One Sided Minimax Differentiability for the Computation of Control, Shape, and Topological Derivatives

Michel C. Delfour[1] and Kevin Sturm[2]

[1] Centre de recherches mathématiques
Département de mathématiques et de statistique
Université de Montréal, Canada

[2] Johann Radon Institute
Altenberger Strasse 69
4040 Linz, Austria

NUMOC - June 23, 2017, Roma

Outline

Outline

Let $\Omega \subset \mathbb{R}^N$, $N \geq 1$, bounded open and $a \in L^2(\Omega)$ be the *control variable* to which is associated the *state* $u = u(a) \in H_0^1(\Omega)$ solution of the variational *state equation*

$$\int_\Omega \nabla u(a) \cdot \nabla \psi - a\,\psi\,dx = 0, \quad \forall \psi \in H_0^1(\Omega), \tag{1.1}$$

where $x \cdot y$ denotes the inner product of x and y in $\mathbb{R}^N$.

Given a target function $g \in L^2(\Omega)$, associate with $u(a)$ the *objective function*

$$f(a) \overset{\text{def}}{=} \int_\Omega \frac{1}{2}|u(a) - g|^2\,dx. \tag{1.2}$$

By introducing the Lagrangian, we get an unconstrained minimax formulation

$$G(a, \varphi, \psi) \overset{\text{def}}{=} \int_\Omega \frac{1}{2}|\varphi - g|^2\,dx + \int_\Omega \nabla\varphi \cdot \nabla\psi - a\,\psi\,dx$$

$$\boxed{f(a) = \inf_{\varphi \in H^1(\Omega)} \sup_{\psi \in H^1(\Omega)} G(a, \varphi, \psi).}$$

Let $\Omega \subset \mathbb{R}^N$, $N \geq 1$, bounded open and $a \in L^2(\Omega)$ be the *control variable* to which is associated the *state* $u = u(a) \in H_0^1(\Omega)$ solution of the variational *state equation*

$$\int_\Omega \nabla u(a) \cdot \nabla \psi - a\,\psi\, dx = 0, \quad \forall \psi \in H_0^1(\Omega), \tag{1.1}$$

where $x \cdot y$ denotes the inner product of x and y in $\mathbb{R}^N$.

Given a target function $g \in L^2(\Omega)$, associate with $u(a)$ the *objective function*

$$f(a) \stackrel{\text{def}}{=} \int_\Omega \frac{1}{2}|u(a) - g|^2\, dx. \tag{1.2}$$

By introducing the Lagrangian, we get an unconstrained minimax formulation

$$G(a, \varphi, \psi) \stackrel{\text{def}}{=} \int_\Omega \frac{1}{2}|\varphi - g|^2\, dx + \int_\Omega \nabla\varphi \cdot \nabla\psi - a\,\psi\, dx$$

$$\boxed{f(a) = \inf_{\varphi \in H^1(\Omega)} \sup_{\psi \in H^1(\Omega)} G(a, \varphi, \psi).}$$

If we are only interested in a descent method, we can obtain the semidifferential of $f(a)$ by a similar minimax formulation. Given the direction $b \in L^2(\Omega)$, to compute

$$df(a;b) = \lim_{t \searrow 0} \frac{f(a+tb) - f(a)}{t},$$

where the *state* $u^t \in H_0^1(\Omega)$ at $t > 0$ is solution of

$$\int_\Omega \nabla u^t \cdot \nabla \psi - (a+tb)\,\psi\,dx = 0, \quad \forall \psi \in H_0^1(\Omega). \tag{1.3}$$

The *associated Lagrangian* is

$$L(t,\varphi,\psi) \overset{\text{def}}{=} \int_\Omega \frac{1}{2}|\varphi - g|^2\,dx + \int_\Omega \nabla\varphi \cdot \nabla\psi - (a+tb)\,\psi\,dx.$$

It is readily seen that

$$\boxed{\begin{aligned} g(t) &\overset{\text{def}}{=} \inf_{\varphi \in H_0^1(\Omega)} \sup_{\psi \in H_0^1(\Omega)} L(t,\varphi,\psi) = f(a+tb) \\ dg(0) &\overset{\text{def}}{=} \lim_{t \searrow 0} \frac{g(t) - g(0)}{t} = df(a;b). \end{aligned}}$$

Outline

Consider the state (1.1) and objective function (1.2). Now perturb the domain Ω by a family of diffeomorphisms T_t generated by a smooth velocity field $V(t)$:

$$\frac{dx}{dt}(t;X) = V(t, x(t;X)),\ x(0;X) = X, \quad T_t(X) \stackrel{\text{def}}{=} x(t;X),\ t \geq 0, \quad \boxed{\Omega_t \stackrel{\text{def}}{=} T_t(\Omega).}$$

The state equation and objective function at $t > 0$ become

$$\int_{\Omega_t} \nabla u_t \cdot \nabla \psi - a\,\psi\, dx = 0, \quad \forall \psi \in H_0^1(\Omega_t), \quad f(t) \stackrel{\text{def}}{=} \int_{\Omega_t} |u_t - g|^2\, dx. \tag{1.4}$$

Introducing the composition $u^t = u_t \circ T_t$ to work in the fixed space $H_0^1(\Omega)$:

$$\int_\Omega \left[A(t)\nabla u^t \cdot \nabla \psi - a\,\psi\right] j(t)\, dx = 0, \quad \forall \psi \in H_0^1(\Omega), \tag{1.5}$$

$$A(t) = DT_t^{-1}\,(DT_t^{-1})^*,\ j(t) = \det DT_t, \quad DT_t \text{ is the Jacobian matrix,} \tag{1.6}$$

$$\Rightarrow f(t) = \int_{\Omega_t} |u_t - g|^2\, dx = \int_\Omega |u^t - g \circ T_t|^2\, j(t)\, dx, \tag{1.7}$$

$$\text{Lagrangian}: L(t, \varphi, \psi) \stackrel{\text{def}}{=} \int_\Omega \left[\frac{1}{2}|\varphi - g \circ T_t|^2 + A(t)\nabla\varphi \cdot \nabla\psi - a\,\psi\right] j(t)\, dx.$$

$$\boxed{\Rightarrow g(t) = \inf_{\varphi \in H_0^1(\Omega)} \sup_{\psi \in H_0^1(\Omega)} L(t, \varphi, \psi), \quad dg(0) = \lim_{t \searrow 0} (g(t) - g(0))/t = df(\Omega; V(0)).}$$

Outline

Associate with a real vector space (usually a Banach space) Θ of mappings $\theta : \mathbb{R}^N \to \mathbb{R}^N$ (Micheletti used the space $\Theta = C_0^k(\mathbb{R}^N, \mathbb{R}^N)$, $k \geq 1$), the following space of transformations (endomorphisms) of $\mathbb{R}^N$:

$$\boxed{\mathcal{F}(\Theta) \stackrel{\text{def}}{=} \left\{ F : \mathbb{R}^N \to \mathbb{R}^N \text{ bijective } : F - I \in \Theta, \text{ and } F^{-1} - I \in \Theta \right\},} \tag{1.8}$$

where $x \mapsto I(x) \stackrel{\text{def}}{=} x : \mathbb{R}^N \to \mathbb{R}^N$ is the identity mapping.

Given a fixed set $\Omega_0 \subset \mathbb{R}^N$ (Micheletti used used a bounded open set of class C^k), consider the set of images

$$\mathcal{X}(\Omega_0) \stackrel{\text{def}}{=} \{F(\Omega_0) : \forall F \in \mathcal{F}(\Theta)\} \tag{1.9}$$

of Ω_0 by the elements of $\mathcal{F}(\Theta)$ and the subgroup

$$\mathcal{G}(\Omega_0) \stackrel{\text{def}}{=} \{F \in \mathcal{F}(\Theta) : F(\Omega_0) = \Omega_0\}.$$

So there is a bijection between the set of images of Ω_0 and the quotient space

$$\boxed{\mathcal{X}(\Omega_0) \longleftrightarrow \mathcal{F}(\Theta)/\mathcal{G}(\Omega_0).}$$

The objective is to construct a metric on $\mathcal{F}(\Theta)/\mathcal{G}(\Omega_0)$ that will serve as a distance between two mages $F_1(\Omega_0)$ and $F_2(\Omega_0)$.

Associate with $F \in \mathcal{F}(\Theta)$ the following candidate for a metric

$$d_0(I, F) \stackrel{\text{def}}{=} \|F - I\|_\Theta + \|F^{-1} - I\|_\Theta, \quad d_0(F, G) \stackrel{\text{def}}{=} d_0(I, G \circ F^{-1}). \tag{1.10}$$

Unfortunately, d_0 is only a semi-metric that will not satisfy the triangle inequality.

Consider the following second candidate (called Courant metric by Micheletti)

$$d(I, F) \stackrel{\text{def}}{=} \inf_{\substack{F = F_1 \circ \cdots \circ F_n \\ F_i \in \mathcal{F}(\Theta)}} \sum_{i=1}^{n} \|F_i - I\|_\Theta + \|F_i^{-1} - I\|_\Theta, \tag{1.11}$$

where the infimum is taken over all *finite factorizations* of F in $\mathcal{F}(\Theta)$ of the form

$$F = F_1 \circ \cdots \circ F_n, \quad F_i \in \mathcal{F}(\Theta).$$

In particular $d(I, F) = d(I, F^{-1})$. Extend this function to all F and G in $\mathcal{F}(\Theta)$

$$d(F, G) \stackrel{\text{def}}{=} d(I, G \circ F^{-1}). \tag{1.12}$$

By definition, d is right-invariant since for all F, G and H in $\mathcal{F}(\Theta)$

$$d(F, G) = d(F \circ H, G \circ H).$$

The objective is to construct a metric on $\mathcal{F}(\Theta)/\mathcal{G}(\Omega_0)$ that will serve as a distance between two mages $F_1(\Omega_0)$ and $F_2(\Omega_0)$.

Associate with $F \in \mathcal{F}(\Theta)$ the following candidate for a metric

$$d_0(I, F) \stackrel{\text{def}}{=} \|F - I\|_\Theta + \|F^{-1} - I\|_\Theta, \quad d_0(F, G) \stackrel{\text{def}}{=} d_0(I, G \circ F^{-1}). \tag{1.10}$$

Unfortunately, d_0 is only a semi-metric that will not satisfy the triangle inequality.

Consider the following second candidate (called Courant metric by Micheletti)

$$d(I, F) \stackrel{\text{def}}{=} \inf_{\substack{F = F_1 \circ \cdots \circ F_n \\ F_i \in \mathcal{F}(\Theta)}} \sum_{i=1}^{n} \|F_i - I\|_\Theta + \|F_i^{-1} - I\|_\Theta, \tag{1.11}$$

where the infimum is taken over all *finite factorizations* of F in $\mathcal{F}(\Theta)$ of the form

$$F = F_1 \circ \cdots \circ F_n, \quad F_i \in \mathcal{F}(\Theta).$$

In particular $d(I, F) = d(I, F^{-1})$. Extend this function to all F and G in $\mathcal{F}(\Theta)$

$$d(F, G) \stackrel{\text{def}}{=} d(I, G \circ F^{-1}). \tag{1.12}$$

By definition, d is right-invariant since for all F, G and H in $\mathcal{F}(\Theta)$

$$d(F, G) = d(F \circ H, G \circ H).$$

The objective is to construct a metric on $\mathcal{F}(\Theta)/\mathcal{G}(\Omega_0)$ that will serve as a distance between two mages $F_1(\Omega_0)$ and $F_2(\Omega_0)$.

Associate with $F \in \mathcal{F}(\Theta)$ the following candidate for a metric

$$d_0(I,F) \overset{\text{def}}{=} \|F-I\|_\Theta + \|F^{-1}-I\|_\Theta, \quad d_0(F,G) \overset{\text{def}}{=} d_0(I, G\circ F^{-1}). \tag{1.10}$$

Unfortunately, d_0 is only a semi-metric that will not satisfy the triangle inequality.

Consider the following second candidate (called Courant metric by Micheletti)

$$d(I,F) \overset{\text{def}}{=} \inf_{\substack{F=F_1\circ\cdots\circ F_n\\ F_i\in\mathcal{F}(\Theta)}} \sum_{i=1}^{n} \|F_i-I\|_\Theta + \|F_i^{-1}-I\|_\Theta, \tag{1.11}$$

where the infimum is taken over all *finite factorizations* of F in $\mathcal{F}(\Theta)$ of the form

$$F = F_1\circ\cdots\circ F_n, \quad F_i\in\mathcal{F}(\Theta).$$

In particular $d(I,F) = d(I,F^{-1})$. Extend this function to all F and G in $\mathcal{F}(\Theta)$

$$d(F,G) \overset{\text{def}}{=} d(I, G\circ F^{-1}). \tag{1.12}$$

By definition, d is right-invariant since for all F, G and H in $\mathcal{F}(\Theta)$

$$d(F,G) = d(F\circ H, G\circ H).$$

$(\mathcal{F}(\Theta), d)$ is complete for Θ equal to the Banach spaces

$$C_0^k(\mathbb{R}^N, \mathbb{R}^N),\ C^k(\overline{\mathbb{R}^N}, \mathbb{R}^N) \subset \mathcal{B}^k(\mathbb{R}^N, \mathbb{R}^N) \text{ and } C^{k,1}(\overline{\mathbb{R}^N}, \mathbb{R}^N), \quad k \geq 0,$$

and, through special constructions, for the Fréchet spaces

$$C_0^\infty(\mathbb{R}^N, \mathbb{R}^N) \subset \mathcal{B}(\mathbb{R}^N, \mathbb{R}^N) = \cap_{k\geq 0}\mathcal{B}^k(\mathbb{R}^N, \mathbb{R}^N).$$

For any Banach or Fréchet space $\Theta \subset C^{0,1}(\overline{\mathbb{R}^N}, \mathbb{R}^N)$, $\mathcal{F}(\Theta)$ is an open subset of $I + \Theta$

- the tangent space is Θ at each point $F \in \mathcal{F}(\Theta)$
- and the associated smooth structure is trivial.

The analogue would be the general linear group $GL(n)$ of invertible linear maps from $\mathbb{R}^N$ to $\mathbb{R}^N$ which is an open subset of $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$. So, the tangent space is $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$.

Choose $\Theta = C_0^k(\mathbb{R}^N, \mathbb{R}^N)$, $k \geq 1$, $\mathcal{F}(\Theta)$, and the set $\mathcal{X}(\Omega_0)$ of the images of an open crack free set $\Omega_0 \subset \mathbb{R}^N$. Consider a function $J : \mathcal{X}(\Omega_0) \to \mathbb{R}$.

THEOREM

Let $\Omega = F(\Omega_0) \in \mathcal{X}(\Omega_0)$ for some $F \in \mathcal{F}(\Omega_0)$. Then J is continuous at Ω for the Courant metric if and only if

$$\lim_{t\searrow 0} J(T_t(\Omega)) = J(\Omega), \quad \frac{dT_t}{dt} = V(t) \circ T_t, \quad T_0 = F,$$

for all families of velocity fields $V \in C^0([0,\tau]; C_0^k(\mathbb{R}^N, \mathbb{R}^N))$.

$(\mathcal{F}(\Theta), d)$ is complete for Θ equal to the Banach spaces

$$C_0^k(\mathbb{R}^N, \mathbb{R}^N),\ C^k(\overline{\mathbb{R}^N}, \mathbb{R}^N) \subset \mathcal{B}^k(\mathbb{R}^N, \mathbb{R}^N) \text{ and } C^{k,1}(\overline{\mathbb{R}^N}, \mathbb{R}^N), \quad k \geq 0,$$

and, through special constructions, for the Fréchet spaces

$$C_0^\infty(\mathbb{R}^N, \mathbb{R}^N) \subset \mathcal{B}(\mathbb{R}^N, \mathbb{R}^N) = \cap_{k\geq 0}\mathcal{B}^k(\mathbb{R}^N, \mathbb{R}^N).$$

For any Banach or Fréchet space $\Theta \subset C^{0,1}(\overline{\mathbb{R}^N}, \mathbb{R}^N)$, $\mathcal{F}(\Theta)$ is an open subset of $I + \Theta$

- the tangent space is Θ at each point $F \in \mathcal{F}(\Theta)$
- and the associated smooth structure is trivial.

The analogue would be the general linear group $GL(n)$ of invertible linear maps from $\mathbb{R}^N$ to $\mathbb{R}^N$ which is an open subset of $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$. So, the tangent space is $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$.

Choose $\Theta = C_0^k(\mathbb{R}^N, \mathbb{R}^N)$, $k \geq 1$, $\mathcal{F}(\Theta)$, and the set $\mathcal{X}(\Omega_0)$ of the images of an open crack free set $\Omega_0 \subset \mathbb{R}^N$. Consider a function $J : \mathcal{X}(\Omega_0) \to \mathbb{R}$.

THEOREM

Let $\Omega = F(\Omega_0) \in \mathcal{X}(\Omega_0)$ for some $F \in \mathcal{F}(\Omega_0)$. Then J is continuous at Ω for the Courant metric if and only if

$$\lim_{t\searrow 0} J(T_t(\Omega)) = J(\Omega), \quad \frac{dT_t}{dt} = V(t)\circ T_t, \quad T_0 = F,$$

for all families of velocity fields $V \in C^0([0,\tau]; C_0^k(\mathbb{R}^N, \mathbb{R}^N))$.

$(\mathcal{F}(\Theta), d)$ is complete for Θ equal to the Banach spaces

$$C_0^k(\mathbb{R}^N, \mathbb{R}^N),\ C^k(\overline{\mathbb{R}^N}, \mathbb{R}^N) \subset \mathcal{B}^k(\mathbb{R}^N, \mathbb{R}^N) \text{ and } C^{k,1}(\overline{\mathbb{R}^N}, \mathbb{R}^N), \quad k \geq 0,$$

and, through special constructions, for the Fréchet spaces

$$C_0^\infty(\mathbb{R}^N, \mathbb{R}^N) \subset \mathcal{B}(\mathbb{R}^N, \mathbb{R}^N) = \cap_{k \geq 0} \mathcal{B}^k(\mathbb{R}^N, \mathbb{R}^N).$$

For any Banach or Fréchet space $\Theta \subset C^{0,1}(\overline{\mathbb{R}^N}, \mathbb{R}^N)$, $\mathcal{F}(\Theta)$ is an open subset of $I + \Theta$
- the tangent space is Θ at each point $F \in \mathcal{F}(\Theta)$
- and the associated smooth structure is trivial.

The analogue would be the general linear group $GL(n)$ of invertible linear maps from $\mathbb{R}^N$ to $\mathbb{R}^N$ which is an open subset of $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$. So, the tangent space is $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^N)$.

Choose $\Theta = C_0^k(\mathbb{R}^N, \mathbb{R}^N)$, $k \geq 1$, $\mathcal{F}(\Theta)$, and the set $\mathcal{X}(\Omega_0)$ of the images of an open crack free set $\Omega_0 \subset \mathbb{R}^N$. Consider a function $J : \mathcal{X}(\Omega_0) \to \mathbb{R}$.

Theorem

Let $\Omega = F(\Omega_0) \in \mathcal{X}(\Omega_0)$ for some $F \in \mathcal{F}(\Omega_0)$. Then J is continuous at Ω for the Courant metric if and only if

$$\lim_{t \searrow 0} J(T_t(\Omega)) = J(\Omega), \quad \frac{dT_t}{dt} = V(t) \circ T_t, \quad T_0 = F,$$

for all families of velocity fields $V \in C^0([0, \tau]; C_0^k(\mathbb{R}^N, \mathbb{R}^N))$.

DEFINITION

Let $\Theta = C_0^k(\mathbb{R}^N, \mathbb{R}^N)$. The function $J : \mathcal{X}(\Omega_0) = \{F(\Omega_0) : F \in \mathcal{F}(\Theta)\} \to \mathbb{R}$ is Hadamard semidifferentiable at $F(\Omega_0)$, $F \in \mathcal{F}(\Theta)$, if

(i) for all $V \in C^0([0,\tau]; \mathcal{D}(\mathbb{R}^N, \mathbb{R}^N))$

$$dJ(F(\Omega_0); V) \overset{\text{def}}{=} \lim_{t \searrow 0} \frac{J(T_t(V)(F(\Omega_0)) - J(F(\Omega_0))}{t} \quad \text{exists,} \quad \frac{dT_t}{dt} = V(t) \circ T_t,\ T_0 = F,$$

(ii) and there exists a function $dJ(F(\Omega_0)) : \mathcal{D}(\mathbb{R}^N, \mathbb{R}^N) \to \mathbb{R}$ such that for all $V \in C^0([0,\tau]; \mathcal{D}(\mathbb{R}^N, \mathbb{R}^N))$

$$dJ(F(\Omega_0); V) = dJ(F(\Omega_0))(V(0)).$$

DEFINITION

$J : \mathcal{X}(\Omega_0) \to \mathbb{R}$ is Hadamard differentiable at $F(\Omega_0)$, $F \in \mathcal{F}(\Theta)$, if
- it is Hadamard semidifferentiable at $F(\Omega_0)$
- and the function $dJ(F(\Omega_0)) : \mathcal{D}(\mathbb{R}^N, \mathbb{R}^N) \to \mathbb{R}$ is linear and continuous.

DEFINITION

Let $\Theta = C_0^k(\mathbb{R}^N, \mathbb{R}^N)$. The function $J : \mathcal{X}(\Omega_0) = \{F(\Omega_0) : F \in \mathcal{F}(\Theta)\} \to \mathbb{R}$ is Hadamard semidifferentiable at $F(\Omega_0)$, $F \in \mathcal{F}(\Theta)$, if

(i) for all $V \in C^0([0,\tau]; \mathcal{D}(\mathbb{R}^N, \mathbb{R}^N))$

$$dJ(F(\Omega_0); V) \overset{\text{def}}{=} \lim_{t \searrow 0} \frac{J(T_t(V)(F(\Omega_0)) - J(F(\Omega_0))}{t} \quad \text{exists}, \quad \frac{dT_t}{dt} = V(t) \circ T_t, \; T_0 = F,$$

Definition of the *Eulerian semiderivative* of Zolésio in his 1979 thesis [Zolésio (1979), p. 12].

(ii) and there exists a function $dJ(F(\Omega_0)) : \mathcal{D}(\mathbb{R}^N, \mathbb{R}^N) \to \mathbb{R}$ such that for all $V \in C^0([0,\tau]; \mathcal{D}(\mathbb{R}^N, \mathbb{R}^N))$

$$dJ(F(\Omega_0); V) = dJ(F(\Omega_0))(V(0)).$$

DEFINITION

$J : \mathcal{X}(\Omega_0) \to \mathbb{R}$ is Hadamard differentiable at $F(\Omega_0)$, $F \in \mathcal{F}(\Theta)$, if
- it is Hadamard semidifferentiable at $F(\Omega_0)$
- and the function $dJ(F(\Omega_0)) : \mathcal{D}(\mathbb{R}^N, \mathbb{R}^N) \to \mathbb{R}$ is linear and continuous.

Definition of the *Gradient of J* of Zolésio in his 1979 thesis [Zolésio (1979), p. 13]

Outline

Differentiability of MInima and Minimax

Danskin (1966, 1967)
Dem'janov (1969)
Lemaire (1970)
Dem'janov-Malozemov (1974)
Correa-Seeger (1985)
Delfour-Zolésio (1988)
Bonnans-Shapiro (1998, 2000)
Sturm (2014, 2015)
Delfour-Sturm (2017, 2016), . . .

Shape Derivative

Zolésio (1979)
Delfour-Zolésio (1988, 2001, 2011)

Topological Derivative

Sokołowski-Żochowski (1999)
Novotny-Sokolowski (2013)
Delfour (2017, 2017), . . .

Outline

In this paper, a *Lagrangian* is a function of the form

$$(t, x, y) \mapsto G(t, x, y) : [0, \tau] \times X \times Y \to \mathbb{R}, \quad \tau > 0,$$

where Y is a *vector space*, X is a subset of a vector space, and $y \mapsto G(t, x, y)$ is *affine*. Associate with the *parameter* $t \geq 0$ the *parametrized minimax function*

$$t \mapsto g(t) \stackrel{\text{def}}{=} \inf_{x \in X} \sup_{y \in Y} G(t, x, y) : [0, \tau] \to \mathbb{R}\,. \tag{2.1}$$

When the limits exist we shall use the following compact notation:

$$dg(0) \stackrel{\text{def}}{=} \lim_{t \searrow 0} \frac{g(t) - g(0)}{t} \qquad \begin{cases} \underline{d}g(0) \stackrel{\text{def}}{=} \liminf_{t \searrow 0} \, (g(t) - g(0))\,/t \\ \overline{d}g(0) \stackrel{\text{def}}{=} \limsup_{t \searrow 0} \, (g(t) - g(0))\,/t \end{cases}$$

$$d_t G(0, x, y) \stackrel{\text{def}}{=} \lim_{t \searrow 0} \frac{G(t, x, y) - G(0, x, y)}{t}$$

$$\varphi \in X, \quad d_x G(t, x, y; \varphi) \stackrel{\text{def}}{=} \lim_{\theta \searrow 0} \frac{G(t, x + \theta\varphi, y) - G(t, x, y)}{\theta}$$

$$\psi \in Y, \quad d_y G(t, x, y; \psi) \stackrel{\text{def}}{=} \lim_{\theta \searrow 0} \frac{G(t, x, y + \theta\psi) - G(t, x, y)}{\theta}.$$

The notation $t \searrow 0$ and $\theta \searrow 0$ means that t and θ go to 0 by strictly positive values

Since $G(t,x,y)$ is affine in y, for all $(t,x) \in [0,\tau] \times X$,

$$\forall y, \psi \in Y, \quad d_y G(t,x,y;\psi) = G(t,x,\psi) - G(t,x,0) = d_y G(t,x,0;\psi).$$

The state equation at $t \geq 0$:

$$\boxed{\text{to find } x^t \in X \text{ such that for all } \psi \in Y,\ d_y G(t,x^t,0;\psi) = 0.}$$

The set of solutions (*states*) x^t at $t \geq 0$ is denoted

$$E(t) \stackrel{\text{def}}{=} \left\{ x^t \in X : \forall \varphi \in Y,\ d_y G(t,x^t,0;\varphi) = 0 \right\}$$

The standard adjoint state equation at $t \geq 0$:

$$\boxed{\text{to find } p^t \in Y \text{ such that } \forall \varphi \in X, \quad d_x G(t,x^t,p^t;\varphi) = 0}, \quad Y(t,x^t) \stackrel{\text{def}}{=} \text{ set of solutions.}$$

Under appropriate conditions and uniqueness of the pair (x^t,p^t),

$$\boxed{dg(0) = d_t G(0,x^0,p^0),}$$

where (x^0,p^0) is the solution of the coupled state-adjoint state equations at $t = 0$.

$$\text{states:} \quad E(t) \stackrel{\text{def}}{=} \left\{ x^t \in X : \forall \varphi \in Y,\ d_y G(t, x^t, 0; \varphi) = 0 \right\}$$

$$\text{minimizers:} \quad X(t) \stackrel{\text{def}}{=} \left\{ x^t \in X : g(t) = \inf_{x \in X} \sup_{y \in Y} G(t, x, y) = \sup_{y \in Y} G(t, x^t, y) \right\}.$$

LEMMA (CONSTRAINED INFIMUM AND MINIMAX)

(i) $\inf_{x \in X} \sup_{y \in Y} G(t, x, y) = \inf_{x \in E(t)} G(t, x, 0)$.

(ii) *The minimax* $g(t) = +\infty$ *if and only if* $E(t) = \emptyset$. *In that case* $X(t) = X$.

(iii) *If* $E(t) \neq \emptyset$, *then*

$$\boxed{X(t) = \{x^t \in E(t) : G(t, x^t, 0) = \inf_{x \in E(t)} G(t, x, 0)\}} \subset E(t) \tag{2.2}$$

and $g(t) < +\infty$.

Hypothesis (H0). Let X be a vector space.

(i) For all $t \in [0, \tau]$, $x^0 \in X(0)$, $x^t \in X(t)$, and $y \in Y$, the function

$$s \mapsto G(t, x^0 + s(x^t - x^0), y) : [0, 1] \to \mathbb{R} \tag{2.3}$$

is absolutely continuous. This implies that, for almost all s, the derivative exists and is equal to $d_x G(t, x^0 + s(x^t - x^0), y; x^t - x^0)$ and that it is the integral of its derivative. In particular,

$$\boxed{G(t, x^t, y) = G(t, x^0, y) + \int_0^1 d_x G(t, x^0 + s(x^t - x^0), y; x^t - x^0)\, ds.} \tag{2.4}$$

(ii) For all $t \in [0, \tau]$, $x^0 \in X(0)$, $x^t \in X(t)$, $y \in Y$, $\varphi \in X$, and almost all $s \in (0, 1)$, $d_x G(t, x^0 + s(x^t - x^0), y; \varphi)$ exists
and the function $s \mapsto d_x G(t, x^0 + s(x^t - x^0), y; \varphi)$ belongs to $L^1(0, 1)$.

Standard adjoint at $t \geq 0$: to find $p^t \in Y$ such that $\forall \varphi \in X, \quad d_x G(t, x^t, p^t; \varphi) = 0$.

DEFINITION (K. STURM)

Given $x^0 \in X(0)$ and $x^t \in X(t)$, the *averaged adjoint state equation*:

$$\text{to find } y^t \in Y \text{ such that } \forall \varphi \in X, \quad \int_0^1 d_x G(t, x^0 + s(x^t - x^0), y^t; \varphi)\, ds = 0. \tag{2.5}$$

The set of solutions will be denoted $Y(t, x^0, x^t)$.
At $t = 0$, $Y(0, x^0, x^0)$ reduces to the set of *standard adjoint states*

$$Y(0, x^0) \stackrel{\text{def}}{=} \left\{ p^0 \in Y : \forall \varphi \in X,\ d_x G(0, x^0, p^0; \varphi) = 0 \right\}. \tag{2.6}$$

An *important consequence* of the introduction of the averaged adjoint state is the following identity: for all $x^0 \in X(0)$, $x^t \in X(t)$, and $y^t \in Y(t, x^0, x^t)$,

$$g(t) = G(t, x^t, 0) = G(t, x^t, y^t) = G(t, x^0, y^t). \tag{2.7}$$

Standard adjoint at $t \geq 0$: to find $p^t \in Y$ such that $\forall \varphi \in X, \quad d_x G(t, x^t, p^t; \varphi) = 0.$

DEFINITION (K. STURM)

Given $x^0 \in X(0)$ and $x^t \in X(t)$, the *averaged adjoint state equation*:

$$\text{to find } y^t \in Y \text{ such that } \forall \varphi \in X, \quad \int_0^1 d_x G(t, x^0 + s(x^t - x^0), y^t; \varphi)\, ds = 0. \tag{2.5}$$

The set of solutions will be denoted $Y(t, x^0, x^t)$.
At $t = 0$, $Y(0, x^0, x^0)$ reduces to the set of *standard adjoint states*

$$Y(0, x^0) \stackrel{\text{def}}{=} \left\{ p^0 \in Y : \forall \varphi \in X,\ d_x G(0, x^0, p^0; \varphi) = 0 \right\}. \tag{2.6}$$

An *important consequence* of the introduction of the averaged adjoint state is the following identity: for all $x^0 \in X(0)$, $x^t \in X(t)$, and $y^t \in Y(t, x^0, x^t)$,

$$g(t) = G(t, x^t, 0) = G(t, x^t, y^t) = G(t, x^0, y^t). \tag{2.7}$$

An *important consequence* of the introduction of the averaged adjoint state is the following identity: for all $x^0 \in X(0)$, $x^t \in X(t)$, and $y^t \in Y(t, x^0, x^t)$,

$$\boxed{g(t) = G(t, x^t, 0) = G(t, x^t, y^t) = G(t, x^0, y^t)} \tag{2.8}$$

$$g(0) = G(0, x^0, 0) = G(0, x^0, y^0). \tag{2.9}$$

As a result

$$g(t) - g(0) = G(t, x^0, y^t) - G(0, x^0, y^0)$$

$$\boxed{dg(0) = \lim_{t \searrow 0} \frac{g(t) - g(0)}{t} = \lim_{t \searrow 0} \frac{G(t, x^0, y^t) - G(0, x^0, y^0)}{t}.}$$

THEOREM (THESIS [STURM (2014)], SIAM [STURM (2015), THM. 3.1])

Consider the Lagrangian functional

$$(t, x, y) \mapsto G(t, x, y) : [0, \tau] \times X \times Y \to \mathbb{R}, \quad \tau > 0,$$

where X and Y are vector spaces and the function $y \mapsto G(t, x, y)$ is affine. Let (H0) and the following hypotheses be satisfied:

(H1) *for all $t \in [0, \tau]$, $g(t)$ is finite, $X(t) = \{x^t\}$ and $Y(t, x^0, x^t) = \{y^t\}$ are singletons;*

(H2) *$d_t G(t, x^0, y)$ exists for all $t \in [0, \tau]$ and all $y \in Y$;*

(H3) *the following limit exists*

$$\boxed{\lim_{s \searrow 0,\, t \searrow 0} d_t G(s, x^0, y^t) = d_t G(0, x^0, y^0).} \tag{2.10}$$

Then, $dg(0)$ exists and

$$dg(0) = d_t G(0, x^0, y^0).$$

Condition $(H3)$ is similar and typical of what can be found in the literature. See, for instance, [Correa-Seeger (1985)]).

PROOF.

From Hypothesis (H2), $d_t G(t, x^0, y)$ exists for all $t \in [0, \tau]$ and $y \in Y$. Hence, there exists $\theta_t \in (0,1)$ such that

$$G(t, x^0, y^t) - G(0, x^0, y^0) = G(0, x^0, y^t) + t\, d_t G(\theta_t t, x^0, y^t) - G(0, x^0, y^0)$$

$$= \underbrace{d_y G(0, x^0, 0; y^t - y^0)}_{=0} + t\, d_t G(\theta_t t, x^0, y^t) = t\, d_t G(\theta_t t, x^0, y^t)$$

$$\Rightarrow \frac{G(t, x^0, y^t) - G(0, x^0, y^0)}{t} = d_t G(\theta_t t, x^0, y^t)$$

since $d_y G(0, x^0, 0; y^t - y^0) = 0$. From hypothesis (H3)

$$\lim_{s \searrow 0,\, t \searrow 0} d_t G(s, x^0, y^t) = d_t G(0, x^0, y^0). \tag{2.11}$$

$$\Rightarrow \boxed{dg(0) = \lim_{t \searrow 0} \frac{G(t, x^0, y^t) - G(0, x^0, y^0)}{t} = d_t G(0, x^0, y^0).} \tag{2.12}$$

□

This is an extention of [Sturm (2014)] [Sturm (2015), Thm. 3.1] with only a local differentiability condition at $t = 0$. To our best knowledge, the *extra term* $R(0, x^0, y^0)$ is new. An example of a topological derivative will be given later.

THEOREM (SINGLETON CASE, [DELFOUR-STURM (2017), DELFOUR-STURM (2016)])

Consider the Lagrangian functional

$$(t, x, y) \mapsto G(t, x, y) : [0, \tau] \times X \times Y \to \mathbb{R}, \quad \tau > 0,$$

where X and Y are vector spaces and the function $y \mapsto G(t, x, y)$ is affine. Let (H0) and the following hypotheses be satisfied:

(H1) *for all $t \in [0, \tau]$, $g(t)$ is finite, $X(t) = \{x^t\}$ and $Y(t, x^0, x^t) = \{y^t\}$ are singletons;*

(H2) *$d_t G(0, x^0, y^0)$ exists;*

(H3) *the following limit exists*

$$R(0, x^0, y^0) \overset{\text{def}}{=} \lim_{t \searrow 0} d_y G\left(t, x^0, 0; \frac{y^t - y^0}{t}\right). \tag{2.13}$$

Then, $dg(0)$ exists and

$$dg(0) = d_t G(0, x^0, y^0) + R(0, x^0, y^0).$$

PROOF.

Recalling that $g(t) = G(t, x^t, y^t) = G(t, x^0, y^t)$,

$$\begin{aligned} g(t) - g(0) &= G(t, x^0, y^t) - G(0, x^0, y^0) \\ &= G(t, x^0, y^0) + d_y G(t, x^0, 0; y^t - y^0) - G(0, x^0, y^0) \end{aligned}$$

$$\Rightarrow \quad \frac{g(t) - g(0)}{t} = d_y G\left(t, x^0, 0; \frac{y^t - y^0}{t}\right) + \frac{G(t, x^0, y^0) - G(0, x^0, y^0)}{t}$$

$$\Rightarrow \; dg(0) = \lim_{t \searrow 0} d_y G\left(t, x^0, 0; \frac{y^t - y^0}{t}\right) + d_t G(0, x^0, y^0)$$

from hypotheses (H2) and (H3). □

Condition (H3) is optimal since under hypotheses (H1)

$$dg(0) \text{ exists} \iff \lim_{t \searrow 0} d_y G\left(t, x^0, 0; \frac{y^t - y^0}{t}\right) \text{ exists}$$

Hypotheses (H2) and (H3) are weaker and more general than (H2) and (H3).

(H2) It is only assumed that $d_t G(0, x^0, y^0)$ exists.
Hypothesis (H2) assumes that $d_t G(t, x^0, y)$ exists for all $t \in [0, \tau]$ and $y \in Y$.

(H3) Hypothesis (H3) assumes that

$$\boxed{\lim_{s \searrow 0,\, t \searrow 0} d_t G(s, x^0, y^t) = d_t G(0, x^0, y^0)}. \tag{2.14}$$

wich implies

$$\boxed{R(0, x^0, y^0) = \lim_{t \searrow 0} d_y G\left(t, x^0, 0; \frac{y^t - y^0)}{t}\right) = 0.} \tag{2.15}$$

Hence, condition (H3) with $R(0, x^0, y^0) = 0$ is weaker and potentially more general (when the limit is not zero) than (H3).

All this is possible since $G(t, x, y)$ is a Lagrangian. For zero-sum games, condition $(H3)$ and a similar condition for the max min would not be as interesting.

Recalling that $g(t) = G(t, x^t, y)$ and $g(0) = G(0, x^0, y)$ for any $y \in Y$, then for the standard adjoint state p^0 at $t = 0$

$$g(t) - g(0) = G(t, x^t, p^0) - G(t, x^0, p^0) + \left(G(t, x^0, p^0) - G(0, x^0, p^0)\right).$$

Dividing by $t > 0$

$$\begin{aligned}\frac{g(t) - g(0)}{t} &= \frac{G(t, x^t, p^0) - G(t, x^0, p^0)}{t} + \frac{G(t, x^0, p^0) - G(0, x^0, p^0)}{t} \\ &= \int_0^1 d_x G\left(t, (1-\theta)x^0 + \theta x^t, p^0; \frac{x^t - x^0}{t}\right) d\theta + \frac{G(t, x^0, p^0) - G(0, x^0, p^0)}{t}.\end{aligned}$$

Therefore, in view of hypothesis (H2), the limit $dg(0)$ exists if and only if the limit of the first term exists

$$\Rightarrow dg(0) = \lim_{t \searrow 0} \int_0^1 d_x G\left(t, (1-\theta)x^0 + \theta x^t, p^0; \frac{x^t - x^0}{t}\right) d\theta + d_t G(0, x^0, p^0)$$

and the existence of the limit of the first term can replace hypothesis (H3). As a result, we have two ways of expression hypothesis (H3) since

$$\boxed{\lim_{t \searrow 0} \int_0^1 d_x G\left(t, (1-\theta)x^0 + \theta x^t, p^0; \frac{x^t - x^0}{t}\right) d\theta = \lim_{t \searrow 0} d_y G\left(t, x^0, 0; \frac{y^t - y^0}{t}\right).}$$

Since $d_X G$ and $d_x d_y G$ both exist, Hypothesis (H3) can be rewritten as follows

$$d_y G\left(t, x^0, 0; \frac{y^t - y^0}{t}\right) = d_y G\left(t, x^0, 0; \frac{y^t - y^0}{t}\right) - d_y G\left(t, x^t, 0; \frac{y^t - y^0}{t}\right)$$
$$= \int_0^1 d_x d_y G\left(t, \theta x^0 + (1-\theta)x^t, 0; \frac{y^t - y^0}{t^\alpha}; \frac{x^0 - x^t}{t^{1-\alpha}}\right) d\theta,$$

for some $\alpha \in [0,1]$. For instance with $\alpha = 1/2$, it would be sufficient to find bounds on the differential quotients

$$\frac{y^t - y^0}{t^{1/2}} \text{ and } \frac{x^t - x^0}{t^{1/2}}$$

which is less demanding than finding a bound on $(x^t - x^0)/t$ or $(y^t - y^0)/t$.

When the integral can be taken inside

$$d_y G\left(t, x^0, 0; \frac{y^t - y^0}{t}\right) = d_x d_y G\left(t, \frac{x^0 + x^t}{2}, 0; \frac{y^t - y^0}{t^\alpha}; \frac{x^0 - x^t}{t^{1-\alpha}}\right)$$

$$\boxed{\lim_{t \searrow 0} d_y G\left(t, x^0, 0; \frac{y^t - y^0}{t}\right) = \lim_{t \searrow 0} d_x d_y G\left(t, \frac{x^0 + x^t}{2}, 0; \frac{y^t - y^0}{t^\alpha}; \frac{x^0 - x^t}{t^{1-\alpha}}\right).}$$

If we are only interested in a descent method, we can obtain the semidifferential of $f(a)$ by a similar minimax formulation.

Given the direction $b \in L^2(\Omega)$, we want to compute $df(a;b) = \lim_{t\searrow 0}(f(a+tb) - f(a))/t$.

The *state* $u^t \in H_0^1(\Omega)$ at $t > 0$ is solution of

$$\int_\Omega \nabla u^t \cdot \nabla\psi - (a+tb)\,\psi\,dx = 0, \quad \forall\psi \in H_0^1(\Omega), \tag{2.16}$$

and the *associated Lagrangian* is

$$L(t,\varphi,\psi) \overset{\text{def}}{=} \int_\Omega \frac{1}{2}|\varphi - g|^2\,dx + \int_\Omega \nabla\varphi\cdot\nabla\psi - (a+tb)\,\psi\,dx.$$

It is readily seen that

$$g(t) \overset{\text{def}}{=} f(a+tb) = \inf_{\varphi\in H_0^1(\Omega)}\ \sup_{\psi\in H_0^1(\Omega)} L(t,\varphi,\psi)$$

$$dg(0) \overset{\text{def}}{=} \lim_{t\searrow 0}\frac{g(t)-g(0)}{t} = df(a;b).$$

Recall

$$L(t,\varphi,\psi) \stackrel{\text{def}}{=} \int_\Omega \frac{1}{2}|\varphi - g|^2 + \nabla\varphi\cdot\nabla\psi - (a+tb)\,\psi\,dx.$$

It is readily seen that

$$d_y L(t,\varphi,\psi;\psi') = \int_\Omega \nabla\varphi\cdot\nabla\psi' - (a+tb)\,\psi'\,dx$$

$$d_x L(t,\varphi,\psi;\varphi') = \int_\Omega (\varphi-g)\,\varphi' + \nabla\varphi'\cdot\nabla\psi\,dx, \quad d_t L(t,\varphi,\psi) = -\int_\Omega b\,\psi\,dx.$$

Observe that the *derivative of the state* $\dot{u} \in H_0^1(\Omega)$ exists:

$$\int_\Omega \nabla\left(\frac{u^t - u^0}{t}\right)\cdot\nabla\psi - b\,\psi\,dx = 0, \quad \forall\psi\in H_0^1(\Omega), \tag{2.17}$$

implies that $(u^t - u^0)/t = \dot{u} \in H_0^1(\Omega)$ solution of

$$\int_\Omega \nabla\dot{u}\cdot\nabla\psi - b\,\psi\,dx = 0, \quad \psi\in H_0^1(\Omega). \tag{2.18}$$

The *averaged adjoint* $y^t \in H_0^1(\Omega)$ is solution of

$$\int_\Omega \left(\frac{u^t + u^0}{2}\right)\varphi + \nabla y^t\cdot\nabla\varphi\,dx = 0, \quad \forall\varphi\in H_0^1(\Omega).$$

$$\int_\Omega \left(\frac{u^t+u^0}{2}\right)\varphi + \nabla y^t\cdot\nabla\varphi\,dx = 0, \quad \forall\varphi\in H_0^1(\Omega).$$

$$\text{adjoint at } t=0: \int_\Omega u^0\,\varphi + \nabla y^0\cdot\nabla\varphi\,dx = 0, \quad \forall\varphi\in H_0^1(\Omega),$$

$$\Rightarrow \int_\Omega \frac{1}{2}\left(\frac{u^t-u^0}{t}\right)\varphi + \nabla\left(\frac{y^t-y^0}{t}\right)\cdot\nabla\varphi\,dx = 0, \quad \forall\varphi\in H_0^1(\Omega). \tag{2.19}$$

It remains to check that the limit in (2.13) exists: $d_yG(t,x^0,0;(y^t-y^0)/t)\to 0$

$$\int_\Omega \nabla u^0\cdot\nabla\left(\frac{y^t-y^0}{t}\right) - (a+tb)\left(\frac{y^t-y^0}{t}\right)dx = -t\int_\Omega b\left(\frac{y^t-y^0}{t}\right)dx$$

$$= -t\int_\Omega \nabla\left(\frac{u^t-u^0}{t}\right)\cdot\nabla\left(\frac{y^t-y^0}{t}\right)dx = \frac{t}{2}\int_\Omega \left|\frac{u^t-u^0}{t}\right|^2 dx = \frac{t}{2}\int_\Omega |\dot{u}|^2\,dx \to 0$$

as $t\to 0$ using (2.18) and (2.19). Therefore, by Theorem 9,

$$df(a;b) = -\int_\Omega b\,y^0\,dx, \quad y^0\in H_0^1(\Omega), \tag{2.20}$$

$$\int_\Omega (u-g)\varphi + \nabla y^0\cdot\nabla\varphi\,dx = 0, \quad \forall\varphi\in H_0^1(\Omega). \tag{2.21}$$

Outline

The *topological derivative* rigorously introduced by [Sokołowski-Żochowski (1999)] induces topological changes.

For instance, let f be an *objective function* defined on a family of open subsets of $\mathbb{R}^N$. Given a point a in the open set Ω, let $\overline{B}_r(a)$ be a closed ball of radius r and center a such that $\overline{B}_r(a) \subset \Omega$.

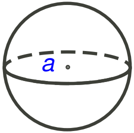
Consider the perturbed domain $\Omega_r \overset{\text{def}}{=} \Omega \backslash \overline{B}_r(a)$: Ω minus the hole $\overline{B}_r(a)$. In this simple case the *topological derivative* is defined as

$$df(0) \overset{\text{def}}{=} \lim_{r \searrow 0} \frac{f(\Omega_r) - f(\Omega)}{|\overline{B}_r(a)|}, \tag{3.1}$$

where $|\overline{B}_r(a)|$ is the volume of $\overline{B}_r(a)$ in $\mathbb{R}^N$.

When f is of the form $f(\Omega) = \int_\Omega \varphi \, dx$, the application of the Lebesgue differentiation theorem gives $df(0) = -\varphi(a)$. Of course, many other types of topological perturbations can be considered (see the recent IFIP paper of [Delfour (2017)]).

$\overline{B}_r(a)$



Outline

Example of a Topological Derivative [Delfour-Sturm (2016)]
Problem Formulation

Given ε, $0 < \varepsilon < 1$, $a > 0$, and the domain $\Omega = (-a, a)$, consider the problem: to find $u \in W^{1,2-\varepsilon}(-a, a)$ such that

$$\forall \varphi \in W^{1,\frac{2-\varepsilon}{1-\varepsilon}}(-a,a) \quad \boxed{\int_{-a}^{a} \frac{du}{dx}\frac{d\varphi}{dx} + u\,\varphi\,dx = \int_{-a}^{a} \frac{d}{dx}\sqrt{|x|}\,\frac{d\varphi}{dx} + \sqrt{|x|}\,\varphi\,dx.} \tag{3.2}$$

Here, $X = W^{1,2-\varepsilon}(-a,a)$ and $Y = W^{1,\frac{2-\varepsilon}{1-\varepsilon}}(-a,a)$ are reflexive Banach spaces since $2 - \varepsilon > 1$ and $\frac{2-\varepsilon}{1-\varepsilon} > 1$. The elements of X will be denoted u and $x \in (-a, a)$ will be the space variable. There exists a unique[1] solution $u(x) = \sqrt{|x|}$, $-a \le x \le a$, and the injections $W^{1,2-\varepsilon}(-a,a) \to C^0[-a,a]$ and $W^{1,\frac{2-\varepsilon}{1-\varepsilon}}(-a,a) \to C^0[-a,a]$ are continuous and the following *objective function* is well-defined:

$$\boxed{f(\Omega) \overset{\text{def}}{=} |u(a)|^2 + |u(-a)|^2 - 2\,|u(0)|^2.}$$

[1] Given measurable functions $k_1, k_2 : [-a,a] \to \mathbb{R}$ such that $\alpha \le k_i(x) \le \beta$ for some constants $\alpha > 0$ and $\beta > 0$, and real numbers $1 < p < \infty$, $p^{-1} + q^{-1} = 1$, associate with the continuous bilinear mapping

$$\boxed{\varphi, \psi \mapsto b(\varphi,\psi) \overset{\text{def}}{=} \int_{-a}^{a} k_1(x)\frac{d\varphi}{dx}\frac{d\psi}{dx} + k_2(x)\,\varphi\,\psi\,dx : W^{1,p}(-a,a) \times W^{1,q}(-a,a) \to \mathbb{R},}$$

the continuous linear operator $A : W^{1,p}(-a,a) \to W^{1,q}(-a,a)'$ which is a topological isomorphism for all $p \in (1,\infty)$ ([Auscher-Tchamitchian (1998)]). Here, $p = 2 - \varepsilon$ and $q = (2 - \varepsilon)/(1 - \varepsilon)$.

Given ε, $0 < \varepsilon < 1$, $a > 0$, and the domain $\Omega = (-a, a)$, consider the problem: to find $u \in W^{1,2-\varepsilon}(-a, a)$ such that

$$\forall \varphi \in W^{1,\frac{2-\varepsilon}{1-\varepsilon}}(-a,a) \quad \boxed{\int_{-a}^{a} \frac{du}{dx}\frac{d\varphi}{dx} + u\,\varphi\,dx = \int_{-a}^{a} \frac{d}{dx}\sqrt{|x|}\,\frac{d\varphi}{dx} + \sqrt{|x|}\,\varphi\,dx.} \tag{3.2}$$

Here, $X = W^{1,2-\varepsilon}(-a, a)$ and $Y = W^{1,\frac{2-\varepsilon}{1-\varepsilon}}(-a, a)$ are reflexive Banach spaces since $2 - \varepsilon > 1$ and $\frac{2-\varepsilon}{1-\varepsilon} > 1$. The elements of X will be denoted u and $x \in (-a, a)$ will be the space variable.There exists a unique[1] solution $u(x) = \sqrt{|x|}$, $-a \leq x \leq a$, and the injections $W^{1,2-\varepsilon}(-a, a) \to C^0[-a, a]$ and $W^{1,\frac{2-\varepsilon}{1-\varepsilon}}(-a, a) \to C^0[-a, a]$ are continuous and the following *objective function* is well-defined:

$$\boxed{f(\Omega) \overset{\text{def}}{=} |u(a)|^2 + |u(-a)|^2 - 2\,|u(0)|^2.}$$

[1] Given measurable functions $k_1, k_2 : [-a, a] \to \mathbb{R}$ such that $\alpha \leq k_i(x) \leq \beta$ for some constants $\alpha > 0$ and $\beta > 0$, and real numbers $1 < p < \infty$, $p^{-1} + q^{-1} = 1$, associate with the continuous bilinear mapping

$$\boxed{\varphi, \psi \mapsto b(\varphi, \psi) \overset{\text{def}}{=} \int_{-a}^{a} k_1(x)\frac{d\varphi}{dx}\frac{d\psi}{dx} + k_2(x)\,\varphi\,\psi\,dx : W^{1,p}(-a, a) \times W^{1,q}(-a, a) \to \mathbb{R},}$$

the continuous linear operator $A : W^{1,p}(-a, a) \to W^{1,q}(-a, a)'$ which is a topological isomorphism for all $p \in (1, \infty)$ ([Auscher-Tchamitchian (1998)]). Here, $p = 2 - \varepsilon$ and $q = (2 - \varepsilon)/(1 - \varepsilon)$.

$\overline{B}_r(0) \subset \mathbb{R}$ be the closed ball of radius r in 0. Volume: $t = |\overline{B}_r(0)| = 2r$. Perturbed domain is $\Omega_r = \Omega \backslash \overline{B}_r(0) = (-a, -r) \cup (r, a)$ has 2 connected components and it is not possible to construct a bijection between Ω and Ω_r.

Perturbed problems parametrized by r, $0 < r < a/2$: to find $u_r \in W^{1,2-\varepsilon}(\Omega_r)$ s. t.

$$\forall \varphi \in W^{1,\frac{2-\varepsilon}{1-\varepsilon}}(\Omega_r), \quad \int_{\Omega_r} \frac{du_t}{dx}\frac{d\varphi}{dx} + u_t\,\varphi\, dx = \int_{\Omega_r} \frac{d\sqrt{|x|}}{dx}\frac{d\varphi}{dx} + \sqrt{|x|}\,\varphi\, dx$$

with the objective function

$$j(r) \overset{\text{def}}{=} |u_r(a)|^2 - |u_r(r)|^2 + |u_r(-a)^2 - |u_r(-r)|^2.$$

The function $u_r(x) = \sqrt{|x|}$ is the unique solution and

$$j(r) = 2a - 2r \quad \Rightarrow dj(0) \overset{\text{def}}{=} \lim_{r \searrow 0} \frac{1}{2r}\left(j(r) - j(0)\right) = -1.$$

By construction, $\Omega_r = T_r(\Omega \backslash \{0\})$, where

$$\text{Bijection } x \mapsto T_r(x) \overset{\text{def}}{=} \begin{cases} x - r\left(1 + \dfrac{x}{a}\right), & x \in (-a, 0) \\ x + r\left(1 - \dfrac{x}{a}\right), & x \in (0, a) \end{cases} : \Omega \backslash \{0\} \to \Omega_r = \Omega \backslash \overline{B}_r(0)$$

and notice that $T_r(0^-) = -r$, $T_r(0^+) = r$, $T_r(a^-) = a$, and $T_r(-a^+) = -a$.

$\overline{B}_r(0) \subset \mathbb{R}$ be the closed ball of radius r in 0. Volume: $t = |\overline{B}_r(0)| = 2r$. Perturbed domain is $\Omega_r = \Omega \backslash \overline{B}_r(0) = (-a, -r) \cup (r, a)$ has 2 connected components and it is not possible to construct a bijection between Ω and Ω_r.

Perturbed problems parametrized by r, $0 < r < a/2$: to find $u_r \in W^{1,2-\varepsilon}(\Omega_r)$ s. t.

$$\forall \varphi \in W^{1,\frac{2-\varepsilon}{1-\varepsilon}}(\Omega_r), \quad \int_{\Omega_r} \frac{du_t}{dx}\frac{d\varphi}{dx} + u_t\,\varphi\,dx = \int_{\Omega_r} \frac{d\sqrt{|x|}}{dx}\frac{d\varphi}{dx} + \sqrt{|x|}\,\varphi\,dx$$

with the objective function

$$j(r) \overset{\text{def}}{=} |u_r(a)|^2 - |u_r(r)|^2 + |u_r(-a)^2 - |u_r(-r)|^2.$$

The function $u_r(x) = \sqrt{|x|}$ is the unique solution and

$$j(r) = 2a - 2r \quad \Rightarrow dj(0) \overset{\text{def}}{=} \lim_{r \searrow 0} \frac{1}{2r}(j(r) - j(0)) = -1.$$

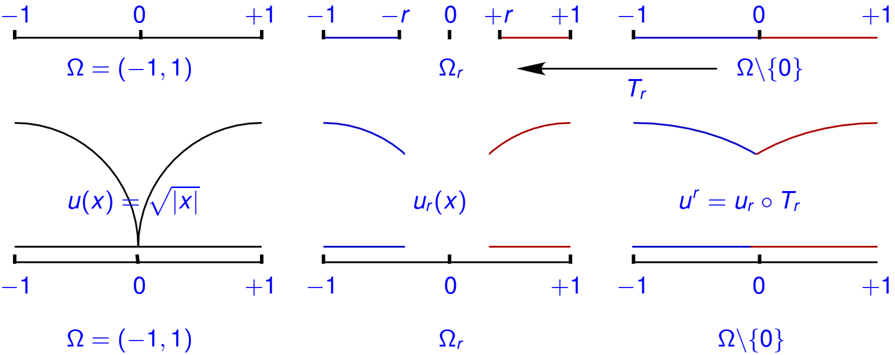
By construction, $\Omega_r = T_r(\Omega \backslash \{0\})$, where

$$\text{Bijection } x \mapsto T_r(x) \overset{\text{def}}{=} \left\{ \begin{array}{ll} x - r\left(1 + \dfrac{x}{a}\right), & x \in (-a, 0) \\ x + r\left(1 - \dfrac{x}{a}\right), & x \in (0, a) \end{array} \right\} : \Omega \backslash \{0\} \to \Omega_r = \Omega \backslash \overline{B}_r(0)$$

and notice that $T_r(0^-) = -r$, $T_r(0^+) = r$, $T_r(a^-) = a$, and $T_r(-a^+) = -a$.

for $a = 1$

the function u has a cusp at the point 0



Prior to proceeding, it is advantageous to simplify the computations by observing that the function $u^r(x) = \sqrt{|T_r(x)|}$ is symmetrical with respect to $x = 0$, that is, $u^r(-x) = u^r(x)$ and

$$j(r) = 2\left[u^r(a)^2 - u^r(0^+)^2\right]. \tag{3.3}$$

As a result

$$dj(0) = \lim_{r \searrow 0} \frac{j(r) - j(0)}{2r} = \lim_{r \searrow 0} \frac{u^r(a)^2 - u^r(0^+)^2}{r}$$

By changing the variable r to t, it is sufficient to apply Theorem 9 to the following problem on $(0, a)$: to find $u^t \in W^{1,2-\varepsilon}(0, a)$ such that for all $\varphi \in W^{1,\frac{2-\varepsilon}{1-\varepsilon}}(0, a)$

$$\boxed{\int_0^a \frac{a}{a-t}\frac{du^t}{dx}\frac{d\varphi}{dx} + \frac{a-t}{a}u^t\varphi\, dx = \int_0^a \frac{1}{2\sqrt{T_t(x)}}\frac{d\varphi}{dx} + \frac{a-t}{a}\sqrt{T_t(x)}\,\varphi\, dx} \tag{3.4}$$

with the objective function

$$\boxed{j^+(t) \overset{\text{def}}{=} u^t(a)^2 - u^t(0^+)^2,} \quad dj^+(0) = \lim_{t \searrow 0}(j^+(t) - j^+(0))/t.$$

From Theorem 9, the Lagrangian associated with the perturbed problems is

$$\begin{aligned} G(t,\varphi,\psi) \overset{\text{def}}{=} & |\varphi(a)|^2 - |\varphi(0)|^2 \\ & + \int_0^a \left(\frac{a}{a-t}\right) \frac{d\varphi}{dx}\,\frac{d\psi}{dx} + \left(\frac{a-t}{a}\right) \varphi\,\psi\,dx \\ & - \int_0^a \frac{1}{2\sqrt{T_t(x)}}\,\frac{d\psi}{dx} + \left(\frac{a-t}{a}\right)\sqrt{T_t(x)}\,\psi\,dx. \end{aligned} \tag{3.5}$$

It is non-convex in the φ variable in view of the presence of the term $-|\varphi(0)|^2$. The standard adjoint p^t is solution of the adjoint equation

$$\forall \varphi \in W^{1,2-\varepsilon}(0,a), \quad \begin{cases} 2u^t(a)\,\varphi(a) - 2u^t(0)\,\varphi(0) \\ + \displaystyle\int_0^a \left(\frac{a}{a-t}\right)\frac{d\varphi}{dx}\,\frac{dp^t}{dx} + \left(\frac{a-t}{a}\right)\varphi\,p^t\,dx = 0. \end{cases} \tag{3.6}$$

In particular this is true for all $\varphi \in H^1(0,a) = W^{1,2}(0,a) \subset W^{1,2-\varepsilon}(0,a)$. Since the differential operator is uniformly coercive for $0 \le t \le a/2$, there exist a unique $p^t \in H^1(0,a)$.

But, in view of the fact that for $0 \leq t \leq a/2$, u^t is finite for all x, we get more regularity: $p_t \in H^2(0,a) \cap C^\infty(0,a)$ is solution of

$$
\begin{aligned}
&-\frac{a}{a-t}\frac{d^2p^t}{dx^2} + \frac{a-t}{a}p^t = 0 \text{ in } (0,a)\\
&\frac{a}{a-t}\frac{dp^t}{dx}(a) = -2\,u^t(a), \quad \frac{a}{a-t}\frac{dp^t}{dx}(0) = -2\,u^t(0).
\end{aligned}
$$

The explicit solution for $u^t(0) = \sqrt{t}$ and $u^t(a) = \sqrt{a}$ is

$$
p^t(x) = \frac{a}{a-t}\frac{2}{e^{a-t} - e^{-(a-t)}}\left[\sqrt{t}\left(e^{\frac{a-t}{a}(a-x)} + e^{-\frac{a-t}{a}(a-x)}\right) - \sqrt{a}\left(e^{\frac{a-t}{a}x} + e^{-\frac{a-t}{a}x}\right)\right].
$$

At $t = 0$,

$$
p^0(x) = -2\sqrt{a}\,\frac{e^x + e^{-x}}{e^a - e^{-a}}.
$$

The right-hand side *t*-derivative is

$$\begin{aligned} d_tG(t,\varphi,\psi) = &\int_0^a \frac{a}{(a-t)^2}\frac{d\varphi}{dx}\frac{d\psi}{dx} - \frac{1}{a}\varphi\,\psi\,dx + \frac{1}{a}\int_0^a \sqrt{T_t}\,\psi\,dx \\ &- \frac{1}{2}\int_0^a \left[\frac{-1}{2(T_t)^{3/2}}\frac{d\psi}{dx} + \left(\frac{a-t}{a}\right)\frac{1}{\sqrt{T_t}}\psi\right] dx \\ &+ \frac{1}{2a}\int_0^a \left[\frac{-x}{2(T_t)^{3/2}}\frac{d\psi}{dx} + \left(\frac{a-t}{a}\right)\frac{x}{\sqrt{T_t}}\psi\right] dx. \end{aligned}$$

At $t=0$, substitute $u^0(x)=\sqrt{x}$ and p^0 and Integrate by parts

$$\begin{aligned} d_tG(0,u^0,p^0) = &\int_0^a \frac{1}{a}\frac{d\sqrt{x}}{dx}\frac{dp^0}{dx} - \frac{1}{a}\sqrt{x}\,p^0\,dx + \frac{1}{a}\int_0^a \sqrt{x}\,p^0\,dx \\ &- \frac{1}{2}\int_0^a \left[\frac{-1}{2x^{3/2}}\frac{dp^0}{dx} + \frac{1}{\sqrt{x}}p^0\right] dx + \frac{1}{2a}\int_0^a \left[\frac{-1}{2\sqrt{x}}\frac{dp^0}{dx} + \sqrt{x}p^0\right] dx \\ = &\frac{1}{2a}\int_0^a \frac{d\sqrt{x}}{dx}\frac{dp^0}{dx} + \sqrt{x}p^0\,dx - \frac{1}{2}\int_0^a \left[\frac{d}{dx}\frac{1}{\sqrt{x}}\frac{dp^0}{dx} + \frac{1}{\sqrt{x}}p^0\right] dx = 0. \end{aligned}$$

Go back to the Lagrangian (3.5) of the perturbed problem and compute

$$d_x G(t, \bar{\varphi}, \psi; \varphi) = 2\,\bar{\varphi}(a)\,\varphi(a) - 2\,\bar{\varphi}(0)\,\varphi(0) + \int_0^a \left(\frac{a}{a-t}\right) \frac{d\varphi}{dx}\,\frac{d\psi}{dx} + \left(\frac{a-t}{a}\right) \varphi\,\psi\,dx.$$

The averaged adjoint state equation for y^t must satisfy the equation: for all $\varphi \in W^{1,2-\varepsilon}(0,a)$

$$0 = \int_0^1 d_x G(t, u^0 + s(u^t - u^0), y^t; \varphi)$$

$$= (u^t(a) + u^0(a))\,\varphi(a) - (u^t(0) + u^0(0))\,\varphi(0) + \int_0^a \left(\frac{a}{a-t}\right) \frac{d\varphi}{dx}\,\frac{dy^t}{dx} + \left(\frac{a-t}{a}\right) \varphi\,y^t\,dx.$$

Its solution $y^t \in H^2(0,a) \cap C^\infty(0,a)$ satisfies the following equations

averaged adjoint state

$$-\left(\frac{a}{a-t}\right)\frac{d^2 y^t}{dx^2} + \left(\frac{a-t}{a}\right) y^t = 0, \quad \text{in } (0,a)$$

$$\left(\frac{a}{a-t}\right)\frac{dy^t}{dx}(0) = -(u^t(0) + u^0(0)) \text{ at } x = 0$$

$$\left(\frac{a}{a-t}\right)\frac{dy^t}{dx}(a) = -(u^t(a) + u^0(a)) \text{ at } x = a.$$

Its explicit expression with $u^t(0) = \sqrt{t}$ and $u^t(a) = \sqrt{a}$ is

$$y^t(x) = \frac{a}{a-t}\frac{1}{e^{a-t} - e^{-(a-t)}}\left[\sqrt{t}\left(e^{\frac{a-t}{a}(a-x)} + e^{-\frac{a-t}{a}(a-x)}\right) - 2\sqrt{a}\left(e^{\frac{a-t}{a}x} + e^{-\frac{a-t}{a}x}\right)\right].$$

The condition to be checked is the existence of the limit (the extra term)

$$\boxed{\lim_{t\searrow 0} d_y G\left(t, u^0, 0; \frac{y^t - y^0}{t}\right).}$$

So, for $\psi = (y^t - y^0)/t \in H^2(0,a)$,

$$d_y G(t, u^0, 0; \psi)$$

$$= \int_0^a \left(\frac{a}{a-t}\right)\frac{du^0}{dx}\frac{d\psi}{dx} + \left(\frac{a-t}{a}\right) u^0\,\psi\,dx - \int_0^a \frac{1}{2\sqrt{T_t(x)}}\frac{d\psi}{dx} + \left(\frac{a-t}{a}\right)\sqrt{T_t(x)}\,\psi\,dx$$

$$= \int_0^a \left(\frac{a}{a-t}\right)\frac{du^0}{dx}\frac{d\psi}{dx} + \left(\frac{a-t}{a}\right) u^0\,\psi\,dx - \int_0^a \frac{a}{a-t}\frac{d\sqrt{T_t(x)}}{dx}\frac{d\psi}{dx} + \frac{a-t}{a}\sqrt{T_t(x)}\,\psi\,dx$$

$$= \int_0^a \left(\frac{a}{a-t}\right)\frac{d(u^0 - u^t)}{dx}\frac{d\psi}{dx} + \left(\frac{a-t}{a}\right)(u^0 - u^t)\,\psi\,dx$$

$$= (u^0 - u^t)\left(\frac{a}{a-t}\right)\frac{d}{dx}\left(\frac{y^t - y^0}{t}\right)\Bigg|_{x=0}^{a} \to -1.$$

Its explicit expression with $u^t(0)=\sqrt{t}$ and $u^t(a)=\sqrt{a}$ is

$$y^t(x)=\frac{a}{a-t}\frac{1}{e^{a-t}-e^{-(a-t)}}\left[\sqrt{t}\left(e^{\frac{a-t}{a}(a-x)}+e^{-\frac{a-t}{a}(a-x)}\right)-2\sqrt{a}\left(e^{\frac{a-t}{a}x}+e^{-\frac{a-t}{a}x}\right)\right].$$

The condition to be checked is the existence of the limit (the extra term)

$$\boxed{\lim_{t\searrow 0} d_yG\left(t,u^0,0;\frac{y^t-y^0}{t}\right).}$$

So, for $\psi=(y^t-y^0)/t\in H^2(0,a)$,

$$d_yG(t,u^0,0;\psi)$$

$$=\int_0^a\left(\frac{a}{a-t}\right)\frac{du^0}{dx}\frac{d\psi}{dx}+\left(\frac{a-t}{a}\right)u^0\,\psi\,dx-\int_0^a\frac{1}{2\sqrt{T_t(x)}}\frac{d\psi}{dx}+\left(\frac{a-t}{a}\right)\sqrt{T_t(x)}\,\psi\,dx$$

$$=\int_0^a\left(\frac{a}{a-t}\right)\frac{du^0}{dx}\frac{d\psi}{dx}+\left(\frac{a-t}{a}\right)u^0\,\psi\,dx-\int_0^a\frac{a}{a-t}\frac{d\sqrt{T_t(x)}}{dx}\frac{d\psi}{dx}+\frac{a-t}{a}\sqrt{T_t(x)}\,\psi\,dx$$

$$=\int_0^a\left(\frac{a}{a-t}\right)\frac{d(u^0-u^t)}{dx}\frac{d\psi}{dx}+\left(\frac{a-t}{a}\right)(u^0-u^t)\,\psi\,dx$$

$$=(u^0-u^t)\left(\frac{a}{a-t}\right)\frac{d}{dx}\left(\frac{y^t-y^0}{t}\right)\Big|_{x=0}^{a}\to -1.$$

THEOREM

(i) *For* $0 < \varepsilon < 1$ *and* $t \in [0, a/2]$,

$$\|u^t\|_{W^{1,2-\varepsilon}(0,a)} \leq c(\varepsilon, a), \quad \|u^t - u^0\|_{C^0[0,a]} \leq \sqrt{t}, \quad u^t \to u^0 \text{ in } W^{1,2-\varepsilon}(0,a)\text{-weak}$$

and this rate of convergence is sharp.

(ii) *For* $x \in (0, a)$, *the material derivative is given by*

$$\dot{u}(x) \overset{\text{def}}{=} \lim_{t \searrow 0} \frac{u^t(x) - u^0(x)}{t} = \frac{1}{2}\left(\frac{1}{\sqrt{|x|}} - \frac{\sqrt{|x|}}{a}\right) \geq 0,$$

$\dot{u} \in L^{2-\varepsilon}(0, a)$ *for* $0 < \varepsilon \leq 1$, *but* $\dot{u} \notin L^2(0, a)$. *Moreover, as* $t \to 0$,

$$\left\|(u^t - u^0)/t - \dot{u}\right\|_{L^{2-\varepsilon}(0,a)} \to 0. \tag{3.7}$$

(iii) *As for the derivative of* $\dot{u}$,

$$\frac{d\dot{u}}{dx}(x) = -\frac{1}{4\sqrt{|x|}} \begin{cases} 1/x + 1/a, & x \in (0, a) \\ 1/x - 1/a, & x \in (-a, 0) \end{cases} \quad \frac{d\dot{u}}{dx}(0^+) = -\infty,$$

Therefore,

$$\frac{d\dot{u}}{dx} \notin L^1(0,a), \text{ and, a fortiori, } \frac{d\dot{u}}{dx} \notin L^{2-\varepsilon}(0,a).$$

From part (iii) we cannot apply the chain rule to get $dj(0^+)$ since the expression is undetermined:

$$2\,u^0(a)\,\dot{u}(a) - 2\,u^0(0)\,\dot{u}(0) = 2\,u^0(a)\,\dot{u}(a) - 2\,[0\,(-\infty)]\,!$$

OUTLINE

We give two theorems for the existence and expressions of $dg(0)$ in the multivalued case where only a right-hand side derivative of g is expected.

- New conditions and quadratic examples were given in [Delfour-Sturm (2017)] without the extra term.

- Complete conditions including the extra term were published in [Delfour-Sturm (2016)] at an IFAC meeting in 2016 prior to the publication of [Delfour-Sturm (2017)] due to longer publication delays in the Journal of Convex Analysis.

Here, we give the latest version from [Delfour-Sturm (2016)].

The first theorem is a mild generalization of the singleton case. Yet, it can be applied to PDE problems with non-homogeneous Dirichlet boundary conditions where non-unique extensions are used (cf. [Delfour-Zolésio (2011)]).

A new non-convex multivalued example will be given for the second more general theorem.

Outline

THEOREM (A FIRST EXTENSION)

Given X, Y, and G, let (H0) *and the following hypotheses be satisfied:*

(H1) *for all t in $[0,\tau]$, $X(t) \neq \varnothing$ and $g(t)$ is finite, and for all $x^t \in X(t)$ and $x^0 \in X(0)$, $Y(t, x^0, x^t) \neq \varnothing$;*

(H2) *for all $x \in X(0)$ and $y \in Y(0, x)$, $d_t G(0, x, y)$ exists;*

(H3) *there exist $\hat{x}^0 \in X(0)$, $\hat{y}^0 \in Y(0, \hat{x}^0)$, and $R(0, \hat{x}^0, \hat{y}^0)$ such that for each sequence $t_n \to 0$, $0 < t_n \leq \tau$, there exist a subsequence $\{t_{n_k}\}$ of $\{t_n\}$, $x^{t_{n_k}} \in X(t_{n_k})$, and $y^{t_{n_k}} \in Y(t_{n_k}, \hat{x}^0, x^{t_{n_k}})$ such that*

$$\lim_{k\to\infty} d_y G\left(t_{n_k}, \hat{x}^0, 0; (y^{t_{n_k}} - \hat{y}^0)/t_{n_k}\right) = R(0, \hat{x}^0, \hat{y}^0).$$

Then, $dg(0)$ exists and there exist $\hat{x}^0 \in X(0)$ and $\hat{y}^0 \in Y(0, \hat{x}^0)$ such that

$$dg(0) = d_t G(0, \hat{x}^0, \hat{y}^0) + R(0, \hat{x}^0, \hat{y}^0).$$

When $X(0) = \{x^0\}$ and $Y(0, x^0) = \{y^0\}$ are singletons, the above hypotheses are equivalent to the ones of Thm. 9.

Outline

THEOREM (GENERAL CASE)

Given X, Y, and G, let (H0) *and the following hypotheses be satisfied:*

(H1) $\forall t \in [0,\tau]$, $X(t) \neq \varnothing$, $g(t)$ *is finite, and* $\forall x^t \in X(t)$ *and* $x^0 \in X(0)$, $Y(t, x^0, x^t) \neq \varnothing$;

(H2) *for all* $x \in X(0)$ *and* $y \in Y(0,x)$, $d_t G(0,x,y)$ *exists and, for each* $x \in X(0)$*, there exists a function* $y \mapsto R(0,x,y) : Y(0,x) \to \mathbb{R}$ *satisfying (H3) and (H4) below;*

(H3) *for each sequence* $t_n \to 0$, $0 < t_n \leq \tau$, $\exists x^0 \in X(0)$ *such that for all* $y^0 \in Y(0,x^0)$, $\exists$*a subsequence* $\{t_{n_k}\}$ *of* $\{t_n\}$, $x^{t_{n_k}} \in X(t_{n_k})$*, and* $y^{t_{n_k}} \in Y(t_{n_k}, x^0, x^{t_{n_k}})$ *such that*

$$\liminf_{k\to\infty} d_y G\left(t_{n_k}, x^0, 0; (y^{t_{n_k}} - y^0)/t_{n_k}\right) \geq R(0, x^0, y^0);$$

(H4) *for each sequence* $t_n \to 0$, $0 < t_n \leq \tau$ *and all* $x^0 \in X(0)$*, there exist* $y^0 \in Y(0,x^0)$*, a subsequence* $\{t_{n_k}\}$ *of* $\{t_n\}$, $x^{t_{n_k}} \in X(t_{n_k})$*, and* $y^{t_{n_k}} \in Y(t_{n_k}, x^0, x^{t_{n_k}})$ *such that*

$$\limsup_{k\to\infty} d_y G\left(t, x^0, 0; (y^{t_{n_k}} - y^0)/t_{n_k}\right) \leq R(0, x^0, y^0).$$

Then, $dg(0)$ *exists and there exists* $\hat{x}^0 \in X(0)$ *and* $\hat{y}^0 \in Y(0,\hat{x}^0)$ *such that*

$$\begin{aligned} dg(0) &= d_t G(0, \hat{x}^0, \hat{y}^0) + R(0, \hat{x}^0, \hat{y}^0) \\ &= \sup_{y \in Y(0,\hat{x}^0)} d_t G(0, \hat{x}^0, y) + R(0, \hat{x}^0, y) = \inf_{x \in X(0)} \sup_{y \in Y(0,x)} d_t G(0,x,y) + R(0,x,y). \end{aligned}$$

Outline

Consider the objective function and the constraint set

$$f(x) \stackrel{\text{def}}{=} Qx \cdot x, \quad U \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n : Ax \cdot x = 1\|, \quad \boxed{\inf f(U)} \tag{4.1}$$

Where Q is an arbitrary symmetrical $n \times n$ matrix and $A > 0$ is a symmetrical $n \times n$ positive definite matrix. $U \neq \varnothing$ is compact and the function f is not necessarily convex.

The minimization problem is equivalent to the generalized eigenvalue problem

$$\boxed{\lambda(Q, A) \stackrel{\text{def}}{=} \inf_{x \neq 0} \frac{Qx \cdot x}{Ax \cdot x}} \tag{4.2}$$

where the minimizer $\hat{x}$ is solution of the problem

$$[Q - \lambda(Q, A)A]\,\hat{x} = 0, \quad A\hat{x} \cdot \hat{x} = 1. \tag{4.3}$$

The semidifferential of $\lambda(Q, A)$ with respect to Q in a direction Q' and A in the direction A' can be found in [Delfour 2011, pp. 166–168] for symmetrical matrices:

$$\begin{aligned} d\lambda(Q, A; Q', A') &= \inf_{x \in X(0)} Q'x \cdot x\,(Ax \cdot x) - (Qx \cdot x)\,A'x \cdot x \\ &= \inf_{x \in X(0)} Q'x \cdot x - \lambda(Q, A)\,A'x \cdot x, \end{aligned} \tag{4.4}$$

$$\text{minimizers } X(0) \stackrel{\text{def}}{=} \left\{x \in \mathbb{R}^n : [Q - \lambda(Q, A)A]x = 0 \text{ and } Ax \cdot x = 1\right\} \tag{4.5}$$

$$\text{states } E(0) \stackrel{\text{def}}{=} \left\{x \in \mathbb{R}^n : Ax \cdot x = 1\right\}. \tag{4.6}$$

For $t \geq 0$, $x \in \mathbb{R}^n$, and $y \in \mathbb{R}$, introduce the Lagrangian

$$G(t, x.y) \stackrel{\text{def}}{=} (Q + tQ')x \cdot x + y\,[(A + tA')x \cdot x - 1] \tag{4.7}$$

$$g(t) \stackrel{\text{def}}{=} \inf_{x \in \mathbb{R}^n} \sup_{y \in \mathbb{R}} G(t, x, y), \quad dg(0) \stackrel{\text{def}}{=} \frac{g(t) - g(0)}{t}. \tag{4.8}$$

where A' and Q' are symmetrical matrices. Set $Q(t) = Q + tQ'$ and $A(t) = A + tA'$. It is easy to check that

$$d_t G(t, x, y) = Q'x \cdot x + y\,A'x \cdot x \tag{4.9}$$

$$d_x G(t, x, y; x') = 2\,[Q(t) + y\,A(t)]\,x \cdot x' \tag{4.10}$$

$$d_y G(t, x, y; y') = y'\,[A(t)x \cdot x - 1]. \tag{4.11}$$

Since A is positive definite, there exists $\alpha > 0$ such that for all $x \in \mathbb{R}^n$, $Ax \cdot x \geq \alpha \|x\|^2$. Hence, there exists $\tau > 0$ such that for all $0 \leq t \leq \tau$

$$\forall t,\ 0 \leq t \leq \tau,\ \forall x \in \mathbb{R}^n, \quad A(t)x \cdot x \geq \frac{\alpha}{2} \|x\|^2$$

and for such t, the set of constraints $E(t) \stackrel{\text{def}}{=} \{x : A(t)x \cdot x = 1\} \neq \varnothing$ is compact. So there exist minimizers $x^t \in \mathbb{R}^n$ and $X(t)$ is not empty for $0 \leq t \leq \tau$

$$\lambda^t \stackrel{\text{def}}{=} \inf_{A(t)x \cdot x = 1} Q(t)x \cdot x = Q(t)x^t \cdot x^t \tag{4.12}$$

To summarize,

$$d_t G(t,x,y) = Q'x\cdot x + y\,A'x\cdot x \tag{4.13}$$

$$\left[Q(t) + y^t\,A(t)\right]\frac{x^t + x^0}{2} = 0 \text{ (average adjoint equation)} \tag{4.14}$$

$$\forall y',\quad d_y G(t,x^t,0;y') = y'\,[A(t)x^t\cdot x^t - 1] = 0 \text{ (state equation)} \tag{4.15}$$

$$d_y G\left(t,x^0,0;\frac{y^t - y^0}{t}\right) = \frac{y^t - y^0}{t}\,[A(t)x^0\cdot x^0 - 1]. \tag{4.16}$$

From the Lagrange Multiplier rule, the standard adjoint is solution of

$$\boxed{\left[Q(t) + p^t\,A(t)\right]x^t = 0} \quad \Rightarrow p^t = -Q(t)x^t\cdot x^t = -\lambda^t. \tag{4.17}$$

Tthe set of minimizers is given by the expression

$$X(t) = \left\{x\in\mathbb{R}^n : [Q(t) + p^t A(t)x = 0 \text{ and } A(t)x\cdot x = 1\right\}. \tag{4.18}$$

For all $x^t\in X(t)$, $x^t\neq 0$ and $-x^t\in X(t)$. So $X(t)$ is not a singleton. However,

$$\forall x^t\in X(t),\quad Y(t,x^t) = \{-\lambda^t\}$$

and $Y(t,x^t)$ is a singleton independent of the choice of the minimizer $x^t\in X(t)$.

Given $x^0 \in X(0)$ and $x^t \in X(t)$, the *averaged adjoint* is solution of the equation:

$$\begin{aligned} \forall x', \quad 0 &= \int_0^1 d_x G(t, x^0 + s(x^t - x^0), y^t; x') ds \\ &= 2 \int_0^1 \left[Q(t) + y^t A(t) \right] (x^0 + s(x^t - x^0)) \cdot x' \, ds \\ &= 2 \left[Q(t) + y^t A(t) \right] \frac{x^t + x^0}{2} \cdot x' \\ &\Rightarrow \boxed{\left[Q(t) + y^t A(t) \right] \frac{x^t + x^0}{2} = 0.} \end{aligned} \tag{4.19}$$

$$\Rightarrow Y(t, x^0, x^t) = \begin{cases} \left\{ -\dfrac{Q(t)\frac{x^t+x^0}{2} \cdot \frac{x^t+x^0}{2}}{A(t)\frac{x^t+x^0}{2} \cdot \frac{x^t+x^0}{2}} \right\}, & \text{if } x^t + x^0 \neq 0 \\ \mathbb{R}, & \text{if } x^t + x^0 = 0 \end{cases} \tag{4.20}$$

Therefore, $Y(t, x^0, x^t) \neq \varnothing$.

A preliminary lemma.

(i) For all t, $0 \leq t \leq \tau$,

$$\forall x^t \in X(t), \quad Y(t, x^t, x^t) = \{-\lambda^t\} \tag{4.21}$$

where λ^t is the minimum of the objective function $Q(t)x \cdot x$ with respect to $E(t) = \{x \in \mathbb{R}^n : A(t)x \cdot x = 1\}$ as seen in (4.12).

(ii) For each sequence $\{t_n : 0 < t_n \leq \tau\}$, there exist $\bar{x} \in X(0)$, $x^{t_n} \in X(t_n)$, and $y^{t_n} \in Y(t_n, \bar{x}, x^{t_n})$ such that

$$x^{t_n} \to \bar{x}, \quad \lambda^{t_n} \to \lambda^0, \quad \text{and} \quad y^{t_n} \to y^0 = -\lambda^0, \tag{4.22}$$

and the set of averaged adjoint states $Y(t_n, \bar{x}, x^{t_n}) = \{y^{t_n}\}$ is a singleton.

(iii) As $t \searrow 0$, the quotient

$$\frac{\lambda^t - \lambda^0}{t} \tag{4.23}$$

is bounded.

(iv) For the sequences of part (ii), the quotients

$$\frac{\lambda^{t_n} - \lambda^0}{t_n} \quad \text{and} \quad \frac{y^{t_n} - y^0}{t_n} \tag{4.24}$$

are bounded.

THEOREM

Given symmetrical $n \times n$ matrices A, A', Q, and Q' such that A is positive definite, there exists at least one x^0 such that $Ax^0 \cdot x^0 = 1$ and

$$\lambda(Q,A) = \inf_{Ax\cdot x=1} Qx \cdot x = Qx^0 \cdot x^0. \tag{4.25}$$

Moreover

$$\boxed{\begin{aligned} d\lambda(Q,A;Q',A') &\overset{\text{def}}{=} \lim_{t\searrow 0} \frac{\lambda(Q+tQ',A+tA')-\lambda(Q,A)}{t} \\ &= \inf_{x^0\in X(0)} \left[Q' - \lambda(Q,A)\,A'\right] x^0 \cdot x^0, \end{aligned}} \tag{4.26}$$

$$X(0) \overset{\text{def}}{=} \left\{x \in \mathbb{R}^n : Ax\cdot x = 1 \text{ and } [Q-\lambda(Q,A)A]x = 0\right\}. \tag{4.27}$$

If $X(0)$ is not simple the dimension of the space $X(0)$ is greater or equal to 2 and we only have a semi-différential.

Proof.

(i) *Hypothesis (H1).* We have seen that for all $0 \le t \le \tau$, $X(t) \neq \varnothing$ and that, for all $x^t \in X(t)$, $Y(t, x^t) = \{-\lambda^t\}$. For the averaged adjoint y^t

$$\Rightarrow \; Y(t, x^0, x^t) = \begin{cases} \left\{ -\dfrac{Q(t)\frac{x^t+x^0}{2} \cdot \frac{x^t+x^0}{2}}{A(t)\frac{x^t+x^0}{2} \cdot \frac{x^t+x^0}{2}} \right\}, & \text{if } x^t + x^0 \neq 0 \\ \mathbb{R}, & \text{if } x^t + x^0 = 0 \end{cases}$$

(ii) *Hypothesis (H2).* We have seen that $d_t G(t, x, y) = Q'x \cdot x + y\, A'x \cdot x$. So for all $x^0 \in X(0)$ and the singleton $Y(0, x^0) = \{-\lambda^0\}$

$$d_t G(t, x^0, y^0) = Q'x^0 \cdot x^0 - \lambda^0 A'x^0 \cdot x^0.$$

(iii) *Hypothesis (H3).* For each sequence $t_n \to 0$, $0 < t_n \le \tau$, choose the sequence $\{x^{t_n}\}$ and its limit $\bar{x} \in X(0)$ from the Lemma (ii) and use the fact that the corresponding sequence $\frac{y^{t_n} - y^0}{t_n}$ is bounded by some constant c from the Lemma (iv):

$$\left| d_y G\left(t_n, \bar{x}, 0; \frac{y^{t_n} - y^0}{t_n}\right) \right| = \left| \frac{y^{t_n} - y^0}{t} \; [A(t_n)\bar{x} \cdot \bar{x} - 1] \right|$$

$$\le \left| \frac{y^{t_n} - y^0}{t_n} \right| \, |A(t_n)\bar{x} \cdot \bar{x} - 1| \le c\, |A(t_n)\bar{x} \cdot \bar{x} - 1| \to c\, |A(0)\bar{x} \cdot \bar{x} - 1| = 0$$

(iv) *Hypothesis (H4).* For all $x^0 \in X(0)$ $Y(0,x^0) = \{-\lambda^0\}$ is a singleton independent of $x^0 \in X(0)$. As in (iii), for each sequence $t_n \to 0$, $0 < t_n \le \tau$, choose the sequence $\{x^{t_n}\}$ and its limit $\bar{x} \in X(0)$ from the Lemma (ii) and use the fact that the corresponding sequence $\frac{y^{t_n}-y^0}{t_n}$ is bounded by some constant c from the Lemma (iv):

$$\begin{aligned}\left| d_y G\left(t_n, x^0, 0; \frac{y^{t_n}-y^0}{t_n}\right)\right| &= \left|\frac{y^{t_n}-y^0}{t}\ \left[A(t_n)x^0\cdot x^0 - 1\right]\right| \\ &\le \left|\frac{y^{t_n}-y^0}{t_n}\right|\ \left|A(t_n)x^0\cdot x^0 - 1\right| \\ &\le c\left|A(t_n)x^0\cdot x^0 - 1\right| \to c\left|A(0)x^0\cdot x^0 - 1\right| = 0.\end{aligned}$$

(v) The conclusion follows from Theorem 12 where the sup disappears since $Y(0,x^0) = \{-\lambda^0\} = \{-\lambda(Q,A)\}$ is a singleton independent of $x^0 \in X(0)$.

- Thank you for your attention -

[Auscher-Tchamitchian (1998)] Auscher P. and Tchamitchian Ph. (1998), *Square root problem for divergence operators and related topics*, Astérisque 249.

[Bonnans-Shapiro (2000)] Bonnans J. F. and Shapiro A. (2000), *Perturbation analysis of optimization problems*, Springer-Verlag, N. Y..

[Correa-Seeger (1985)] Correa R. and Seeger A. (1985), *Directional derivatives of a minimax function*, Nonlinear Anal. Theory Methods and Appl. **9**, 13–22.

[Danskin (1966)] Danskin J. M. (1966), *The theory of max-min, with applications*, SIAM J. on Appl. Math. **14**, no. 4, 641–644.

[Delfour 2011] M. C. Delfour, *Introduction to optimization and semidifferential calculus*, MOS-SIAM Series, Philadelphia, USA, 2012.

[Delfour 2012] M. C. Delfour, *Groups of Transformations for Geometrical Identification Problems: Metrics, Geodesics*, pp. 3403–3406, Mini-Workshop: Geometries, Shapes and Topologies in PDE-based Applications, M. Hintermüller, G. Leugering, and J. Sokolowski, organizers, Mathematisches Forschungsinstitut Oberwolfach, 2012 (Report No. 57/2012, DOI: 10.4171/OWR/2012/57).

[Delfour (2015)] Delfour M. C. (2015), *Metrics spaces of shapes and geometries from set parametrized functions*, in “New Trends in Shape Optimization", A. Pratelli and G. Leugering, eds., pp. 57–101, International Series of Numerical Mathematics vol. 166, Birkhäuser Basel.

[Delfour (2017)] Delfour M. C. (2017), *Differentials and Semidifferentials for Metric Spaces of Shapes and Geometries*, in “System Modeling and Optimization, (CSMO 2015),” L. Bociu, J. A. Desideri and A. Habbal, eds., pp. 230–239, AICT Series, Springer, 2017.

[Delfour (2017-2)] M. C. Delfour, *Topological Derivative: a Semidifferential via the Minkowski Content*, 2016, submitted.

[Delfour-Sturm (2015)] Delfour M. C. and Sturm K. (2015), *Parametric semidifferentiability of minimax of Lagrangians: averaged adjoint state approach*, Report CRM-3351, Centre de recherches mathématiques, Univ. de Montréal, Canada.

[Delfour-Sturm (2017)] M. C. Delfour and K. Sturm, *Parametric Semidifferentiability of Minimax of Lagrangians: Averaged Adjoint Approach*, Journal of Convex Analysis **24** (2017), No. 4, to appear.

[Delfour-Sturm (2016)] M. C. Delfour and K. Sturm, *Minimax differentiability via the averaged adjoint for control/shape sensitivity*, Proc. of the 2nd IFAC Workshop on Control of Systems Governed by Partial Differential quations, IFAC-PaperOnLine **49-8** (2016), pp. 142–149.

[Delfour-Zolésio (1988)] Delfour M. C. and Zolésio J. P. (1988), *Shape sensitivity analysis via min max differentiability*, SIAM J. on Control and Optimization **26**, 834–862.

[Delfour-Zolésio (2011)] Delfour M. C. and Zolésio J. P. (2011), *Shapes and geometries: metrics, analysis, differential calculus, and optimization*, 2nd edition, SIAM series on Advances in Design and Control, Society for Industrial and Applied Mathematics, Philadelphia, USA.

[Dem'janov (1969)] Dem'janov V. F. (1969), *Differentiation of the maximin function. I, II*, (Russian) Ž. Vyčisl. Mat. i Mat. Fiz. **8** (1968), 1186–1195; ibid. **9**, 55–67.

[Dem'janov (1974)] Dem'janov V. F. (1974), *Minimax: Directional Differentiation*, Izdat. Leningrad Gos. Univ., Leningrad.

[Dem'janov-Malozemov (1974)] Dem'janov V. F. and Malozemov V. N. (1974), *Introduction to the Minimax*, Trans from Russian by D. Louvish, Halsted Press [John Wiley & Sons], New York-Toronto.

[Eschenauer et al (1994)] Eschenauer H A., Kobelev V. V., and Schumacher A. (1994), *Bubble method for topology and shape optimization of structures*, Structural Optimization **8**, 42–51.

[Lemaire (1970)] Lemaire B. , *Problèmes min-max et applications au contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles linéaires*, Thèse de doctorat d'état, Univ. de Montpellier, France 1970.

[Novotny-Sokolowski (2013)] Novotny A. A. and Sokolowski J. (2013), *Topological Derivatives in Shape Optimization*, Interaction of Mechanics and Mathematics, Springer, Heidelberg, N. Y..

[Sokołowski-Zòchowski (1999)] Sokolowski J. and Zòchowski A. (1999), *On the topological derivative in shape optimization*, SIAM J. Control Optim. **37**, no. 4, 1251–1272.

[Sturm (2014)] Sturm K. (2014), *On shape optimization with non-linear partial differential equations*, Doctoral thesis, Technische Universiltät of Berlin, Germany.

[Sturm (2015)] Sturm K. (2015), *Minimax Lagrangian approach to the differentiability of non-linear PDE constrained shape functions without saddle point assumption*, SIAM J. on Control and Optim., **53**, no. 4, 2017–2039.

[Zolésio (1979)] J. P. Zolésio, *Identification de domaines par déformation*, Thèse de doctorat d'état, Université de Nice, France, 1979.