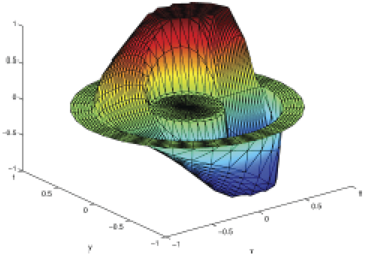
R. Herzog, G. Stadler, G. Wachsmuth
Directional sparsity in optimal control of partial differential equations,
SIAM Journal on Control and Optimization (2012)

Annular sparsity

Sparsity by the norm

$$\int_0^R \left(\int_0^{2\pi} u(r,\varphi)^2 \, d\varphi \right)^{1/2} r \, dr.$$





Copied out of

R. Herzog, J. Obermeier, G. Wachsmuth

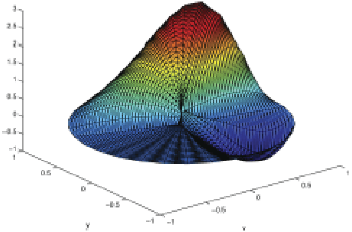
Annular and Sectorial Sparsity in Optimal Control of Elliptic Equations

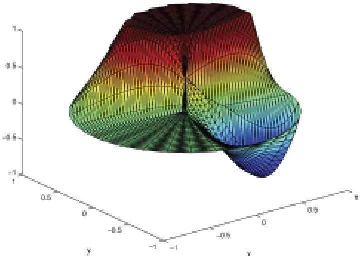
Computational Optimization and Applications 62, 2015

Sectorial sparsity

Sparsity by the norm

$$\int_0^{2\pi} \left(\int_0^R u(r,\varphi)^2 \, r \, dr \right)^{1/2} d\varphi.$$





Copied out of

R. Herzog, J. Obermeier, G. Wachsmuth

Annular and Sectorial Sparsity in Optimal Control of Elliptic Equations

Computational Optimization and Applications 62, 2015