Weierstrass Institute for
Applied Analysis and Stochastics

Leibniz Gemeinschaft

Mini-Course "Introduction to Optimal Control Problems for PDEs"

Jürgen Sprekels, Weierstrass Institute for Applied Analysis and Stochastics (WIAS), Berlin, Germany

Lecture 1: Basic Concepts
Lecture 2: Elliptic Control Problems
Lecture 3: Linear-quadratic Parabolic Control Problems
Lecture 4: Necessary Optimality Conditions for General Optimization Problems in Banach Spaces

1. What is an optimal control problem (OCP)?

A typical OCP consists of 4 ingredients:

- a **control function** u;
- a **state equation**, which associates with every control u a **state function** y;
- a **cost functional**, which depends on u and y, to be minimized;
- various **constraints** to be obeyed by the control u and the state y.

Constraints on u are called **control constraints**, constraints on y **state constraints**. Very often, the control and state constraints are **pointwise constraints**. In this summer school, the state equation will be given by an initial-boundary value problem for a PDE or by a variational inequality.

Steps to be taken in the analysis on an OCP:

- Existence of a solution. Is it unique?
- Derivation of first-order necessary and second-order sufficient optimality conditions
- Numerical approximation and analysis of the discretized problem. Do we have convergence?
- Implementation of the discretized problem and calculation of an approximating solution

Example 1: Optimal stationary heating

Let $\Omega \subset \mathbb{R}^3$ be the spatial location of a body to be heated/cooled at its boundary Γ. We apply a heat source u (the control) to Γ that is constant in time: $u = u(x)$. We aim to choose u in such a way that the corresponding temperature $y = y(x)$ (the state) is the "best approximation" to some desired $y_\Omega = y_\Omega(x)$ in Ω. Model:

$$\min \; J(y,u) := \frac{1}{2}\int_\Omega |y(x) - y_\Omega(x)|^2 \, dx + \frac{\lambda}{2}\int_\Gamma |u(x)|^2 \, ds(x),$$

subject to the **state equation**

$$\boxed{\begin{array}{rcll} -\Delta y & = & 0 & \text{in } \Omega \\ \dfrac{\partial y}{\partial \nu} & = & \alpha\,(u - y) & \text{on } \Gamma \end{array}}$$

and the **pointwise control constraints**

$$u_a(x) \le u(x) \le u_b(x) \quad \text{on } \Gamma.$$

Here, $\lambda > 0$ is a parameter, and u_a, u_b are given.
This is a **linear-quadratic elliptic boundary control problem**.

Example 2: Optimal heat source

Assume here that Ω is heated by a distributed heat source u acting in Ω (e.g., by electromagnetic induction). We then get:

$$\min \; J(y,u) := \frac{1}{2}\int_\Omega \left|y(x) - y_\Omega(x)\right|^2 dx + \frac{\lambda}{2}\int_\Omega |u(x)|^2\, dx,$$

subject to

$$\boxed{\begin{array}{rcll} -\Delta y & = & \beta u & \text{in } \Omega \\ \dfrac{\partial y}{\partial \nu} & = & \alpha\,(y_a - y) & \text{on } \Gamma \end{array}}$$

and

$$u_a(x) \le u(x) \le u_b(x) \quad \text{in } \Omega.$$

Here, $y_a =$ ext. temperature, $\beta = \beta(x)$ given, e.g., $\beta = \chi_D$ for $D \subset \Omega$.

This is a **linear-quadratic elliptic control problem with distributed control**.

Example 3: Optimal nonstationary boundary control

This time we assume that $\Omega \subset \mathbb{R}^3$ is heated over some time $[0,T]$, $T>0$. We want to apply a control $u=u(x,t)$ in such a way that a desired temperature distribution y_Ω is reached at $t=T$. We put $Q := \Omega \times (0,T)$, $\Sigma := \Gamma \times (0,T)$. Problem:

$$\min J(y,u) := \frac{1}{2}\int_\Omega \left|y(x,T) - y_\Omega(x)\right|^2 dx + \frac{\lambda}{2}\int_0^T\int_\Gamma |u(x,t)|^2\, ds(x)\, dt,$$

subject to

$$\boxed{\begin{array}{rcll} y_t - \Delta y &=& 0 & \text{in } Q \\ \dfrac{\partial y}{\partial \nu} &=& \alpha\,(u-y) & \text{on } \Sigma \\ y(x,0) &=& y_0(x) & \text{in } \Omega \end{array}}$$

and

$$u_a(x,t) \le u(x,t) \le u_b(x,t) \quad \text{on } \Sigma.$$

This is a **linear-quadratic parabolic boundary control problem**.

Example 4: Control of nonstationary flows

Nonstationary flows of incompressible media in $\Omega \subset \mathbb{R}^2$ (or $\mathbb{R}^3$) are described by the **Navier–Stokes equations**

$$
\begin{aligned}
u_t - \frac{1}{Re}\Delta u + (u\cdot\nabla)u + \nabla p &= f && \text{in } Q \\
\operatorname{div} u &= 0 && \text{in } Q \\
u &= 0 && \text{on } \Sigma \\
u(\cdot,0) &= u_0 && \text{in } \Omega.
\end{aligned}
$$

Here, typically, the velocity field u (which is in this case the state variable) is to be controlled by the application of electromagnetic fields. In this case, $f =$ gravity $g +$ Lorentz force $j \times B$.

One then obtains a control problem with a nonlinear state equation.

Example 5: Control of phase-field equations (C. Lefter + J. S., AMSA 17 (2007))

Consider a liquid-solid phase transition, and let $\chi \in [0,1]$ denote the **order parameter**: $\{\chi = 0\} \Leftrightarrow$ solid phase, $\{\chi = 1\} \Leftrightarrow$ liquid phase, $\{0 < \chi < 1\} \Leftrightarrow$ "mush". A typical model is ($\theta =$ absolute temperature)

$$\mu(\theta)\chi_t = -F_1'(\chi) - \left(\frac{\beta_1}{\theta} + \beta_2\right) F_2'(\chi) - \frac{F_3'(\chi)}{\theta} \quad \text{in } Q,$$

$$C_V\,\theta_t + (\beta_1 F_2'(\chi) + F_3'(\chi))\chi_t - \Delta\theta = 0 \quad \text{in } Q,$$

$$\frac{\partial\theta}{\partial\nu} = u - \theta \quad \text{on } \Sigma,$$

$$\chi(\cdot,0) = \chi_0,\ \theta(\cdot,0) = \theta_0, \quad \text{in } \Omega.$$

F_1, F_3 : smooth
$F_2(\chi) = \kappa(\chi\ln(\chi) + (1-\chi)\ln(1-\chi)),$
$\kappa > 0.$

Control problem:

$$\min J(\chi,\theta,u) := \frac{1}{2}\int_\Omega \left(|\theta(x,T) - \theta_\Omega(x)|^2 + |\chi(x,T) - \chi_\Omega(x)|^2\right) dx + \frac{\lambda}{2}\int_\Sigma |u(x,t)|^2\, ds(x)\, dt$$

subject to the IBVP and the control constraints
$0 < u_a \le u(x,t) \le u_b$ on Σ.

Example 6: A control-into-coefficients problem

Consider a clamped plate $\Omega \subset \mathbb{R}^2$, which is described by the BVP

$$\boxed{\begin{aligned} \Delta(b\,u^3(x)\,\Delta y) &= f \quad \text{in } \Omega \\ y = \frac{\partial y}{\partial \nu} &= 0 \qquad \text{on } \Gamma. \end{aligned}}$$

Here, $b > 0$ and the load f are given; $u \in L^\infty(\Omega)$ stands for the thickness (> 0!) of the plate. A typical problem is:

$$\min J(y,u) := \int_\Omega u(x)\,dx \quad \text{(= the weight of the plate)}$$

subject to the BVP and the control and state constraints (= safety requirements)

$$\begin{aligned} 0 < m \le u(x) \le M \quad & \text{in } \Omega \\ y(x) \ge -\tau \qquad & \text{in } \Omega, \end{aligned}$$

where $m, M, \tau > 0$ are given.

Example 7: Optimal layout of materials

Let $\Omega \subset \mathbb{R}^3$ be a body composed of m different materials M_i having thermal conductivities k_i, $i = 1, \ldots, m$. Then the total thermal conductivity k of Ω is $k(x) = \sum_{i=1}^{m} \chi_i(x)\, k_i$, where χ_i is the characteristic function of the region occupied by M_i.

Problem: Given a heat source f, what is the optimal distribution of the materials in Ω that maximizes the temperature y in a given subdomain $\omega \subset \Omega$? Model:

$$\min_k \left\{ -\int_\omega y(x)\, dx \right\}$$

subject to

$$\begin{aligned} -\nabla \cdot (k(x)\, \nabla y) &= f && \text{in } \Omega \\ \frac{\partial y}{\partial \nu} &= 0 && \text{on } \Gamma. \end{aligned}$$

The optimization parameters are the subsets of Ω occupied by the various materials.

Example 8: Electrochemical machining

This is an OCP with a variational inequality as state equation:

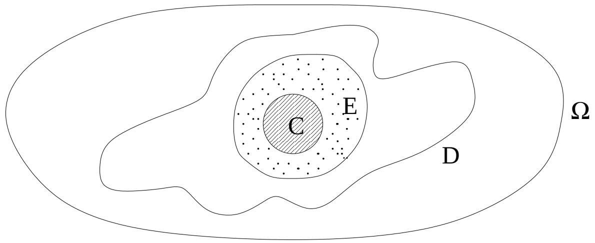
$$\min_{E \subset D \subset \Omega} J(D) = \frac{1}{2} \int_{E \setminus C} |y(x)|^2 \, dx \,,$$

subject to

$$\int_{D \setminus C} \nabla y(x) \cdot \nabla (y - z)(x) \, dx \le \int_{D \setminus C} f(x) \, (y(x) - z(x)) \, dx,$$

$$\forall z \in S = \{ w \in H^1(D \setminus C) : w|_{\partial C} = 0,$$

$$w|_{\partial D} = 1, \; w \ge 0 \text{ a.e. in } D \setminus C \} \,,$$



C: core

D: machine

$\partial C, \partial D$: electrodes

$E \setminus C$: desired shape (choice of E)

$E_y = \{ x \in D \setminus C : y(x) = 0 \}$: final shape obtained

To guarantee: $E_y \supset E \setminus C$!

1. Let: $J = J(y,u) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ cost functional to be minimized
$A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $U_{ad} \subset \mathbb{R}^m$ nonempty. We consider:

$$\boxed{\begin{array}{c} \min J(y,u) \\ Ay = Bu,\ u \in U_{ad} \end{array}}$$

Convention: All vectors are column vectors!

Example: $J(y,u) = \frac{1}{2}|y - y_d|^2 + \frac{\lambda}{2}|u|^2$, $y_d \in \mathbb{R}^n$ given; $|\cdot| =$ Euclidean norm

Assume: $\exists A^{-1} \implies y = Su$, with the solution matrix $S := A^{-1}B$.

$\implies$ **reduced cost functional** $f(u) := J(y,u) = J(Su,u)$.

In the example: $f(u) = \frac{1}{2}|Su - y_d|^2 + \frac{\lambda}{2}|u|^2$.

$\implies$ **reduced problem** $\boxed{\min f(u)\,, \quad u \in U_{ad}}$

2. Existence:

Def.: If $\bar{u} \in U_{ad}$ satisfies $f(\bar{u}) \leq f(u) \quad \forall\, u \in U_{ad}$, then $\bar{u}$ is called **optimal control**, and $\bar{y} := S\bar{u}$ is called associated **optimal state**.

Theorem: Let $J : \mathbb{R}^n \times U_{ad} \to \mathbb{R}$ be continuous, $U_{ad} \neq \emptyset$, closed, bounded. If $\exists A^{-1}$, then $\exists$ an optimal pair $(\bar{u}, \bar{y})$.

Proof: f is continuous on the compact set U_{ad}.

3. First-order necessary optimality conditions:

We use the notation: for differentiable $f : \mathbb{R}^m \to \mathbb{R}$ we write
$f'(u)\,h = \nabla f(u) \cdot h$, $h \in \mathbb{R}^m$. Let J be continuously differentiable w.r.t. y, u.
$\Longrightarrow$ $f(u) = J(Su, u)$ is C^1

Theorem: Let U_{ad} be convex. Then any optimal control $\bar{u}$ satisfies the variational inequality: $f'(u)(u - \bar{u}) \geq 0 \quad \forall\, u \in U_{ad}$.

Chain rule: $f' = D_y J\, S + D_u J$, and thus

$$\begin{aligned} f'(\bar u)\,h &= D_y J(S\bar u,\bar u)\,Sh + D_u J(S\bar u,\bar u)\,h \\ &= \big(\nabla_y J(\bar y,\bar u)\,,\, A^{-1}Bh\big)_{\mathbb{R}^n} + \big(\nabla_u J(\bar y,\bar u)\,,\, h\big)_{\mathbb{R}^m} \\ &= \big(B^\top (A^\top)^{-1}\nabla_y J(\bar y,\bar u) + \nabla_u J(\bar y,\bar u)\,,\, h\big)_{\mathbb{R}^m}. \end{aligned}$$

$\Longrightarrow$ Variational inequality becomes

$$\big(B^\top (A^\top)^{-1}\nabla_y J(\bar y,\bar u) + \nabla_u J(\bar y,\bar u)\,,\, u-\bar u\big)_{\mathbb{R}^m} \ge 0 \qquad \forall u \in U_{ad}.$$

This requires to evaluate the inverse of $A^\top$, which is usually a bad idea. We therefore introduce the **adjoint state** p, given by $p = (A^\top)^{-1}\nabla_y J(\bar y,\bar u)$, which solves the **adjoint state equation**

$$\boxed{A^\top p = \nabla_y J(\bar y,\bar u)\,.}$$

Example: $J(y,u) = \frac{1}{2}|y - y_d|^2 + \frac{\lambda}{2}|u|^2 \implies A^\top p = y - y_d\,.$

Advantages: First-order condition simplifies, use of A^{-1} is avoided. Also the form of $\nabla f(\bar{u})$ simplifies: $\nabla f(\bar{u}) = B^\top p + \nabla_u J(\bar{y},\bar{u})\,.$

$$\implies f'(\bar{u})\,h = (B^\top p + \nabla_u J(\bar{y},\bar{u}), h)_{\mathbb{R}^m}\,.$$

Theorem 1: Suppose $\exists A^{-1}$. Let $(\bar{y},\bar{u})$ be an optimal pair. Then there exists an associated adjoint state $\bar{p}$ such that $(y,u,p) = (\bar{y},\bar{u},\bar{p})$ solves

$$\boxed{\begin{aligned} &A\,y = B\,u, \quad u \in U_{ad} \\ &A^\top p = \nabla_y J(y,u) \\ &\big(B^\top p + \nabla_u J(y,u)\,,\, v - u\big)_{\mathbb{R}^m} \geq 0 \quad \forall v \in U_{ad}. \end{aligned}}$$

In the case $U_{ad} = \mathbb{R}^m$ follows: $B^\top \bar{p} + \nabla_u J(\bar{y},\bar{u}) = 0.$

Example: $J(y,u) = \frac{1}{2}|Cy - y_d|^2 + \frac{\lambda}{2}|u|^2, \quad C \in \mathbb{R}^{n\times n}$.

$\Longrightarrow \nabla_y J(y,u) = C^\top(Cy - y_d), \qquad \nabla_u J(y,u) = \lambda u$

$\Longrightarrow$ Optimality system:

$$
\begin{aligned}
&Ay = Bu, \quad u \in U_{ad}\\
&A^\top p = C^\top(Cy - y_d)\\
&\big(B^\top p + \lambda u\,,\, v - u\big)_{\mathbb{R}^m} \geq 0 \quad \forall v \in U_{ad}.
\end{aligned}
$$

If $U_{ad} = \mathbb{R}^m$, then $B^\top \bar{p} + \lambda \bar{u} = 0$, and if $\lambda > 0$, then $\bar{u} = -\frac{1}{\lambda}B^\top \bar{p}$. We obtain the optimality system

$$
\boxed{
\begin{aligned}
&Ay = -\frac{1}{\lambda}BB^\top p\\
&A^\top p = C^\top(Cy - y_d)
\end{aligned}
}
$$

4. Lagrangians

Def.: The function $L:\mathbb{R}^{2n+m}\to\mathbb{R}$, $L(y,u,p):=J(y,u)-(Ay-Bu,p)_{\mathbb{R}^n}$, is called the **Lagrangian function** of the OCP.

Using L, we can eliminate the equality constraint $Ay=Bu$ from the problem. The second and third optimality conditions read:

$$\nabla_y L(\bar{y},\bar{u},\bar{p})=0$$
$$\big(\nabla_u L(\bar{y},\bar{u},\bar{p})\,,\,u-\bar{u}\big)_{\mathbb{R}^m}\geq 0\quad\forall u\in U_{ad}.$$

$\Longrightarrow$ $(\bar{y},\bar{u})$ solves the necessary optimality conditions of a minimization problem **without** equality constraints:

$$\min_{(y,u)} L(y,u,p),\qquad u\in U_{ad},\ y\in\mathbb{R}^n.$$

Remark: p plays the role of a **Lagrange multiplier** to $Ay=Bu$.

5. Discussion of the variational inequality

We now assume **box constraints** for u:

$$U_{ad} = \{u \in \mathbb{R}^m : u_a \le u \le u_b\}, \quad u_a, u_b \in \mathbb{R}^m \quad \text{(componentwise)}$$

Variational inequality yields

$$\big(B^\top \bar{p} + \nabla_u J(\bar{y}, \bar{u}) \, , \, \bar{u}\big)_{\mathbb{R}^m} \le \big(B^\top \bar{p} + \nabla_u J(\bar{y}, \bar{u}) \, , \, u\big)_{\mathbb{R}^m} \quad \forall u \in U_{ad}$$

$\Longrightarrow$ $\bar{u}$ solves

$$\min_{u \in U_{ad}} \big(B^\top \bar{p} + \nabla_u J(\bar{y}, \bar{u}) \, , \, u\big)_{\mathbb{R}^m} = \min_{u \in U_{ad}} \sum_{i=1}^m \big(B^\top \bar{p} + \nabla_u J(\bar{y}, \bar{u})\big)_i u_i$$

Since u_i are independent of each other $\Longrightarrow$

$$\big(B^\top \bar{p} + \nabla_u J(\bar{y}, \bar{u})\big)_i \bar{u}_i = \min_{u_{a,i} \le u_i \le u_{b,i}} \big(B^\top \bar{p} + \nabla_u J(\bar{y}, \bar{u})\big)_i u_i, \quad 1 \le i \le m.$$

Hence, we must have:

$$\bar{u}_i = \begin{cases} u_{b,i} & \text{if} \quad \left(B^\top \bar{p} + \nabla_u J(\bar{y},\bar{u})\right)_i < 0 \\ u_{a,i} & \text{if} \quad \left(B^\top \bar{p} + \nabla_u J(\bar{y},\bar{u})\right)_i > 0. \end{cases}$$

We have no direct information if $(B^\top \bar{p} + \nabla_u J(\bar{y},\bar{u}))_i = 0$.

Notice that we only need information about $(B^\top \bar{p} + \nabla_u J(\bar{y},\bar{u}))$, not about $\bar{p}, \bar{u}$ directly!

Moreover, putting (with $z_+ := \frac{1}{2}(|z| + z)$, $z_- := \frac{1}{2}(|z| - z)$)

$$\begin{aligned} \mu_a &:= \left(B^\top \bar{p} + \nabla_u J(\bar{y},\bar{u})\right)_+ \\ \mu_b &:= \left(B^\top \bar{p} + \nabla_u J(\bar{y},\bar{u})\right)_-, \end{aligned}$$

We find that

$$
\begin{array}{|ccccccc|}
\hline
\mu_a & \geq & 0, \quad u_a - \bar{u} & \leq & 0, \quad \big(u_a - \bar{u}\,,\, \mu_a\big)_{\mathbb{R}^m} & = & 0 \\
\mu_b & \geq & 0, \quad \bar{u} - u_b & \leq & 0, \quad \big(\bar{u} - u_b\,,\, \mu_b\big)_{\mathbb{R}^m} & = & 0 \\
\hline
\end{array}
$$

These are the so-called **complementary slackness conditions**.

Remark: We obviously have $\mu_a - \mu_b = \nabla_u J(\bar{y}, \bar{u}) + B^\top \bar{p}$

(1) $$\Longrightarrow \nabla_u J(\bar{y}, \bar{u}) + B^\top \bar{p} - \mu_a + \mu_b = 0$$

6. The Karush–Kuhn–Tucker conditions

Introduce the **extended Lagrangian** $\mathscr{L}$ by adding the inequality constraints in the form

$$\mathscr{L}(y,u,p,\mu_a,\mu_b) := J(y,u) - \big(Ay - Bu\,,\,p\big)_{\mathbb{R}^n} + \big(u_a - u\,,\,\mu_a\big)_{\mathbb{R}^m} + \big(u - u_b\,,\,\mu_b\big)_{\mathbb{R}^m}.$$

Then, using (1),

$$\nabla_u \mathscr{L}(\bar{y},\bar{u},\bar{p},\mu_a,\mu_b) = 0.$$

Moreover, since $\nabla_y \mathscr{L} = \nabla_y L$, the adjoint equation can be written as

$$\nabla_y \mathscr{L}(\bar{y},\bar{u},\bar{p},\mu_a,\mu_b) = 0.$$

Hence, μ_a,μ_b are the Lagrange multipliers corresponding to the inequality constraints, and we have:

Theorem 2 (Karush–Kuhn–Tucker conditions)

Suppose $\exists A^{-1}$. Let $U_{ad} = \{u \in \mathbb{R}^m : u_a \leq u \leq u_b\}$. If $(\bar{y}, \bar{u})$ is an optimal pair, then $\exists$ multipliers $p \in \mathbb{R}^n$ and $\mu_a, \mu_b \in \mathbb{R}^m$ such that:

$$\boxed{\begin{aligned} &\nabla_y \mathscr{L}(\bar{y}, \bar{u}, \bar{p}, \mu_a, \mu_b) = 0 \\ &\nabla_u \mathscr{L}(\bar{y}, \bar{u}, \bar{p}, \mu_a, \mu_b) = 0 \\ &\mu_a \geq 0, \quad \mu_b \geq 0 \\ &\big(u_a - \bar{u}\,,\, \mu_a\big) = \big(\bar{u} - u_b\,,\, \mu_b\big) = 0 \end{aligned}}$$

We assume:

(A1) V, U are reflexive Banach spaces with duals V^*, U^*; H is a Hilbert space with $V \subset H$ with continuous embedding

(A2) U_{ad} is nonempty, closed and convex in U;
C is nonempty, closed and convex in H

(A3) $A \in \mathscr{L}(V, V^*)$, $\quad B \in \mathscr{L}(U, V^*)$, $\quad f \in V^*$

(A4) $J : V \times U \to (-\infty, +\infty]$ is proper, convex, l.s.c.

We consider the control and state constrained OCP:

$$\min J(y,u) \text{ subject to } Ay = Bu + f, \quad u \in U_{ad}, \quad y \in C.$$

We assume **admissibility**:

(A5) $\exists (y_0, u_0) \in C \times U_{ad}$ such that $Ay_0 = Bu_0 + f$ and $J(y_0, u_0) < +\infty$.

Theorem 3: Suppose $\exists A^{-1} \in \mathscr{L}(V^*, V)$. If U_{ad} is bounded or J is uniformly coercive in u, i.e.,

(2) $$\lim_{\|u\|_U \to \infty} J(y,u) = +\infty \quad \text{uniformly in } y,$$

then OCP has at least one optimal pair $(\bar{y}, \bar{u}) \in C \times U_{ad}$.

Proof: If U_{ad} is bounded, redefine J outside $V \times U_{ad}$ by $+\infty$. We thus may assume that (2) holds. By **(A5)**, $\exists$ minimizing sequence $\{u_n\}$ such that, with $y_n = A^{-1} B u_n \in C$, we have

$$\lim_{n \to \infty} J(y_n, u_n) = \inf(\text{OCP}) =: \delta .$$

By (2) $\{u_n\}$ is bounded in U, and therefore $\{y_n\}$ is bounded in V. We thus may assume that $(y_n, u_n) \to (\bar{y}, \bar{u})$ weakly in $V \times U$, and hence in $H \times U$. Consequently, by **(A2)**, we have $(\bar{y}, \bar{u}) \in C \times U_{ad}$. Moreover, $\bar{y} = A^{-1} B \bar{u}$. Since J is weakly l.s.c. on $V \times U$, we find that

$$J(\bar{y}, \bar{u}) \le \liminf_{n \to \infty} J(y_n, u_n) = \delta . \qquad \square$$

We now consider the more general state equation

(3) $$A(u)\,y + \partial\varphi(y) \ni B\,u + f\,.$$

Here, we assume $B \in \mathscr{L}(U,V^*)$, $f \in V^*$ (as above) and:

(A6) $\varphi : V \to (-\infty,+\infty]$ is proper, convex, l.s.c., with subdifferential

$$\partial\varphi(x) = \{w \in V^*;\, \varphi(x) - \varphi(v) \le (w, x - v)_{V^* \times V} \quad \forall\, v \in V\}$$

(A7) **(i)** $A(u) \in \mathscr{L}(V,V^*) \quad \forall\, u \in U_{ad}$

(ii) $u_n \to u$ strongly in $U \implies A(u_n) \to A(u)$ strongly in $\mathscr{L}(V,V^*)$

(iii) $(A(u)y,y)_{V^* \times V} \ge m\|y\|_V^2 \quad \forall\, y \in V$, for some $m > 0$.

J, U_{ad}, C have the same properties as above.

Theorem 4: Let the above conditions be satisfied, and suppose $\exists (y_0, u_0) \in C \times U_{ad}$ such that $A(u_0)\, y_0 + \partial\varphi(y_0) \ni B\, u_0 + f$ and $L(y_0, u_0) < +\infty$ (admissibility).

Then the OCP

$$\min J(y,u)\,, \text{ subject to } (y,u) \in C \times U_{ad} \text{ and (3)}$$

has a solution provided that one of the following conditions is satisfied:

(a) $U_{ad} \subset U$ is compact (but not necessarily convex!)

(b) $C \subset V$ is compact (but not necessarily convex!), $U_{ad} \subset U$ is bounded, and the mapping $u \mapsto A(u)$ is linear and bounded.

Proof (only of (a)): Pick a minimizing sequence $\{u_n\} \subset U_{ad}$, $J(y_n, u_n) \to \inf(\text{OCP})$, where $\{y_n\} \subset C$ and $A(u_n)\, y_n + \partial\varphi(y_n) \ni B\, u_n + f$.

Since U_{ad} is compact, we may assume that $\|u_n - u\|_U \to 0$ for some $u \in U_{ad}$.

Then $B\, u_n \to B\, u$ strongly in V^* , and $A(u_n) \to A(u)$ strongly in $\mathscr{L}(V, V^*)$.

Now, $\forall\, n \in \mathbb{N}\ \exists\, v_n \in \partial\varphi(y_n) : A(u_n)\, y_n + v_n = B\, u_n + f$.

The monotonicity of the graph $\partial\varphi$ (φ is convex!) yields:

$$
\begin{aligned}
&(Bu_n + f, y_n - y_1)_{V^* \times V} \\
&= (A(u_n)y_n, y_n - y_1)_{V^* \times V} + (v_n - v_1, y_n - y_1)_{V^* \times V} + (v_1, y_n - y_1)_{V^* \times V} \\
&\geq (A(u_n)(y_n - y_1), y_n - y_1)_{V^* \times V} + (A(u_n)y_1, y_n - y_1)_{V^* \times V} - \|v_1\|_{V^*} \|y_n - y_1\|_V .
\end{aligned}
$$

By **(A7)**, we thus can find some $C_1 > 0$ such that

$$
m\, \|y_n - y_1\|_V^2 \leq C_1\, \|y_n - y_1\|_V \quad \forall n \in \mathbb{N}.
$$

$\Longrightarrow$ $\{y_n\}$ is bounded in V , and we may assume there is some subsequence, again indexed by n, such that

$$y_n \longrightarrow y \text{ weakly in } V, \text{ and clearly } y \in C.$$

Denoting by $A(u)^* : V \simeq V^{**} \longrightarrow V^*$ the dual operator of $A(u)$, we find $\forall\, v \in V$:

$$
\begin{aligned}
&|(A(u_n)y_n, v)_{V^*\times V} - (A(u)y, v)_{V^*\times V}| \\
&\quad \le |(A(u_n)y_n, v)_{V^*\times V} - (A(u)y_n, v)_{V^*\times V}| \\
&\qquad + |(A(u)y_n, v)_{V^*\times V} - (A(u)y, v)_{V^*\times V}| \\
&\quad \le |v|_V\, |A(u_n)y_n - A(u)y_n|_{V^*} + |(y_n - y, A(u)^* v)_{V\times V^*}| \\
&\quad \le |v|_V\, |A(u_n) - A(u)|_{L(V,V^*)}\, |y_n|_V + |(y_n - y, A(u)^* v)_{V\times V^*}| \longrightarrow 0
\end{aligned}
$$

as $n \longrightarrow \infty$.

Moreover, it follows from **(A7)** that

(4)
$$\liminf_{n\to\infty} (A(u_n)y_n, y_n - y)_{V^*\times V}$$
$$= \liminf_{n\to\infty} \left[(A(u_n)(y_n - y), y_n - y)_{V^*\times V} + (A(u_n)\, y, y_n - y)_{V^*\times V}\right]$$
$$\geq \liminf_{n\to\infty} (A(u_n)y, y_n - y)_{V^*\times V}$$
$$= \lim_{n\to\infty} (A(u_n)y, y_n)_{V^*\times V} - \lim_{n\to\infty} (A(u_n)y, y)_{V^*\times V} = 0.$$

Now, the definition of $\partial\varphi$ implies that

$$\forall\, v \in V : \varphi(v) - \varphi(y_n) \geq (B u_n + f - A(u_n)\, y_n, v - u_n)_{V^*\times V}$$

$$\overset{v=y}{\Longrightarrow} (A(u_n)y_n, y_n - y)_{V^*\times V} + \varphi(y_n) \leq (Bu_n + f, y_n - y)_{V^*\times V} + \varphi(y)$$

But then, owing to the lower semicontinuity of φ,

(5)
$$\limsup_{n\to\infty} (A(u_n)\, y_n, y - y_n)_{V^*\times V} \leq \varphi(y) - \liminf_{n\to\infty} \varphi(y_n) \leq 0$$

From (4), (5), we conclude that

$$\lim_{n\to\infty} (A(u_n)y_n, y_n - y) = 0.$$

On the other hand,

$$(A(u_n)y_n - A(u_n)y, y_n - y)_{V^*\times V} \geq m\|y_n - y\|_V^2,$$

that is, $y_n \to y$ strongly in V, for a subsequence indexed by n.

Moreover, since $A(u_n) \to A(u)$ strongly in $\mathscr{L}(V, V^*)$ and $y_n \to y$ strongly in V, we have

$$A(u_n)\,y_n \to A(u)\,y \quad \text{weakly in } V^*.$$

Finally, observe that the **maximal monotone** operator $\partial\varphi$ is **demiclosed**, i.e., we have: whenever $v_n \in \partial\varphi(z_n)$, $\|z_n - z\|_V \to 0$ and $v_n \to v$ weakly in V^*, then $v \in \partial\varphi(z)$.

Now,

$$v_n = Bu_n + f - A(u_n)y_n \to v = Bu + f - A(u)y \quad \text{weakly in } V^*.$$

Then,

$$v = Bu + f - A(u)y \in \partial\varphi(y),$$

and the pair $(y,u) \in C \times U_{ad}$ is admissible. The minimality follows again from the lower semicontinuity of J. □