AIDE-MÉMOIRE ON THE OPTIMAL CONTROL OF PDES

March 11, 2023

J. FRÉDÉRIC BONNANS

Contents

This material is taken from the six first chapters in the lecture notes on 'Optimal control of partial differential equations' by J.F. Bonnans. Additional material on
https://pages.saclay.inria.fr/frederic.bonnans/co-edp.html
The complement sections refer to additional material presented in these lecture notes.

1. Various examples and motivation

See the lecture notes.

2. Abstract optimization

2.1. **Minimizing over a convex set.**

2.1.1. *Convex problems.* In what follows we assume that X is a Banach space. We say that $K \subset X$ is **convex** if $tx + (1-t)y \in K$, for all $t \in]0,1[$, and x, y in K. We say that $f : X \to \mathbb{R} \cup \{\pm\infty\}$ is **convex** if, for all $t \in]0,1[$, and x, y in $\operatorname{dom}(f)$:

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y). \tag{1}$$

Note that this implies that the domain of f is convex. We say that f is **strictly convex** if the above inequality is strict whenever $x \neq y$. The following is well-known.

Lemma 2.1. (i) Assume that f and K are convex. Then $F(P)$ and $S(P)$ are convex. (ii) If in addition f is strictly convex, then $S(P)$ has at most one element.

We often have **composite cost functions** of the form

$$\begin{cases} f(x) = g(x) + h(Ax); \quad A \in L(X,Y), \\ g : X \to \mathbb{R} \cup \{\pm\infty\}; \quad h : Y \to \mathbb{R} \cup \{\pm\infty\}, \end{cases} \tag{2}$$

Y being another Banach space. If g and h are convex, so is f. If in addition g is strictly convex, so is f. If g is convex, and h is strictly convex, we have a kind of partial uniqueness of solutions:

Lemma 2.2. (i) Assume that K is convex, g is convex, and h is strictly convex. Let x' and x'' belong to $S(P)$. Then $Ax' = Ax''$.

2.1.2. *Strong convexity.* We say that $f : X \to \mathbb{R} \cup \{\pm\infty\}$ is, for $\alpha > 0$, α-**strongly convex** if, for all $t \in]0,1[$, and x, y in $\operatorname{dom}(f)$:

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \tfrac{1}{2}\alpha t(1-t)\|y - x\|^2. \tag{3}$$

Clearly the α-convexity implies the strict convexity.

Lemma 2.3. (i) Assume that f is l.s.c. and α-convex, K is closed and convex, and $\operatorname{dom}(f) \cap K \neq \emptyset$. Then (P) has a unique solution say $\bar{x}$, and any minimizing sequence converges to $\bar{x}$.

Remark 2.4. In the case of a composite cost function of the form (2), if K is closed and convex, and g and h are l.s.c. convex, with h α-convex, and if $S(P)$ is nonempty, then we can prove in the same way that there exists $\bar{y} \in Y$ such that $\bar{y} = A\bar{x}$ for all $\bar{x} \in S(P)$, and if x_k is a minimizing sequence, then $Ax_k \to \bar{y}$.

2.1.3. *First-order optimality conditions.*

Definition 2.5. Let X be a Banach space, K be a convex subset of X, and $f : X \to \mathbb{R}$ We say that $\bar{x} \in K$ is a **local minimum** of f over K if, for some $\varepsilon > 0$:

$$f(\bar{x}) \leq f(x), \quad \text{for all } x \in K, \quad \|x - \bar{x}\| \leq \varepsilon. \tag{4}$$

Lemma 2.6. *Let $\bar{x} \in K$ be a local minimum of f over K, and f be differentiable at $\bar{x}$. Then*

$$Df(\bar{x})(x - \bar{x}) \geq 0, \quad \textit{for all } x \in K. \tag{5}$$

Conversely, if $\bar{x} \in K$ satisfies (5)*, and f is convex, then f attains its minimum over K at $\bar{x}$.*

2.2. **Normal cones.**

2.2.1. *Setting.* Let X be a Banach space. We say that $C \subset X$ is a cone if $\alpha x \in C$, for all $\alpha > 0$ and $x \in C$. If C is a convex cone, it is stable by addition:

$$\text{If } x \in C \text{ and } x' \in C, \text{ then } x + x' \in C. \tag{6}$$

Let K be a closed convex subset of X. The **normal cone** to K at $\bar{x} \in K$ is defined as

$$N_K(\bar{x}) := \{x^* \in X^*; \quad \langle x^*, x - \bar{x}\rangle_X \leq 0, \ \text{ for all } x \in K\}. \tag{7}$$

For instance, if $f : X \to \mathbb{R}$ attains a local minimum over K at $\bar{x} \in K$, and is differentiable at $\bar{x}$, we have established that

$$Df(\bar{x})(x - \bar{x}) \geq 0, \ \text{ for all } x \in K. \tag{8}$$

This is equivalent to

$$-Df(\bar{x}) \in N_K(\bar{x}). \tag{9}$$

So, it is useful to characterize normal cones.

Example 2.7. If $X = \mathbb{R}^n$ and $K = \mathbb{R}^n_-$, corresponding to the constraint $x \leq 0$, one easily checks that

$$N_K(\bar{x}) = \{\lambda \in \mathbb{R}^n_+; \ \lambda_i \cdot \bar{x}_i = 0, \quad i = 1, \ldots, n\}, \tag{10}$$

or equivalently

$$N_K(\bar{x}) = \{\lambda \in \mathbb{R}^n_+; \ \lambda \cdot \bar{x} = 0\} = \mathbb{R}^n_+ \cap \bar{x}^\perp, \tag{11}$$

where we denote the orthogonal of $x \in X$ by $x^\perp := \{x^* \in \mathbb{R}^n; \ \ x^* \cdot x = 0\}$.

With a cone C, subset of a Banach space X, is associated the (convex) dual **polar cone**

$$C^- := \{x^* \in X^*; \ \langle x^*, x\rangle_X \leq 0, \ \text{ for all } x \in C\}. \tag{12}$$

For instance, the polar cone of $\mathbb{R}^n_-$ is $\mathbb{R}^n_+$.

Lemma 2.8. *If K is a convex closed cone, then*

$$N_K(\bar{x}) = K^- \cap \bar{x}^\perp. \tag{13}$$

2.3. **Abstract optimal control.**

2.3.1. *Linear state equations.* Consider the problem

$$\operatorname*{Min}_{u\in U} J(u,y); \quad Ay = Bu + f; \quad u \in K_U, \ \ y \in K_Y. \tag{14}$$

Here K_U is a subset of the Banach space U, K_Y is a subset of the Banach space Y, $J : U \times Y \to \mathbb{R}$ is of class C^1, $A \in L(Y,W)$, where W is another Banach space, $B \in L(U,W)$, and $f \in W$.

We assume that A is bijective. Remember that, by the **open mapping theorem** (see Brézis [2]), its inverse is continuous. So, in the sequel, by **invertible** operator we will mean a continuous and bijective linear mapping.

Then the **state equation**

$$Ay = Bu + f \quad \text{in } W \tag{15}$$

has, for any $u \in U$, a unique solution denoted by

$$y[u] := A^{-1}(Bu + f). \tag{16}$$

The **linearized state equation**

$$Az = Bv \quad \text{in } W \tag{17}$$

has, for any $v \in U$, a unique solution in Y denoted by

$$z[v] := A^{-1}Bv. \tag{18}$$

The **reduced cost** $J_R(u) := J(u, y[u])$is of class C^1 and, by the chain rule, for any u and v in U, and $y = y[u]$:

$$\langle DJ_R(u), v\rangle_U = \langle D_u J(u,y), v\rangle_U + \langle D_y J(u,y), z[v]\rangle_Y. \tag{19}$$

The last term satisfies

$$\langle D_y J(u,y), z[v]\rangle_Y = \langle D_y J(u,y), A^{-1}Bv\rangle_Y = \langle B^\dagger A^{-\dagger} D_y J(u,y), v\rangle_Y. \tag{20}$$

Remember that since A is invertible, so is $A^\dagger$, and $(A^\dagger)^{-1} = (A^{-1})^\dagger$ can be denoted by $A^{-\dagger}$. Therefore the **costate equation**

$$A^\dagger p = D_y J(u, y[u]) \quad \text{in } Y^* \tag{21}$$

has a unique solution $p[u] \in W^*$, called the **costate associated with u**. We deduce that

$$\langle DJ_R(u), v\rangle_U = \langle D_u J(u,y), v\rangle_U + \langle B^\dagger p[u], v\rangle_U. \tag{22}$$

We have proved that

Lemma 2.9. *The derivative of the reduced cost at $u \in U$ is, for $y = y[u]$ and $p = p[u]$:*

$$DJ_R(u) = D_u J(u,y) + B^\dagger p. \tag{23}$$

2.3.2. *Control constraints.* We come back to the optimal control problem (14), assuming that K_U is nonempty, closed, convex and that $K_Y = Y$. Let $\bar{u}$ be a local minimum. By lemma 2.6, we have that

$$DJ_R(\bar{u})(v - \bar{u}) \geq 0, \quad \text{for all } v \in K_U. \tag{24}$$

In particular, consider the case when U is a Hilbert space and

$$J(u,y) = \tfrac{1}{2}\|u\|_U^2 + g(y), \tag{25}$$

with $g : Y \to \mathbb{R}$ of class C^1. For $\bar{u} \in K_U$, with associated state $\bar{y}$ and costate $\bar{p}$ solution of the costate equation $A^\dagger \bar{p} = Dg(\bar{y})$, the derivative of the reduced cost is $DJ_R(\bar{u}) = B^\dagger \bar{p} + \bar{u}$. So, the first-order optimality condition is

$$(B^\dagger \bar{p} + \bar{u}, v - \bar{u})_U \geq 0, \quad \text{for all } v \in K_U. \tag{26}$$

This is equivalent to

$$\bar{u} = P_{K_U}(-B^\dagger \bar{p}). \tag{27}$$

2.3.3. *Nonlinear state equations.* We now consider a variant of the previous problem, using the **implicit function theorem (IFT)** in order to establish the local well-posedness of a nonlinear state equation. Consider the problem

$$\operatorname*{Min}_{u\in U}\ J(u,y); \quad \mathcal{A}(y,u)=0; \quad u\in K_U,\ \ y\in K_Y. \tag{28}$$

Here K_U is a subset of the Banach space U, K_Y is a subset of the Banach space Y, $J : U \times Y \to \mathbb{R}$ is of class C^1, and $\mathcal{A} : U \times Y \to W$ is of class C^1, where W is another Banach space.

Theorem 2.10. [Implicit function theorem] *Let $(u_0, y_0) \in U \times Y$ be a zero of $\mathcal{A}$, i.e., $\mathcal{A}(u_0, y_0) = 0$, with $D_y\mathcal{A}(u_0, y_0)$, element of $L(Y,W)$, invertible.*
Then there exist open neighborhoods $\mathcal{U}$ of u_0 and $\mathcal{Y}$ of y_0, and a C^1 function $\varphi : \mathcal{U} \to Y$ such that, for any $(u,y) \in \mathcal{U}\times\mathcal{Y}$, (u,y) is a zero of $\mathcal{A}$ iff $y = \varphi(u)$. If in addition $\mathcal{A}$ is of class C^m for some $m \geq 2$, then φ is also of class C^m.

For the proof we refer to classical textbooks. In particular $y_0 = \varphi(u_0)$. Differentiating the relation $\mathcal{A}(u, \varphi(u)) = 0$ for u in $\mathcal{U}$, we obtain that

$$D_y\mathcal{A}(u,\varphi(u))D\varphi(u) + D_u\mathcal{A}(u,\varphi(u)) = 0. \tag{29}$$

It can be proved that the set of invertible mappings is open. For u close enough to u_0, $D_y\mathcal{A}(u, \varphi(u))$ is close to $D_y\mathcal{A}(u_0, y_0)$ and therefore invertible, so that

$$D\varphi(u) = -[D_y\mathcal{A}(u,\varphi(u))]^{-1}D_u\mathcal{A}(u,\varphi(u)). \tag{30}$$

In particular,

$$D\varphi(u_0) = -[D_y\mathcal{A}(u_0,y_0)]^{-1}D_u\mathcal{A}(u_0,y_0). \tag{31}$$

In the sequel $(\bar{u}, \bar{y})$ will be a zero of $\mathcal{A}$ such that $D_y\mathcal{A}(\bar{u}, \bar{y})$ is invertible, and for u close to $\bar{u}$, we will denote by $y[u]$ the state associated with u provided by the implicit function theorem (there might exist other solutions of the state equation, far from $\bar{y}$). The derivative of the reduced cost

$$J_R(u) := J(u, y[u]) \tag{32}$$

in the direction $v \in U$ satisfies

$$\begin{aligned} DJ_R(u)v &= \langle D_uJ(u,y[u]), v\rangle + \langle D_yJ(u,y[u]), Dy[u]v\rangle \\ &= \langle D_uJ(u,y[u]) + Dy[u]^\dagger D_yJ(u,y[u]), v\rangle. \end{aligned} \tag{33}$$

By the implicit function theorem,

$$\begin{aligned} Dy[u]^\dagger D_yJ(u,y[u]) &= -[(D_y\mathcal{A}(u,y[u]))^{-1}D_u\mathcal{A}(u,y[u])]^\dagger D_yJ(u,y[u]) \\ &= D_u\mathcal{A}(u,y[u])^\dagger p[u], \end{aligned} \tag{34}$$

where the **costate** $p[u]$ is the unique solution of the **costate equation**

$$-D_y\mathcal{A}(u,y[u])^\dagger p[u] = D_yJ(u,y[u]). \tag{35}$$

It follows that:

Lemma 2.11. *The derivative of the reduced cost at $u \in U$ is*

$$DJ_R(u) = D_uJ(u,y[u]) + D_u\mathcal{A}(u,y[u])^\dagger p[u]. \tag{36}$$

2.4. **Complements.** Rules for computing normal cones, based on a qualification condition.

3. Functional analysis

3.1. Weak convergence.

Given a sequence $\{x_k\}$ in a Banach space X, and $\bar{x} \in X$, we say that x_k **weakly converges** to $\bar{x}$, and write $x_k \rightharpoonup \bar{x}$ if

$$\langle x^*, x_k \rangle_X \to \langle x^*, \bar{x} \rangle_X, \quad \text{for all } x^* \in X^*. \tag{37}$$

Lemma 3.1. *Any weakly convergent sequence is bounded and, if X is reflexive, any bounded sequence has a weakly convergent subsequence.*

Proof. See Brézis [2]. □

Lemma 3.2. *Let X be a Hilbert space and $x_k \rightharpoonup \bar{x}$ in X. Then*

$$\|\bar{x}\|_X \le \liminf_k \|x_k\|_X \le \limsup_k \|x_k\|_X, \tag{38}$$

with equality in both inequalities iff $x_k \to x$.

3.1.1. *Closed convex sets are weakly sequentially closed.* This follows from the Hahn-Banach Theorem. An important consquence is that any l.s.c. convex function on a Banach space is weakly l.s.c.

3.1.2. *Weak∗ convergence.* Let X be a Banach space. We say that the sequence x_k^* in X^* weakly∗ converges to $x^* \in X^*$ if

$$\langle x_k^*, x \rangle \to \langle x^*, x \rangle, \quad \text{for all } x \in X. \tag{39}$$

Note that, if X is reflexive, then the weak and weak∗ convergence coincide. In general, we can only say that the former implies the latter. Then, see Brézis [2], ch. 3, Coro. 3.30:

Lemma 3.3. *Let X be a separable Banach space (i.e., there exists a dense sequence). Then any bounded sequence in X^* has a weakly∗ converging subsequence.*

3.2. Fourier transform.

The *Fourier transform* of $f \in L^1(\mathbb{R}^n)$ is defined as[1] $\mathcal{F}(f) : \mathbb{R}^n \to \mathbb{C}$, such that

$$\mathcal{F}(f)(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i x \cdot \xi} \mathrm{d}x, \tag{40}$$

where "·" denotes the scalar product in $\mathbb{R}^n$. By the dominated convergence theorem, $\hat{f}$ is continuous and bounded, and $\|\hat{f}\|_\infty \le \|f\|_1$. By Malliavin [9], thm 2.4.0.2 p. 127, we also have $\hat{f}(\xi) \to 0 \quad$ when $|\xi| \to \infty$.Let $\hat{C}_0(\mathbb{R}^n)$ be the space of continuous complex valued functions over $\mathbb{R}^n$, with limit 0 at infinity, endowed with the uniform norm. We have proved that $f \to \hat{f}$ is **nonexpansive** $L^1(\mathbb{R}^n) \to \hat{C}_0(\mathbb{R}^n)$.

3.2.1. *Fourier transform in L^2.* It happens that if $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, then $\hat{f} \in L^2(\mathbb{R}^n)$, and the **Plancherel formula** holds:

$$\|\hat{f}\|_2 = \|f\|_2. \tag{41}$$

Since $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ is a dense subset of $L^2(\mathbb{R}^n)$, the Fourier transform has a continuous extension to $L^2(\mathbb{R}^n)$, called the **Fourier transform in** $L^2(\mathbb{R}^n)$. The latter is **isometric**, i.e., (41) holds for all $f \in L^2(\mathbb{R}^n)$.

[1] Several possible conventions: we follow [3].

3.2.2. *Fourier transform of derivatives.* Let $f \in L^1(\mathbb{R}^n)$, such that $\delta_\varepsilon f \to g$ in $L^1(\mathbb{R}^n)$ where for $\varepsilon > 0$:

$$\delta_\varepsilon f(x) := (f(x + \varepsilon e_j) - f(x))/\varepsilon. \tag{42}$$

Lemma 3.4. *We have that $\hat{g}(\xi) = 2\pi i \xi_j \hat{f}(\xi)$ (equality in $\hat{C}_0(\mathbb{R}^n)$), or in other words*

$$\widehat{\frac{\partial f}{\partial x_j}}(\xi) = 2\pi i \xi_j \hat{f}(\xi). \tag{43}$$

Theorem 3.5. *Let $f \in L^2(\mathbb{R}^n)$ has a partial derivative g with respect to x_j, in the L^2 sense. Then $\hat{g}(\xi) = 2\pi i \xi_j \hat{f}(\xi)$ (equality in the space $\hat{L}^2(\mathbb{R}^n)$ of complex valued square summable functions).*

3.2.3. *Spaces $W^{m,p}(\Omega)$.* For Ω domain (= open subset) of $\mathbb{R}^n$, and $p \in [1, \infty)$, the Sobolev space of function over Ω, with weak partial derivatives of order at most $m \in \mathbb{N}$ in $L^p(\Omega)$, is denoted by $W^{m,p}(\Omega)$ and endowed with the norm

$$\|f\|_{W^{m,p}(\Omega)} := \left(\sum_{|\alpha| \le m} \|D^\alpha f\|^p_{L^p(\Omega)} \right)^{1/p}. \tag{44}$$

When $p = 2$ we redefine it as the Hilbert space $H^m(\Omega)$. We define in a similar way $W^{m,\infty}(\Omega)$, endowed with the norm

$$\|f\|_{W^{m,\infty}(\Omega)} := \sum_{|\alpha| \le m} \|D^\alpha f\|_{L^\infty(\Omega)}. \tag{45}$$

3.2.4. *Fractional Sobolev space.* Following Grisvard [4], for $s = m + \sigma$, with $m \in \mathbb{N}$ and $\sigma \in (0, 1)$, and $p \in [1, \infty]$, set

$$W^{s,p}(\mathbb{R}^n) := \{u \in W^{m,p}(\mathbb{R}^n); \quad N_{s,p}(u) < \infty\}, \tag{46}$$

where

$$N_{s,p}(u) := \sum_{|\alpha| = m} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \frac{|D^\alpha u(x) - D^\alpha u(y)|^p}{|x - y|^{n + \sigma p}} \mathrm{d}x \mathrm{d}y. \tag{47}$$

This space is endowed with the norm

$$\|u\|_{W^{s,p}(\mathbb{R}^n)} := \left(\|u\|^p_{W^{m,p}(\mathbb{R}^n)} + N_{s,p}(u) \right)^{1/p}. \tag{48}$$

This is a Banach space (a Hilbert space if $p = 2$).

Theorem 3.6. *Let $f \in L^2(\mathbb{R}^n)$. Then for all $m \in \mathbb{N}$:*

$$f \in H^m(\mathbb{R}^n) \quad \Leftrightarrow \quad (1 + |\xi|^m)\hat{f}(\xi) \in L^2(\mathbb{R}^n). \tag{49}$$

3.2.5. *Spaces $H^s(\mathbb{R}^n)$.* Define, for $s \ge 0$:

$$H^s(\mathbb{R}^n) := \{f \in L^2(\mathbb{R}^n); |\xi|^s \hat{f}(\xi) \in L^2(\mathbb{R}^n)\}, \tag{50}$$

with natural norm

$$\|f\|_{H^s(\mathbb{R}^n)} := \|(1 + |\xi|^2)^{s/2} \hat{f}(\xi)\|_2. \tag{51}$$

For $s \in \mathbb{N}$ we recover the $H^m(\mathbb{R}^n)$ spaces previously defined. Set for $s > 0$:

$$H^{-s}(\mathbb{R}^n) := H^s(\mathbb{R}^n)^*. \tag{52}$$

3.3. **Elliptic equations.**

Theorem 3.7. *Given* $f \in L^2(\mathbb{R}^n)$*, the equation*

$$u(x) - \Delta u(x) = f(x), \quad x \in \mathbb{R}^n, \tag{53}$$

has a unique solution u *in* $H^2(\mathbb{R}^n)$*, and for all* $1 \leq i \leq j \leq n$*:*

$$\|D^2_{x_i x_j} u\|_2 \leq \|f\|_2. \tag{54}$$

3.4. **Linear-quadratic setting.**

$$y(x) - \Delta y(x) = f(x) + u(x), \quad x \in \mathbb{R}^n. \tag{55}$$

with $f \in L^2(\mathbb{R}^n)$ given and $u \in L^2(\mathbb{R}^n)$. Cost function

$$J(u, y) := \tfrac{1}{2} \int_{\mathbb{R}} (y(x) - y_d(x))^2 \mathrm{d}x + \tfrac{1}{2} \int_{\mathbb{R}} u(x)^2 \mathrm{d}x. \tag{56}$$

Optimal control problem

$$\operatorname*{Min}_{u,y} J(u, y) \quad \text{s.t. (55);} \quad u \in K_U, \tag{P}$$

with K_U nonempty closed convex subset of $U := L^2(\mathbb{R}^n)$. The state space is $Y := H^2(\mathbb{R}^n)$. For each $u \in U$, the state equation (55) has a unique solution denoted by $y[u]$ in Y, and $u \mapsto y[u]$ is affine and continuous $U \to Y$. The **reduced cost function** is

$$J_R(u) := J(u, y[u]). \tag{57}$$

Being continuous and strongly convex, it has a unique minimizer $\bar{u}$ over K_U, and each minimizing sequence u_k in K_U strongly converges to $\bar{u}$.

Reduction Lagrangian: sum of cost function and product of state equation by a multiplier:

$$\mathcal{L}(u, y, p) := J(u, y) + \int_{\mathbb{R}^n} p(x) \left(f(x) + u(x) + \Delta y(x) - y(x) \right) \mathrm{d}x. \tag{58}$$

A priori the space for the state equation is $L^2(\mathbb{R}^n)$, so we should take the multiplier in the dual space $L^2(\mathbb{R}^n)$, but we decide to search for it in the smaller space Y. Then by Green's theorem over $Y \times Y$ (valid over the dense subset $\mathcal{D}(\mathbb{R}^n)^2$ and extended by continuity):

$$\int_{\mathbb{R}^n} p(x) \Delta y(x) \mathrm{d}x = \int_{\mathbb{R}^n} y(x) \Delta p(x) \mathrm{d}x, \tag{59}$$

so that for $y = y[u]$ and $z \in Y$:

$$\mathcal{L}_y z = \int_{\mathbb{R}^n} (y(x) - y_d(x) + \Delta p(x) - p(x)) z(x) \mathrm{d}x. \tag{60}$$

For $y = y[u]$, this gives the costate equation, with unique solution $p[u]$ in Y:

$$p(x) - \Delta p(x) = y(x) - y_d(x), \quad x \in \mathbb{R}^n. \tag{61}$$

The linearized state equation, with $(v, z) \in U \times Y$, is:

$$z(x) - \Delta z(x) = v(x), \quad x \in \mathbb{R}^n. \tag{62}$$

And we get as expected, writing $z = z[v]$ and $p = p[v]$:

$$
\begin{aligned}
DJ_R(u)v &= \int_{\mathbb{R}^n} ((y(x) - y_d(x))z(x) + u(x)v(x))\mathrm{d}x, \\
&= \int_{\mathbb{R}^n} ((p(x) - \Delta p(x))z(x) + u(x)v(x))\mathrm{d}x, \\
&= \int_{\mathbb{R}^n} ((z(x) - \Delta z(x))p(x) + u(x)v(x))\mathrm{d}x, \\
&= \int_{\mathbb{R}^n} (p(x) + u(x))v(x)\mathrm{d}x.
\end{aligned} \tag{63}
$$

Since $U = L^2(\Omega)$, this means that the derivative of J_R is

$$DJ_R(u) = p[u] + u. \tag{64}$$

Let $\bar{u}$ be solution of the optimal control problem. Set $\bar{y} = y[\bar{u}]$, $\bar{p} = p[\bar{u}]$. if J_R is strictly convex, $(\bar{u}, \bar{y}, \bar{p})$ is the unique solution in $U \times Y \times Y$ of the optimality system

$$
\left\{
\begin{array}{l}
\bar{y} - \Delta\bar{y} = f + \bar{u}; \\
\bar{p} - \Delta\bar{p} = \bar{y} - y_d; \\
\displaystyle\int_{\mathbb{R}^n} (\bar{p}(x) + \bar{u}(x))(v(x) - \bar{u}(x))\mathrm{d}x \geq 0, \quad \text{for all } v \in K_U.
\end{array}
\right. \tag{65}
$$

The above inequality is equivalent to $\bar{u} = P_{K_U}(-\bar{p})$. Eliminating $\bar{u}$, we obtain that $(\bar{y}, \bar{p})$ is the unique solution in $Y \times Y$ of

$$
\left\{
\begin{array}{ll}
\bar{y} - \Delta\bar{y} & = f + P_{K_U}(-\bar{p}); \\
\bar{p} - \Delta\bar{p} & = \bar{y} - y_d.
\end{array}
\right. \tag{66}
$$

3.5. **Complements.** Parabolic optimal control problems.

4. Lax-Milgram

4.1. **General setting.** Let V be a Hilbert space, $a(\cdot, \cdot)$ a bilinear form on V, that is both **continuous and coercive**, that is, for some constants $c > 0$ and $\alpha > 0$:

$$
\left\{
\begin{array}{llll}
\text{(i)} & |a(u, v)| & \leq & c\|u\|_V\|v\|_V, \quad \text{for all } u,\ v \text{ in } V, \\
\text{(ii)} & a(u, u) & \geq & \alpha\|u\|_V^2, \quad\quad\;\; \text{for all } u \text{ in } V.
\end{array}
\right. \tag{67}
$$

Lemma 4.1 (*Lax-Milgram*)**.** *Given a continuous linear form L on V, there exists a unique $u \in V$ such that*

$$a(u, v) = L(v), \quad \textit{for all } v \in V, \tag{68}$$

and it holds that $\|u\|_V \leq \|L\|/\alpha$.

Given $u \in V$, the mapping $v \mapsto a(u, v)$ defines a linear form on V that we may denote by $A[u]$. We see that the equation (68) is equivalent to

$$A[u] = L \quad \text{in } V^*. \tag{69}$$

4.2. **Monotonicity.** We assume that V is a space of functions defined a.e. on a certain set Ω. For $u \in V$, define its positive part by $u_+(x) := \max(u(x), 0)$, for $x \in \Omega$. Let V_+ denote the set of nonnegative (i.e., equal to their positive part) elements of V. We say that $L \in V^*$ is nonnegative if $L(v) \geq 0$, for all $v \in V_+$.

Lemma 4.2. *Assume that, for all $u \in V$, $u_+ \in V$, and $a(u, u_+) = a(u_+, u_+)$. Then the solution u to* (68) *belongs to V_+ whenever L is nonnegative.*

4.3. **Optimal control.** Let U, Y, W be Hilbert spaces, $B \in L(U, Y^*)$, $A \in L(Y, Y^*)$ be such that the associated bilinear form $a(y, z) := \langle Ay, z\rangle_Y$ is continuous and coercive. Consider the problem

$$\text{Min } J(u, y) := \tfrac{1}{2}\|u\|_U^2 + \tfrac{1}{2}\|Cy[u] - y_d\|_W^2; \quad u \in K, \tag{70}$$

where K is a nonempty closed convex subset of U, $y_d \in W$, $C \in L(Y, W)$ and $y[u]$ is the unique solution of

$$Ay = f + Bu \quad \text{in } Y^*. \tag{71}$$

The reduced cost is $J_R(u) := J(u, y[u])$. The Lagrangian of the problem is

$$\mathcal{L}(u, y, p) := J(u, y) + \langle f + Bu - Ay, p\rangle_Y = J(u, y) - a(y, p) + \langle f + Bu, p\rangle_Y \tag{72}$$

so that the costate equation (in variational form) is

$$0 = \mathcal{L}_y(u, y, p)z = ((Cy - y_d, Cz)_W - a(z, p), \quad \text{for all } z \in Y, \tag{73}$$

or equivalently $p = p[u]$ is defined by, for $y = y[u]$:

$$A^* p = C^*(Cy - y_d) \quad \text{in } Y^*. \tag{74}$$

We obtain the following resut:

Lemma 4.3. *The reduced cost is of class C^1, with derivative defined by*

$$DJ_R(u) = \mathcal{L}_u(u, y, p) = u + B^* p. \tag{75}$$

Remark 4.4. The previous lemma is a particular case of the abstract setting in section 2, the space W for the state equation being here equal to Y^*.

4.4. **Complements.** See the analysis of identification of coefficients.

5. Boundary value problems

5.1. **Green formulas.** Let Ω be an open subset of $\mathbb{R}^n$ with smooth enough boundary $\partial\Omega$ and **outward normal** at $x \in \partial\Omega$, denoted by $\nu(x)$. **Green's formula** says that, for all $i \in \{1, \dots, n\}$ and z of class $C^1 : \bar{\Omega} \to \mathbb{R}$:

$$\int_\Omega z_{x_i}(x)\mathrm{d}x = \int_{\partial\Omega} z(x)\nu_i(x)\mathrm{d}x. \tag{76}$$

If $z(x) = v(x)w(x)$, with v and w of class C^1 over $\bar{\Omega}$, we get:

$$\int_\Omega v w_{x_i}\mathrm{d}x + \int_\Omega v_{x_i} w\mathrm{d}x = \int_{\partial\Omega} v w \nu_i \mathrm{d}x. \tag{77}$$

Corollary 5.1. *Let $\Phi(x)$ be a C^1 vector field (function with image in $\mathbb{R}^n$) over $\bar{\Omega}$, and $v(x) : \bar{\Omega} \to \mathbb{R}$ be C^1. Then*

$$\int_\Omega v \operatorname{div}\Phi\mathrm{d}x + \int_\Omega \Phi\cdot\nabla v = \int_{\partial\Omega} v\Phi\cdot\nu\mathrm{d}x. \tag{78}$$

In particular, taking $v = 1$, we get the ***Stokes formula***

$$\int_\Omega \operatorname{div}\Phi\mathrm{d}x = \int_{\partial\Omega} \Phi\cdot\nu\mathrm{d}x. \tag{79}$$

The Stokes formula says that the integral of a flux over a domain is equal to the integral of the (outward) normal flux over the boundary.

Let $u : \bar{\Omega} \to \mathbb{R}$ be of class C^1. Define its **normal derivative** of u at $x \in \partial\Omega$ as

$$u_\nu(x) := \nabla u(x)\cdot\nu(x) = \sum_{i=1}^n \frac{\partial u(x)}{\partial x_i}\nu_i(x), \quad x \in \partial\Omega. \tag{80}$$

Taking $\Phi(x) = \nabla u(x)$, so that $\operatorname{div}\Phi = \Delta u$, we deduce from (78) that

$$\int_\Omega v(x)\Delta u(x)\mathrm{d}x + \int_\Omega \nabla v(x)\cdot\nabla u(x)\mathrm{d}x = \int_{\partial\Omega} v(x)u_\nu(x)\mathrm{d}x. \tag{81}$$

Weighted gradient Consider the elliptic operator, where the a_{ij} are non necessarily symmetric, but positive semidefinite:

$$Au(x) = -\sum_{i=1}^n \frac{\partial}{\partial x_i}\left(\sum_{j=1}^n a_{ij}(x)\frac{\partial u(x)}{\partial x_j}\right). \tag{82}$$

This is a **divergence** of a weighted gradient:

$$Au(x) = -\operatorname{div}\nabla_A u; \quad (\nabla_A u)_i := \sum_{j=1}^n a_{ij}(x)\frac{\partial u(x)}{\partial x_j}. \tag{83}$$

Equivalently, setting $a(x) := (a_{ij}(x))_{ij}$, we get

$$\nabla_A u(x) = a(x)\nabla u(x). \tag{84}$$

Weighted normal derivatives and weighted normal direction For $x \in \partial\Omega$, set :

$$\nu_A(x) := a(x)^\dagger\nu(x), \;\text{ i.e., }\; (\nu_A)_j(x) := \sum_{i=1}^n a_{ij}(x)\nu_i(x),\; 1 \le j \le n. \tag{85}$$

Associated A-normal derivative for $x \in \partial\Omega$:

$$u_{\nu_A} := \nu_A(x)\cdot\nabla u(x) = \nu(x)^\dagger a(x)\nabla u(x) = \nabla_A u(x)\cdot\nu(x). \tag{86}$$

Equivalently

$$u_{\nu_A}(x) = \sum_{i,j=1}^n \nu_i(x)a_{ij}(x)\frac{\partial u(x)}{\partial x_j}. \tag{87}$$

Then by Green's formula:

$$\begin{aligned} -\int_\Omega v(x)Au(x)\mathrm{d}x = & \int_\Omega v(x)\operatorname{div}\nabla_A u(x)\mathrm{d}x \\ = & \int_{\partial\Omega} v(x)u_{\nu_A}(x)\mathrm{d}x - \int_\Omega \nabla v(x)\cdot\nabla_A u(x)\mathrm{d}x. \\ = & \int_{\partial\Omega} v(x)u_{\nu_A}(x)\mathrm{d}x - \int_\Omega \sum_{i,j=1}^n a_{ij}(x)\frac{\partial v(x)}{\partial x_i}\frac{\partial u(x)}{\partial x_j}\mathrm{d}x. \end{aligned} \tag{88}$$

Define the **formal adjoint** $A^\dagger$ of A by

$$A^\dagger u(x) = -\sum_{i=1}^n \frac{\partial}{\partial x_i}\left(\sum_{j=1}^n a_{ji}(x)\frac{\partial u(x)}{\partial x_j}\right), \tag{89}$$

and set

$$\nu_{A^\dagger}(x) := a(x)\nu(x), \;\text{ i.e., }\; (\nu_{A^\dagger})_j(x) := \sum_{i=1}^n a_{ji}(x)\nu_i(x),\; 1 \le j \le n. \tag{90}$$

Associated $A^\dagger$-normal derivative for $x \in \partial\Omega$:

$$u_{\nu_{A^\dagger}} := \nu(x)^\dagger a(x)^\dagger\nabla u(x) = \nabla_{A^\dagger}u(x)\cdot\nu(x). \tag{91}$$

Equivalently

$$u_{\nu_{A^\dagger}}(x) = \sum_{i,j=1}^n \nu_i(x)a_{ji}(x)\frac{\partial u(x)}{\partial x_j}. \tag{92}$$

Then by Green's formula:

(93)
$$
\begin{aligned}
-\int_\Omega v(x)A^\dagger u(x)\mathrm{d}x = & \int_\Omega v(x)\,\mathrm{div}\,\nabla_{A^\dagger} u(x)\mathrm{d}x \\
= & \int_{\partial\Omega} v(x)u_{\nu_{A^\dagger}}(x)\mathrm{d}x - \int_\Omega \nabla v(x)\cdot\nabla_{A^\dagger} u(x)\mathrm{d}x \\
= & \int_{\partial\Omega} v(x)u_{\nu_{A^\dagger}}(x)\mathrm{d}x - \int_\Omega \sum_{i,j=1}^n a_{ji}(x)\frac{\partial v(x)}{\partial x_i}\frac{\partial u(x)}{\partial x_j}\mathrm{d}x.
\end{aligned}
$$

5.2. **Sobolev spaces on general domains.**

5.2.1. *Integer order spaces.* Let Ω be an open subset of $\mathbb{R}^n$, $m \in \mathbb{N}^*$, $p \in [1,\infty]$. We recall that $W^{m,p}(\Omega)$ denotes the space of functions over Ω having weak derivatives up to order m in $L^p(\Omega)$, endowed with the norm

$$
\|f\|_{W^{m,p}(\Omega)} := \left(\sum_{|\alpha|\le m} \|D^\alpha f\|^p_{L^p(\Omega)}\right)^{1/p}, \tag{94}
$$

for $p < \infty$ and $\|f\|_{W^{m,\infty}(\Omega)} := \max\{\|D^\alpha f\|_{L^\infty(\Omega)};\ |\alpha| \le m$, with $\alpha \in \mathbb{N}^n$ and $|\alpha| = \sum_i \alpha_i$. Remember that $H^m(\Omega) := W^{m,2}(\Omega)$ is a Hilbert space. We denote the closure of $\mathcal{D}(\Omega)$ in $W^{m,p}(\Omega)$ by

$$
W_0^{m,p}(\Omega) := \overline{\mathcal{D}(\Omega)}^{W^{m,p}(\Omega)}. \tag{95}
$$

5.3. **Representation of the dual: abstract setting.**

5.3.1. *General results.* Let X, Y be Banach spaces. Given $M \in L(X,Y)$ and injective, assume that

$$
\|x\|_X = \|Mx\|_Y. \tag{96}
$$

This applies for instance to the case of Sobolev spaces. We next characterize the dual of such spaces.

Lemma 5.2. *We have that $X^* = \{M^\dagger y^*;\ y^* \in Y^*\}$, and for all $x^* \in X^*$:*

$$
\|x^*\|_{X^*} = \min\{\|y^*\|_{Y^*};\ x^* = M^\dagger y^*\}. \tag{97}
$$

5.3.2. *Case of Hilbert spaces.* Now let X and Y be Hilbert spaces, with again $\|x\|_V = \|Mx\|_Y$ for some $M \in L(X,Y)$. Let $x^* \in X^*$. By the Riesz theorem, there exists a unique $x \in X$ solution of

$$
(x,x')_X = \langle x^*, x'\rangle_X, \quad \text{for all } x' \in X. \tag{98}
$$

We know that the primality mapping $x^* \mapsto x$, with inverse denoted by $\mathbf{\Lambda}_X$, is isometric: $X^* \to X$, and such that

$$
(x,x')_X = \langle \mathbf{\Lambda}_X x, x'\rangle_X, \quad \text{for all } x' \in X. \tag{99}
$$

Then $\mathcal{A} := M^*\mathbf{\Lambda}_Y M$ belongs to $L(X,X^*)$, and, for all x, x' in X:

$$
(x,x')_X = (Mx,Mx')_Y = \langle \mathbf{\Lambda}_Y Mx, Mx'\rangle_Y = \langle \mathcal{A}x, x'\rangle_X. \tag{100}
$$

It follows that $\mathbf{\Lambda}_X = \mathcal{A}$, which means that given $x^* \in X^*$, computing $\mathbf{\Lambda}_X^{-1}x^*$ amounts to solve $M^*\mathbf{\Lambda}_Y Mx = x^*$, or equivalently

$$
(Mx,Mx')_Y = \langle x^*, x'\rangle_X, \quad \text{for all } x' \in X. \tag{101}
$$

Remark 5.3. Sometimes we need to evaluate a scalar product in X^*. Let x_1^*, x_2^* belong to X^*, and $x_1 := \mathbf{\Lambda}_X^{-1}x_1^*$, $x_2 := \mathbf{\Lambda}_X^{-1}x_2^*$. Then, since $\mathbf{\Lambda}_X$ is isometric:

(102)
$$4(x_1^*,x_2^*)_{X^*} = \|x_1^*+x_2^*\|_{X^*}^2 - \|x_1^*+x_2^*\|_{X^*}^2 = \|x_1+x_2\|_X^2 - \|x_1+x_2\|_X^2 = 4(x_1,x_2)_X.$$

That is,

$$(x_1^*,x_2^*)_{X^*} = (x_1,x_2)_X = (Ax_1,Ax_2)_Y = (AM^{-1}x_1^*,AM^{-1}x_2^*)_Y. \tag{103}$$

Example 5.4. Let V be a closed subset of $H^1(\Omega)$, endowed with the norm induced by $H^1(\Omega)$. The previous setting applies with $X = V$, $Y = (L^2(\Omega))^{n+1}$, and for $v \in V$:

$$Mv := \left(v, \frac{\partial v}{\partial x_1}, \dots, \frac{\partial v}{\partial x_n}\right). \tag{104}$$

Then we can compute, identifying $H = L^2(\Omega)$ with its dual:

$$\mathcal{A}v = v + \sum_{i=1}^n \left(\frac{\partial}{\partial x_i}\right)^* \frac{\partial v}{\partial x_i}. \tag{105}$$

Or, to be more precise, for inV, $\mathcal{A}v \in V^*$ is defined by:

$$\langle \mathcal{A}v, v'\rangle_X = \int_\Omega (v(x)v'(x) + \nabla v(x)\cdot\nabla v'(x))\mathrm{d}x \tag{106}$$

5.4. **Representation of the dual of $W^{m,p}(\Omega)$.**

Lemma 5.5. *Let $p \in [1,\infty)$. With any $f \in W^{m,p}(\Omega)^*$ are associated $g_\alpha \in L^q(\Omega)$, $1/p + 1/q = 1$, $|\alpha| \le m$, such that*

$$\langle f,u\rangle_{W^{m,p}} = \sum_{|\alpha|\le m} \int_\Omega g_\alpha(x) D^\alpha u(x)\mathrm{d}x. \tag{107}$$

5.5. **Representation of the dual of $W_0^{m,p}(\Omega)$.** Let E be a closed subspace of a Banach space X. By the **Hahn Banach** theorem (see Brézis [2], ch.1), any $e^* \in E^*$ has an extension of same norm over X, that is, there exists $x^* \in X^*$ such that $\|x^*\|_{X^*} = \|e^*\|_{E^*}$ and $\langle x^*,e\rangle_X = \langle e^*,e\rangle_E$, for all $e \in E$.

Let $p \in [1,\infty)$. Since $W_0^{m,p}(\Omega)$ is, by the definition, a closed subspace of $W^{m,p}(\Omega)$, by the previous lemma and the Hahn Banach theorem, with any $f \in W_0^{m,p}(\Omega)^*$, are associated $g_\alpha \in L^q(\Omega)$, $1/p+1/q = 1$, $|\alpha| \le m$, such that

$$\langle f,u\rangle_{W_0^{m,p}} = \sum_{|\alpha|\le m} \int_\Omega g_\alpha(x) D^\alpha u(x)\mathrm{d}x. \tag{108}$$

Since $\mathcal{D}(\Omega)$ is a dense subset of $W_0^{m,p}(\Omega)$, we deduce that, in a weak sense: of **distributions**,

$$f = \sum_{|\alpha|\le m} (-1)^{|\alpha|} D^\alpha g_\alpha. \tag{109}$$

Actually the above display is a definition of elements of $W^{-m,p}(\Omega)$. So, the dual of $W_0^{m,p}(\Omega)$ is $W^{-m,p}(\Omega)$.

Example 5.6. Any $f \in H_0^1(\Omega)^*$ is of the form

$$f = g_0 - \sum_{i=1}^n \frac{\partial g_i}{\partial x_i}; \quad g_0,\dots,g_n \;\text{ in }\; L^2(\Omega). \tag{110}$$

Of course, the decomposition is not unique !

5.6. **Sobolev embedding theorems.** The domain $\Omega \subset \mathbb{R}^n$ is either equal to $\mathbb{R}^n$, an open half subspace, or a bounded set with C^∞ boundary. Some useful spaces are (if natural the norms are not explicited):

(1) $C^m_B(\Omega)$: space of functions having bounded, continuous derivatives of order α over Ω, whenever $|\alpha| \leq m$.
(2) $C^m(\bar{\Omega})$: space of functions having bounded, *uniformly* continuous derivatives of order α over Ω, whenever $|\alpha| \leq m$.
(3) $C^{m,\lambda}(\bar{\Omega})$: space of elements of $C^m(\bar{\Omega})$ whose derivatives up to order m are Hölder of exponent $\lambda \in (0,1)$; norm

$$\|f\|_{m,\lambda} := \max_{0\leq|\alpha|\leq m} \sup_{x\in\Omega} |D^\alpha f(x)| + \max_{0\leq|\alpha|\leq m} \sup_{\substack{x,y\in\Omega \\ x\neq y}} \frac{|D^\alpha f(x) - D^\alpha f(y)|}{|x-y|^\lambda}. \tag{111}$$

With the Sobolev space $W^{m,p}(\Omega)$, $\Omega \subset \mathbb{R}^n$, with $m \in \mathbb{N}$, $m \neq 0$, and $p \in [1,\infty]$, we associate the amount

$$D(m,n,p) := 1/p - m/n. \tag{112}$$

Theorem 5.7. *The following holds:*

(1) *If $D(m,n,p) > 0$, let $q^* := 1/D(m,n,p) = pn/(n-pm)$, or equivalently*
$$1/q^* = 1/p - m/n. \tag{113}$$
Then $q^* > p$ and $W^{m,p}(\Omega) \subset L^q(\Omega)$, for all $q \in [p, q^*]$. The embedding is compact if Ω is bounded and $q < q^*$.
(2) *If $D(m,n,p) = 0$, then $W^{m,p}(\Omega) \subset L^q(\Omega)$, for all $p \leq q < \infty$, with compact embedding if Ω is bounded.*
(3) *If $D(m,n,p) < 0$, then $W^{m,p}(\Omega) \subset C^0_B(\bar{\Omega})$, with compact embedding when Ω is bounded. Even more, if*
$$(m-1)p < n < mp, \tag{114}$$
then $W^{m,p}(\Omega) \subset C^{0,\lambda}(\bar{\Omega})$, where $\lambda \leq m - n/p$ (compact embedding if Ω is bounded and strict inequality).

Example 5.8. Let $m \in \mathbb{N}_*$. If $1/2 - m/n > 0$, then $H^m(\Omega) \subset L^{q^*}(\Omega)$, $1/q^* := 1/2 - m/n$. In particular, for $n = 3$, $H^1(\Omega) \subset L^6(\Omega)$, and elements of the unit ball of $H^m(\Omega)$ are uniformly Hölder continuous as soon as $m \geq 2$. For $n = 4$, $H^1(\Omega) \subset L^4(\Omega)$ and $H^2(\Omega) \subset L^q(\Omega)$, for all $q \in [2,\infty)$.

5.7. **Homogeneous Dirichlet conditions and Neumann conditions.**

5.7.1. *Homogeneous Dirichlet conditions.* Let $f \in L^2(\Omega)$. We want to solve

$$u(x) - \Delta u(x) = f(x), \; x \in \Omega \tag{115}$$

with **homogeneous Dirichlet boundary condition**

$$u = 0 \text{ on } \partial\Omega. \tag{116}$$

The corresponding variational formulation is: find $u \in H^1_0(\Omega)$ such that

$$\int_\Omega (u(x)v(x) + \nabla u(x)\cdot\nabla v(x))\mathrm{d}x = \int_\Omega f(x)v(x)\mathrm{d}x, \quad \text{for all } v \in H^1_0(\Omega). \tag{117}$$

More generally, given $f \in H^1_0(\Omega)^*$, we want to find $u \in H^1_0(\Omega)$ such that

$$u - \Delta u = f \quad \text{in } H^1_0(\Omega)^*. \tag{118}$$

Here, as already done before, the operator $-\Delta : H^1_0(\Omega) \to H^1_0(\Omega)^*$ is defined by, for u, v in $H^1_0(\Omega)$:

$$-\langle \Delta u, v\rangle_{H^1_0(\Omega)} := \int_\Omega \nabla u(x)\cdot\nabla v(x)\mathrm{d}x. \tag{119}$$

So, the variational formulation of (118) is: find $u \in H_0^1(\Omega)$ such that

$$\int_\Omega (u(x)v(x) + \nabla u(x)\cdot\nabla v(x))\mathrm{d}x = \langle f, u\rangle_{H_0^1(\Omega)}, \quad \text{for all } v \in H_0^1(\Omega). \tag{120}$$

Remark 5.9. (i) By the Lax-Milgram the variational formulation has a unique solution.
(ii) Note that we have no hypothesis on the domain Ω.

5.7.2. *Neumann conditions.* For $f \in H^1(\Omega)^*$, consider the problem of finding $u \in H^1(\Omega)$ such that

$$\int_\Omega (u(x)v(x) + \nabla u(x)\cdot\nabla v(x))\mathrm{d}x = \langle f, v\rangle_{H^1(\Omega)}, \quad \text{for all } v \in H^1(\Omega). \tag{121}$$

By the Lax Milgram lemma it has a unique solution. If u is smooth enough, by Green's formula we have that

$$\int_\Omega (u(x) - \Delta u(x))v(x)\mathrm{d}x + \int_{\partial\Omega} u_\nu(x)v(x)\mathrm{d}x = \langle f, v\rangle_{H^1(\Omega)}, \quad \text{for all } v \in H^1(\Omega). \tag{122}$$

Assume now that for some $f_0 \in L^2(\Omega)$ and $g \in L^2(\partial\Omega)$:

$$\langle f, v\rangle_{H^1(\Omega)} = \int_\Omega f_0(x)v(x)\mathrm{d}x + \int_{\partial\Omega} g(x)v(x)\mathrm{d}x. \tag{123}$$

We will see that function in $H^1(\Omega)$ have traces over $\partial\Omega$ in $L^2(\partial\Omega)$, so this really defines a continuous linear form over $H^1(\Omega)$. Taking $v \in \mathcal{D}(\Omega)$, we deduce from (122) that (in the weak sense for Δu):

$$u - \Delta u = f_0 \quad \text{in } L^2(\Omega). \tag{124}$$

Combining with (122) we see that this implies

$$\int_{\partial\Omega} (u_\nu(x) - g(x))v(x)\mathrm{d}x = 0, \quad \text{for all } v \in H^1(\Omega). \tag{125}$$

Since (as seen later) the trace of $H^1(\Omega)$ over $\partial\Omega$ is a dense subset of $L^2(\partial\Omega)$ this means that we have the **formal Neumann boundary condition**

$$u_\nu(x) = g(x) \quad \text{a.e. over } \partial\Omega. \tag{126}$$

5.8. Trace theorem for $H^1(\Omega)$.

5.8.1. *Main result.*

Definition 5.10. Let $\mathcal{N}$ be a neighbourhood of $\bar{x} \in \partial\Omega$. We say that $\Xi(x) : \mathcal{N} \to \mathbb{R}^n$ is a (Lipschitz) local change of coordinates whenever it is injective, and Lipschitz with Lipschitz inverse. In addition we say that Ξ is **flattening** $\partial\Omega$ near $\bar{x}$, whenever

$$\Xi(\mathcal{N} \cap \Omega) \subset \mathbb{R}^{n-1} \times \mathbb{R}_+; \quad \Xi(\mathcal{N} \cap \partial\Omega) \subset \mathbb{R}^{n-1} \times \{0\}. \tag{127}$$

Then we say that $\partial\Omega$ is Lipschitz near $\bar{x}$ and that $\mathcal{N}$ is a flattening neighbourhood. If Ξ is of class C^p, we say that $\partial\Omega$ is of class C^p near $\bar{x}$. If $\partial\Omega$ is Lipschitz, or of class C^p near each of its element, we say that it is Lipschitz, or of class C^p.

We will mainly restrict the analysis to the case when Ω is bounded, with Lipschitz boundary. It is easily seen that $\partial\Omega$ is compact, and can be covered with finitely many flattening neighbourhoods.

Let us assume that Ω is **bounded with a C^∞ boundary**, or is a half-space. Then the following properties hold, see [7]:

- $C^\infty(\bar{\Omega}) \cap H^1(\Omega)$ is a dense subset of $H^1(\Omega)$. Consider the **trace mapping** $\tau : C^\infty(\bar{\Omega}) \cap H^1(\Omega) \to C(\partial\Omega)$.

- For some $c > 0$,

$$\|\tau v\|_{L^2(\partial\Omega)} \leq c\|v\|_{H^1(\Omega)}, \quad \text{for all } v \in H^1(\Omega), \tag{128}$$

so that τ has a **unique continuous extension** (still denoted by τ) from $H^1(\Omega)$ into $L^2(\partial\Omega)$, which has dense range, and kernel equal to $H^1_0(\Omega)$.

- The range of τ, denoted by $H_\tau(\partial\Omega)$, is endowed with the **trace norm**

$$\|z\|_\tau := \min_{u\in H^1(\Omega)} \|u\|_{H^1(\Omega)}; \quad \tau u = z. \tag{129}$$

Since $H^1(\Omega)$ is a Hilbert space, we may write

$$H^1(\Omega) = H^1_0(\Omega) \oplus H^1_0(\Omega)^\perp. \tag{130}$$

The solution of problem (129) is the unique $u \in H^1_0(\Omega)^\perp$, with trace z, or $u = \tau^- z$, where τ^- denotes the (continuous) **pseudo-inverse** (inverse of minimum norm) of τ, and so,

$$\|z\|_\tau = \|\tau^- z\|_{H^1(\Omega)}. \tag{131}$$

Clearly $H_\tau(\partial\Omega)$ is a Hilbert space, isometric to $H^1_0(\Omega)^\perp$.

5.8.2. *More on the pseudo-inverse* τ^-. Given $z \in H_\tau(\partial\Omega)$, since $u = \tau^- z$ is the unique solution of $\tau u = z$ that is orthogonal to $H^1_0(\Omega)$, we have that

$$\int_\Omega (u(x)v(x) + \nabla u(x)\cdot\nabla v(x)\mathrm{d}x = 0, \;\text{ for all } v \in H^1_0(\Omega), \tag{132}$$

and therefore u is the unique solution of the problem with **nonhomogeneous Dirichlet condition**

$$u - \Delta u = 0 \text{ in } \Omega; \quad u = z \quad \text{on } \partial\Omega. \tag{133}$$

5.8.3. *Embeddings of boundary spaces.* It turns out [7] that $H_\tau(\partial\Omega)$ is isomorphic (up to local maps if Ω is not a half space) to $H^{1/2}(\partial\Omega)$, the space of closure of linear combinations of functions with compact support on flattening neighbourhoods, norm evaluated after flattening, as for $H^{1/2}(\mathbb{R}^{n-1})$, by the Fourier transform. We have, if Ω is bounded, the (strict when $n > 1$) **dense and compact embeddings**

$$H^1(\partial\Omega) \subset H^{1/2}(\partial\Omega) \subset L^2(\partial\Omega). \tag{134}$$

In particular, a bounded sequence in $H^{1/2}(\partial\Omega)$ has a strongly converging subsequence in $L^2(\partial\Omega)$.

Remember that for smooth Ω, and u, v smooth enough functions over Ω:

$$\int_{\partial\Omega} u_\nu(x)v(x)\mathrm{d}x = \int_\Omega v(x)\Delta u(x)\mathrm{d}x + \int_\Omega \nabla v(x)\cdot\nabla u(x)\mathrm{d}x. \tag{135}$$

Now take $u \in H^1_\Delta(\Omega)$, $z \in H^{1/2}(\partial\Omega)$ and $v = v_z$, $v_z := \tau^- z$. Then

$$z \mapsto \int_\Omega v_z(x)\Delta u(x)\mathrm{d}x + \int_\Omega \nabla v_z(x)\cdot\nabla u(x)\mathrm{d}x \tag{136}$$

is a continuous linear form over $H^{1/2}(\partial\Omega)$, which **we may call normal derivative** of u.

This normal derivative operator $u \mapsto u_\nu$ is a continuous linear mapping:

$$H^1_\Delta(\Omega) \to H^{1/2}(\partial\Omega)^* \tag{137}$$

defined by: for all $u \in H^1_\Delta(\Omega)$:

$$\langle u_\nu, z\rangle_{H^{1/2}(\partial\Omega)} := \int_\Omega v_z(x)\Delta u(x)\mathrm{d}x + \int_\Omega \nabla v_z(x)\cdot\nabla u(x)\mathrm{d}x. \tag{138}$$

5.9. **Normal fluxes.** Let Ω be a bounded open subset of $\mathbb{R}^n$ with C^∞ boundary. We has seen in Corollary 5.1 that, if $\Phi(x)$ is a C^1 vector field over $\bar{\Omega}$, and $v(x) : \bar{\Omega} \to \mathbb{R}$ is C^1, then

$$\int_\Omega v \operatorname{div} \Phi \mathrm{d}x + \int_\Omega \Phi \cdot \nabla v = \int_{\partial\Omega} v \Phi_\nu \mathrm{d}x, \tag{139}$$

with $\Phi_\nu := \Phi \cdot \nu$ called **normal flux**. Consider the space

$$L^2_{\text{div}}(\Omega)^n := \{\Phi \in L^2(\Omega)^n;\ \ \operatorname{div} \Phi \in L^2(\Omega)\}. \tag{140}$$

Clearly the l.h.s. of (139) extends to a bilinear continuous form say $a(v, \Phi)$ over $H^1(\Omega) \times L^2_{\text{div}}(\Omega)^n$. Adapting the arguments of the previous section, we see that when $\Phi \in L^2_{\text{div}}(\Omega)^n$, this allows to define the **normal flux** Φ_ν as an element of the dual of $H^{1/2}(\Omega)$ (since $H^{1/2}(\Omega)$ is the trace space of $H^1(\Omega)$) and the following holds:

Lemma 5.11. *The normal flux mapping: $\Phi \mapsto \Phi_\nu$ has a continuous extension to $L^2_{\text{div}}(\Omega)^n$, with image in the dual of $H^{1/2}(\partial\Omega)$. If $w \in H^{1/2}(\partial\Omega)$ is the trace of $v \in H^1(\Omega)$, then*

$$\langle \Phi_\nu, w \rangle_{H^{1/2}(\partial\Omega)} = \int_\Omega v \operatorname{div} \Phi + \int_\Omega \Phi \cdot \nabla v. \tag{141}$$

5.10. **General trace theorems.** The following holds, see Grisvard [4, Thm. 1.5.1.2]. Denote by τ the trace and by $\frac{\partial^\ell u}{\partial \nu^\ell}$ the trace of the ℓ times derivative along the normal direction at the boundary of a domain Ω.

Theorem 5.12. *Let Ω have a $C^{k,1}$ boundary (k times differentiable with Lipschitz derivatives of order k). Let $s > 0$ and $p \in (1, \infty)$ be such that $s - 1/p$ is nonnegative and not integer, i.e., $s - 1/p = \ell + \sigma$, with $\ell \in \mathbb{N}$, $\sigma \in (0, 1)$. Then the mapping*

$$u \mapsto \left(\tau u, \frac{\partial u}{\partial \nu}, \ldots, \frac{\partial^\ell u}{\partial \nu^\ell} \right) \tag{142}$$

defined over $C^{k,1}(\bar{\Omega})$ (space of C^k functions over $\bar{\Omega}$ with Lipschitz derivatives up to order k) has a continuous extension as an operator from $W^{s,p}(\Omega)$ onto

$$E^{s,\ell,p} := \Pi_{j=0}^{\ell} W^{s-j-1/p,p}(\partial\Omega). \tag{143}$$

In addition this operator has a right continuous inverse $E^{s,\ell,p} \to W^{s,p}(\Omega)$.

Remark 5.13. In the case when $s - 1/p \in \mathbb{N}$, traces of Sobolev spaces involve Besov spaces, see Lions and Peetre [8]. In particular, in general the trace of $W^{1/p,p}(\Omega)$ (and in particular the trace of $H^{1/2}(\Omega)$) is not equal to $L^p(\partial\Omega)$.

Example 5.14. For $p \in (1, \infty)$, the space $W^{1,p}(\Omega)$ has a trace on $\partial\Omega$ in the space $W^{1-1/p,p}(\partial\Omega)$, and the trace mapping is surjective.

Example 5.15. For $p \in (1, \infty)$, the space $W^{2,p}(\Omega)$ has a trace on $\partial\Omega$ in the space $W^{2-1/p,p}(\partial\Omega)$, and a normal derivative in the space $W^{1-1/p,p}(\partial\Omega)$. In addition, the mapping $u \mapsto (\tau u, \frac{\partial u}{\partial \nu})$ is onto from $W^{2,p}(\Omega)$ to $W^{2-1/p,p}(\partial\Omega) \times W^{1-1/p,p}(\partial\Omega)$.

5.11. **Elliptic optimal control with various boundary conditions.**

5.12. **Periodicity conditions.**

5.12.1. *Framework, state equation.* Domain: **standard torus**

$$\Omega := (0,1)^n. \tag{144}$$

Functions over Ω identified with the functions over $\mathbb{R}^n$, invariant by translation of any element of the natural basis. The torus has $2n$ **facets** (faces of maximal dimension):

$$F_i := \{x \in \bar{\Omega};\ x_i = 0\}; \quad F_i' := \{x \in \bar{\Omega};\ x_i = 1\}, \quad 1 \leq i \leq n. \tag{145}$$

If $y \in H^1(\Omega)$, denote by $y_{|F_i}$ the trace of y over F_i, with a similar convention for F_i'. We will establish later that this trace mapping is well-defined and continuous $H^1(\Omega) \to L^2(\partial\Omega)$. The space of periodic functions over Ω, with square integrable gradient, is:

$$Y := \{y \in H^1(\Omega);\ y_{|F_i} = y_{|F_i'}, 1 \leq i \leq n\}. \tag{146}$$

Consider the continuous bilinear form on $Y \times Y$:

$$a(y,z) := \int_\Omega (y(x)z(x) + \nabla y(x) \cdot \nabla z(x))\mathrm{d}x. \tag{147}$$

Let $U := L^2(\Omega)$. By the Lax-Milgram theory for any $u \in U$, the state equation

$$a(y,z) = \int_\Omega u(x)z(x)\mathrm{d}x, \quad \text{for all } z \in Y, \tag{148}$$

has a unique solution $y[u] \in Y$, and for some $c > 0$ not depending on u, $\|y[u]\|_Y \leq c\|u\|_2$.Taking first $z \in \mathcal{D}(\Omega)$ and integrating by parts, we obtain that in a weak sense

$$y - \Delta y = u \quad \text{over } \Omega. \tag{149}$$

Taking now z arbitrary in Y, integrating by parts in (148), and taking into account the periodicity condition for z, we deduce that

(150)
$$0 = \sum_{i=1}^n \left(\int_{F_i'} \frac{\partial y(x)}{\partial x_i} z(x) - \int_{F_i} \frac{\partial y(x)}{\partial x_i} z(x) \right) = \sum_{i=1}^n \int_{F_i} \left(\frac{\partial y(x+e_i)}{\partial x_i} - \frac{\partial y(x)}{\partial x_i} \right) z(x)\mathrm{d}x.$$

Given φ of class $C^\infty : F_i \to \mathbb{R}$ with support in the interior of F_i, it is not difficult to construct $z \in Y$ with trace φ over F_I, trace $\varphi(\cdot - e_i)$ over F_I', and zero trace on other facets. Therefore (150) implies that

$$\int_{F_i} \left(\frac{\partial y(x+e_i)}{\partial x_i} - \frac{\partial y(x)}{\partial x_i} \right) \varphi(x)\mathrm{d}x = 0. \tag{151}$$

By a density argument we deduce that

$$\frac{\partial y(x)}{\partial x_i} = \frac{\partial y(x+e_i)}{\partial x_i} \quad \text{for } x \in F_i,\ i = 1 \text{ to } n. \tag{152}$$

This means that the solution y extended by periodicity over $\mathbb{R}^n$ has continuous normal derivatives at the boundary of the cells.

Lemma 5.16. *The solution $y[u]$ of the state equation* (148) *belongs to $H^2(\Omega)$, and $u \mapsto y[u]$ is linear and continuous from U into $H^2(\Omega)$.*

Remark 5.17. Let $u \in L^p(\Omega)$, with $2 < p < \infty$. Using the setting of the previous proof, we can show that $y \in W^{2,p}(\Omega)$ by a **bootstrapping argument**. Indeed, let $y \in W^{2,q}(\Omega)$, for some $2 \leq q < \infty$. In view of the previous lemma, this holds for $q = 2$. By the Sobolev imbeddings, y and ∇y belongs to $L^{q'}(\Omega)$, where q' is such that $1/q' = 1/q - 1/n$ if $q < n$, and $q' := q + 1$ otherwise. Therefore $g \in L^{r(q)}(\Omega)$, where $r(q) := \min(p, q')$. Then $y \in W^{2,r(p)}(\Omega)$. As long as $r(p) < p$, we have that $1/r(p) - 1/q = 1/n$. So the sequence $q_0 = 2$, $q_{k+1} := r(q_k)$, for $k \in \mathbb{N}$, is after finitely many steps constant, equal to p. At each step we check that $u \mapsto y$ is continuous $L^p(\Omega) \to W^{2,q_k}(\Omega)$. The conclusion follows.

5.12.2. *Optimal control with periodicity conditions.* Cost function with $y_d \in L^2(\Omega)$, $\gamma \in L^\infty_+(\partial\Omega)$:

$$J(y,u) := \tfrac{1}{2}\int_\Omega u(x)^2 + \tfrac{1}{2}\int_\Omega (y-y_d)^2 + \tfrac{1}{2}\int_{\partial\Omega} \gamma y^2. \tag{153}$$

Costate equation

$$\int_\Omega (z(x)p(x) + \nabla z(x)\cdot\nabla p(x))\mathrm{d}x = \int_\Omega (y(x)-y_d(x))z(x)\mathrm{d}x + \int_{\partial\Omega} \gamma(x)y(x)z(x)\mathrm{d}x. \tag{154}$$

Taking z in $\mathcal{D}(\Omega))$ we deduce that

$$p - \Delta p = y - y_d \quad \text{in } \Omega. \tag{155}$$

Integrating by parts in (154) we deduce that

$$\sum_{i=1}^n \int_{F_i} \left(\frac{\partial p(x+e_i)}{\partial x_i} - \frac{\partial p(x)}{\partial x_i} \right) z(x)\mathrm{d}x = \int_{\partial\Omega} \gamma(x)y(x)z(x)\mathrm{d}x. \tag{156}$$

Since y and z are periodic (and by the usual density argument for z) this is equivalent to

$$\frac{\partial p(x+e_i)}{\partial x_i} - \frac{\partial p(x)}{\partial x_i} = (\gamma(x)+\gamma(x+e_i))y(x), \quad x \in F_i,\ \ i = 1,\dots,n. \tag{157}$$

The derivative of the reduced cost $J_R(u) := J(u, y[u])$ is therefore

$$DJ_R(u) = u + p. \tag{158}$$

Remark 5.18. Since the state is periodic, we obtain the same cost function by changing $\gamma(x)$ into $\tilde\gamma(x)$ defined by

$$\tilde\gamma(x) = \tilde\gamma(x+e_i) = \tfrac{1}{2}(\gamma(x)+\gamma(x+e_i)), \quad \text{for all } x \in F_i. \tag{159}$$

This leaves the costate invariant.

5.13. **Robin boundary conditions.** Here the domain Ω is an open bounded subset of $\mathbb{R}^n$ with C^∞ boundary. Consider the state equation

$$y - \Delta y = f \ \text{ in } \Omega; \quad \beta y + y_\nu = u \ \text{ on } \partial\Omega, \tag{160}$$

with $u \in L^2(\partial\Omega)$, $\beta \in L^\infty_+(\partial\Omega)$. The variational formulation is

$$\int_\Omega (yz + \nabla y\cdot\nabla z) + \int_{\partial\Omega} \beta yz = \int_\Omega fz + \int_{\partial\Omega} uz, \ \text{for all } z \in H^1(\Omega). \tag{161}$$

Cost function with $y_d \in L^2(\Omega)$, $\gamma \in L^\infty_+(\partial\Omega)$:

$$J(y,u) := \tfrac{1}{2}\int_\Omega (y-y_d)^2 + \tfrac{1}{2}\int_{\partial\Omega} (u^2 + \gamma y^2). \tag{162}$$

The costate equation, in the abstract Lax-Milgram setting, reads

$$a(z,p) = D_y J(u,y)z, \quad \text{for all } z \in Y. \tag{163}$$

Here the corresponding expression is

$$\int_\Omega (pz + \nabla p\cdot\nabla z) + \int_{\partial\Omega} \beta pz = \int_\Omega (y-y_d)z + \int_{\partial\Omega} \gamma yz, \ \text{for all } z \in H^1(\Omega). \tag{164}$$

The corresponding PDE is

$$p - \Delta p = y - y_d \ \text{ in } \Omega; \quad \beta p + p_\nu = \gamma y \ \text{ on } \partial\Omega. \tag{165}$$

Observe that **the boundary condition for the costate is also of Robin type** (and of Neuman type if $\beta = 0$). The derivative of reduced cost $J_R(u := J(u, y[u])$ is:

$$DJ_R(u) = u + p \quad \text{in } L^2(\partial\Omega). \tag{166}$$

5.14. **Static Fokker-Planck equation.** In this section we study a static version of the isotropic Fokker-Planck equation, starting with a general continuity equation. Then we discuss the case of isotropic diffusion, and finally the general Fokker-Planck equation.

5.14.1. *Well-posedness of the state equation.* Consider a **static** form of the continuity equation:

$$\begin{cases} \text{(i)} & \frac{1}{h} y(x) + \operatorname{div} \Phi[y](x) = f(x) + u_d(x) \text{ in } \Omega; \\ \text{(ii)} & -\Phi_\nu[y] = u_b \quad \text{on } \partial\Omega. \end{cases} \tag{167}$$

Here $h > 0$ can as before be interpreted as a time step so that (167) can be viewed at a time discretization of a dynamic equation. The flux $\Phi[y]$ is a function of x (with values in $\mathbb{R}^n$) parameterized by the solution y; $f \in L^2(\Omega)$ is a given r.h.s. and $u_d \in L^2(\Omega)$ is the **distributed control**. In addition $u_b \in L^2(\partial\Omega)$ is the **boundary control**. The control is $u = (u_b, u_d)$, element of $U := L^2(\Omega) \times L^2(\partial\Omega)$.

The Stokes type formula established in lemma 5.11 implies that, if $\Phi[y]$ belongs to $L^2_{\text{div}}(\Omega)^n$, then

$$\int_\Omega z \operatorname{div} \Phi[y] \mathrm{d}x + \int_\Omega \Phi[y] \cdot \nabla z = \int_{\partial\Omega} z \Phi_\nu[y] \mathrm{d}x, \quad \text{for all } z \in H^1(\Omega). \tag{168}$$

The variational formulation of (167) is therefore, choosing $Y = H^1(\Omega)$ as state space:

$$\int_\Omega (\frac{1}{h} yz - \Phi[y] \cdot \nabla z) = \int_\Omega (f + u_d) z + \int_{\partial\Omega} z u_b, \quad \text{for all } z \in Y. \tag{169}$$

In the sequel, we assume that $\Phi[y](x)$ is a linear function of $y(x)$ and $\nabla y(x)$, with coefficients that are measurable in x. We apply the Lax-Milgram framework with

$$a(y, z) := \int_\Omega (\frac{1}{h} yz - \Phi[y] \cdot \nabla z); \quad Lz := \int_\Omega (f + u_d) z + \int_{\partial\Omega} z u_b. \tag{170}$$

Obviously Lz is a continuous linear form over Y. We need the continuity and coercivity of the bilinear form in the left side of (169). Sufficient conditions for this are that, for some $c > 0$ and $\varepsilon > 0$:

$$\begin{cases} \text{(i)} & |\Phi[y](x)| & \leq c(|y(x)| + |\nabla y(x)|), \\ \text{(ii)} & \varepsilon\left(|y(x)^2 + |\nabla y(x)|^2\right) & \leq \frac{1}{h} y(x)^2 - \Phi[y](x) \cdot \nabla y(x). \end{cases} \tag{171}$$

We have proved that:

Lemma 5.19. *Let* (171) *hold. Then the variational formulation* (169) *has a unique solution* $y[u]$ *in* Y*.*

5.14.2. *Optimality conditions.* Consider a cost function of the following form:

$$J(y, u) := \tfrac{1}{2} \int_\Omega (u_d^2 + (y - y_d)^2) + \tfrac{1}{2} \int_{\partial\Omega} (u_b^2 + \gamma y^2), \tag{172}$$

where $y_d \in L^2(\Omega)$, $\gamma \in L^\infty_+(\partial\Omega)$ are given. The reduction Lagrangian is

$$\mathcal{L}(u, y, p) = J + Lp - a(y, p), \tag{173}$$

with $u \in U$, $y \in Y$, $p \in Y$. The variational formulation of the costate equation is

$$\int_\Omega (\frac{1}{h} zp - \Phi[z] \cdot \nabla p) = \int_\Omega (y - y_d) z + \int_{\partial\Omega} \gamma yz, \quad \text{for all } z \in Y. \tag{174}$$

Denote by $J_R(u_b, u_d) := J(u_b, u_d, y[u_b, u_d])$ the reduced cost. By the above discussion, its derivative in direction $v = (v_b, v_d) \in U$, is:

$$\begin{cases} D_{u_d} J_R(u_b, u_d) v_d &= \mathcal{L}_{u_d} v_d = \int_\Omega (u_d + p) v_d, \\ D_{u_d} J_R(u_b, u_d) v_b &= \mathcal{L}_{u_b} v_b = \int_{\partial\Omega} (u_b + p) v_b. \end{cases} \tag{175}$$

5.14.3. *Isotropic Fokker-Planck equation.* This is the special case of the continuity equation, when the flux $\Phi[y]$ is the sum of transportation and isotropic diffusion terms:

$$\Phi[y](x) := y(x) b(x) - \nabla y(x), \tag{176}$$

for some vector field $b \in L^\infty(\Omega)^n$, and then

$$a(y, z) := \int_\Omega \left(\frac{1}{h} yz - \Phi[y] \cdot \nabla z \right) = \int_\Omega \left(\frac{1}{h} yz - yb \cdot \nabla z + \nabla y \cdot \nabla z \right). \tag{177}$$

The r.h.s. $L \in Y^*$ is as in (170).

Lemma 5.20. *Let $b \in L^\infty(\Omega)^n$. If h is small enough, then* (171) *holds, and consequently, the variational formulation* (169) *has a unique solution in $H^1(\Omega)$.*

We consider again the cost function in (172). The costate equation reads

$$\int_\Omega (\frac{1}{h} pz - (zb - \nabla z) \cdot \nabla p) = \int_\Omega (y - y_d) z + \int_{\partial\Omega} \gamma yz. \tag{178}$$

It has a unique solution in $H^1(\Omega)$, and since

$$\int_\Omega \nabla z \cdot \nabla p = \int_{\partial\Omega} z p_\nu - \int_\Omega z \Delta p, \tag{179}$$

we see that formally, the costate is solution of the PDE

$$\begin{cases} \frac{1}{h} p - b \cdot \nabla p - \Delta p &= y - y_d \text{ in } \Omega; \\ p_\nu &= \gamma y \text{ on } \partial\Omega. \end{cases} \tag{180}$$

Note that **the boundary condition for the adjoint is of Neumann type**.

5.15. **Case of an anisotropic operator.** We next consider a similar problem with a general diffusion term, namely

$$\frac{1}{h} y + \operatorname{div}(yb) - \sum_{i,j=1}^n \frac{\partial^2 [a_{ij}(x) y(x)]}{\partial x_i \partial x_j} = f + u_d \text{ in } \Omega, \tag{181}$$

with $\{a_{ij}(x)\}$ nonnegative diffusion matrix (not necessarily symmetric). This is a special case of the continuity equation (167)(i) where, for $i = 1$ to n:

$$\Phi_i[y](x) := y(x) b_i(x) - \sum_{j=1}^n \frac{\partial [a_{ij}(x) y(x)]}{\partial x_j}. \tag{182}$$

We still assume a boundary condition on the normal flux:

$$-\Phi_\nu[y] = u_b \quad \text{on } \partial\Omega. \tag{183}$$

So, the variational formulation (169) remains valid with the new expression of the flux. The corresponding bilinear form is now

$$a(y, z) = \int_\Omega \left(\frac{1}{h} yz - yb \cdot \nabla z + \sum_{i,j=1}^n \frac{\partial z(x)}{\partial x_i} \frac{\partial [a_{ij}(x) y(x)]}{\partial x_j} \right), \tag{184}$$

and the expression for the r.h.s. of the state equation is still valid. Assuming the cost function to be still (172), the expression (178) for the costate equation is still valid. That is,

$$\left\{\begin{array}{ll} a(z,p) & = \int_\Omega \left(\frac{1}{h} zp - zb\cdot\nabla p + \sum_{i,j=1}^n \frac{\partial p(x)}{\partial x_i} \frac{\partial [a_{ij}(x) z(x))]}{\partial x_j} \right) \\ & = \int_\Omega (y(x) - y_d(x)) z(x) + \int_{\partial\Omega} \gamma y z. \end{array}\right. \tag{185}$$

We need to integrate by parts w.r.t. x_j the term

$$\int_\Omega \frac{\partial p(x)}{\partial x_i} \frac{\partial [a_{ij}(x) z(x)]}{\partial x_j} = -\int_\Omega z(x) a_{ij}(x) \frac{\partial^2 p(x)}{\partial x_i \partial x_j} + \int_{\partial\Omega} z(x) a_{ij}(x) \frac{\partial p(x)}{\partial x_i} \nu_j. \tag{186}$$

We see that the corresponding PDE is

$$\frac{1}{h} p - b\cdot\nabla p - \sum_{i,j=1}^n a_{ij}(x) \frac{\partial^2 p(x)}{\partial x_i \partial x_j} = y - y_d \text{ in } \Omega; \quad \frac{\partial p}{\partial \nu_{A^\dagger}} = \gamma y \text{ on } \partial\Omega, \tag{187}$$

where we recall, see (92), that

$$p_{\nu_{A^\dagger}}(x) = \sum_{i,j=1}^n \nu_j(x) a_{ij}(x) \frac{\partial p(x)}{\partial x_i}. \tag{188}$$

We skip the material about Bochner's integrals with value in a Banach space.

6. Control of parabolic equations

6.1. **Setting.**

6.1.1. *Gelfand triple.* We consider the Lions-Magenes variational framework for parabolic equations [6]. The **Gelfand triple setting** is as follows. Let V, H be Hilbert spaces and $J \in L(V,H)$ be injective. Then we may identify $v \in V$ with Jv and call J the operator of inclusion from V info H. By identifying we mean that in practice, we sometimes may omit the operator J and write v instead of Jv, for $v \in V$.

The adjoint operator $J^* : H^* \to V^*$ associates with each $h^* \in H^*$ the element $J^* h^* \in V^*$ defined by

$$\langle J^* h^*, v\rangle_V = \langle h^*, Jv\rangle_H. \tag{189}$$

It can be interpreted as the restriction of the linear form h^* to elements of V. We may say that **the transpose of an inclusion is a restriction**. We say that (V,H,V^*) is a **Gelfand triple** if the operator J of inclusion of V in H (necessarily injective) has **dense image**.

Lemma 6.1. *Let (V,H,V^*) be a Gelfand triple. Then J^* is injective with dense image.*

By the Riesz theorem, there exists a bijective isometry $\mathbf{\Lambda}_H : H \to H^*$ such that

$$(h,h')_H = \langle \mathbf{\Lambda}_H h, h'\rangle_H, \quad \text{for all } h,\, h' \in H. \tag{190}$$

So, we have the dense and continuous injections, identified with corresponding inclusions, called **Gelfand triple**:

$$V \overset{J}{\subset} H \overset{\mathbf{\Lambda}_H}{\equiv} H^* \overset{J^*}{\subset} V^*. \tag{191}$$

We call H the **pivot space**. In our applications V is separable, i.e., it contains a dense sequence v_k which is then also dense in H and v^*. We speak then of a **separable Gelfand triple**.

Remark 6.2. In the Gelfand triple setting, the essential rule is that (understating in practice the injection operators), if $v^* \in V^*$ can be expressed as $v^* = J^* h^*$ for some (necessarily unique) $h^* \in H^*$, then

$$\langle v^*, v\rangle_V = \langle h^*, Jv\rangle_H. \tag{192}$$

In particular, let v, v' belong to V. One should not confuse $\langle v, v'\rangle_V = (v, v')_H$ with $(v, v')_V$.

Remark 6.3. Let $M : V \to V^*$, $Mv := J^* \mathbf{\Lambda}_H Jv$. Then, for all v, w in V:

$$\langle Mv, w\rangle_V = \langle \mathbf{\Lambda}_H Jv, Jw\rangle_H = (Jv, Jw)_H. \tag{193}$$

Example 6.4. Given a smooth open subset Ω of $\mathbb{R}^n$, a typical example is when $V = H^1(\Omega)$ and $H = L^2(\Omega)$. Understating the operators J, J^*, $\mathbf{\Lambda}_H$, we obtain that, for v, v' in V:

$$\langle v, v'\rangle_V = (v, v')_H = \int_\Omega v(x)v'(x)\mathrm{d}x \tag{194}$$

is different from

$$(v, v')_V = \int_\Omega (v(x)v'(x) + \nabla v(x) \cdot \nabla v'(x))\mathrm{d}x. \tag{195}$$

6.1.2. *The $W(0,T)$ space.* Now define

$$W(0,T) := \{y \in L^2(0,T;V);\ \ \dot{y} \in L^2(0,T;V^*)\}. \tag{196}$$

It can be proved, see [7], that the following **continuous inclusion** holds:

$$W(0,T) \subset C(0,T;H) \tag{197}$$

as well as the **integration by parts formula**: for all $0 \le t' \le t'' \le T$:

$$(y(t''), z(t''))_H - (y(t'), z(t'))_H = \int_{t'}^{t''} (\langle \dot{y}(t), z(t)\rangle_V + \langle \dot{z}(t), y(t)\rangle_V)\,\mathrm{d}t \tag{198}$$

holds for all y, z in $W(0,T)$. Since the r.h.s. is the primitive of an integrable function, (198) is equivalent to

$$(y(t), z(t))_H \text{ is primitive of } \langle \dot{y}(t), z(t)\rangle_V + \langle \dot{z}(t), y(t)\rangle_V. \tag{199}$$

Being the primitive of an integrable function, $(y(t), z(t))_H$ is absolutely continuous, belongs to $W^{1,1}(0,T)$, and for a.a. t, $(y(t), z(t))_H$ is Fréchet differentiable and the Fréchet and weak derivative coincide:

$$\frac{\mathrm{d}}{\mathrm{d}t}(y(t), z(t))_H = \langle \dot{y}(t), z(t)\rangle_V + \langle \dot{z}(t), y(t)\rangle_V. \tag{200}$$

For more reading about the primitives of integrable functions, see e.g. Malliavin [9, Ch. 5, sect. 6].

6.1.3. *The parabolic equation.* In the sequel we assume that (V, H, V^*) is a separable Gelfand triple. Let the bilinear form $a(t; y, z)$ over $V \times V$ be uniformly (in time) continuous, i.e., for some $M > 0$, we have for a.a. t:

$$|a(t; y, z) \le M\|y\|_V \|z\|_V \quad \text{for all } y,\ z \text{ in } V, \tag{201}$$

and uniformly **semicoercive**, i.e. for some $\alpha > 0$ and $\lambda \in \mathbb{R}$, we have for a.a. t:

$$a(t; y, y) \ge \alpha\|y\|_V^2 - \lambda\|y\|_H^2, \quad \text{for all } y \in V. \tag{202}$$

Let $f \in L^2(0,T;V^*)$ and $y_0 \in H$. Consider the abstract parabolic equation, with unknown $y \in W(0,T)$, so that the initial condition makes sense:

(203)
$$\begin{cases} \langle \dot{y}(t), z\rangle_V + a(t; y(t), z) & = \langle f(t), z\rangle_V, \ \text{ for all } z \in V, \text{ for a.a. } t \in (0,T); \\ y(0) & = y_0. \end{cases}$$

Remark 6.5. With $a(t;y,z)$, is associated $A(t) \in L(V,V^*)$ such that

$$a(t;y,z) = \langle A(t)y,z\rangle_V \quad \text{for all } y,\ z \text{ in } V. \tag{204}$$

So we may as well write the parabolic equation (203) in the **operational form**

$$\dot y(t) + A(t)y(t) = f(t) \ \text{ in } L^2(0,T;V^*);\ \ y(0) = y_0 \ \text{ in } H. \tag{205}$$

The uniform continuity of $a(t;\cdot)$ is equivalent to the one of $A(t)$, and its semicoercivity with parameters (α,λ) is equivalent to

$$\langle A(t)y,y\rangle_V \geq \alpha\|y\|_V^2 - \lambda\|y\|_H^2, \quad \text{for all } y \in V. \tag{206}$$

Theorem 6.6. *Let $y_0 \in H$ and $f \in L^2(0,T;V^*)$. Then (203) has a unique solution in $W(0,T)$, and for some $c > 0$ not depending on (y_0, f):*

$$\|y\|_{W(0,T)} \leq c\left(\|y_0\|_H + \|f\|_{L^2(0,T;V^*)}\right). \tag{207}$$

Lemma 6.7. *We have that $\bar y$ is solution of the parabolic equation (203).*

6.1.4. *The second a priori estimate.* We have a stronger result under the **semi-symmetry hypothesis**

$$\begin{cases} a(t;y,z) = a_0(t;y,z) + a_1(t;y,z); \\ a_0\ C^1 \text{ in time, symmetric, unif. coercive,} \\ a_1 \text{ unif. continuous on } V \times H. \end{cases} \tag{208}$$

The two last conditions imply that there exists $c_0 > 0$ such that

$$|a_1(t;y,)| \geq c_0\|y\|_V^2, \quad \text{for all } y,\ z \text{ in } V. \tag{209}$$

and that there exists $c_1 > 0$ such that

$$|a_1(t;y,z)| \leq c_1\|y\|_V\|z\|_H, \quad \text{for all } y,\ z \text{ in } V. \tag{210}$$

We denote by $A_0(t)$ and $A_1(t)$ the operators in $L(V,V^*)$ associated with the bilinear forms a_0 and a_1.

Theorem 6.8. *Assume that $y_0 \in V$, $f \in L^2(0,T;H)$, and the semi symmetry hypothesis (208) holds. Then the solution of the parabolic equation (203) satisfies $\dot y \in L^2(0,T;H)$ and $y \in L^\infty(0,T;V)$, and for some $c > 0$ not depending on (y_0, f):*

$$\|y\|_{L^\infty(0,T;V)} + \|\dot y\|_{L^2(0,T;H)} \leq c\left(\|y_0\|_V + \|f\|_{L^2(0,T;H)}\right). \tag{211}$$

6.2. Optimal control of parabolic equations.

6.2.1. *Linear-quadratic setting.* Consider the state equation

$$\begin{cases} \text{For a.a. } t \in (0,T): \\ \langle \dot y(t), z\rangle_V + a(t;y(t),z) = \langle f(t) + Bu(t), z\rangle_V, \ \text{ for all } z \in V; \\ y(0) = y_0. \end{cases} \tag{212}$$

Here (V,H,V^*) is a separable Gelfand triple, $f \in L^2(0,T;V^*)$, $y_0 \in H$, U is a Hilbert space, $B \in L(U,V^*)$, the bilinear form $a(t,\cdot,\cdot)$ is uniformly (in time) continuous and semi coercive. We know that an equivalent expression is

$$\dot y(t) + A(t)y(t) = f(t) + Bu(t) \ \text{ in } L^2(0,T;V^*),\ y(0) = y_0. \tag{213}$$

where $A(t) \in L^\infty(V,V^*)$ is the operator associated with the bilinear form $a(t,\cdot,\cdot)$. Consider the cost function

$$J(u,y) := \tfrac{1}{2}\int_0^T \left(\|y(t) - y_d(t)\|_H^2 + \|u(t)\|_U^2\right) \mathrm{d}t + \tfrac{1}{2}\|y(T) - y_D\|_H^2, \tag{214}$$

where

$$y_d \in L^2(0,T;H); \quad y_D \in H. \tag{215}$$

We have the control constraint

$$u \in K_U; \quad \text{with } K_U \text{ convex, nonempty subset of } L^2(0,T;U). \tag{216}$$

We consider the optimal control problem

$$\operatorname{Min} J(u,y) \ \text{ s.t. (212) and (216).} \tag{217}$$

Remark 6.9. Since this problem is feasible and convex with a strongly convex cost, it has a unique solution.

We denote the Lagrange multiplier associated with the state equation as

$$(p,q) \in L^2(0,T;V) \times H. \tag{218}$$

The reduction Lagrangian function of this problem is
(219)

$$\mathcal{L}(u,y,p,q) := J(u,y) + \int_0^T \langle f(t) + Bu(t) - A(t)y(t) - \dot{y}(t), p(t)\rangle_V \mathrm{d}t + (q, y_0 - y(0))_H.$$

We deduce the costate equation: for arbitrary $z \in W(0,T)$:

$$\begin{aligned} 0 = \mathcal{L}_y z = & \int_0^T [(y(t) - y_d(t), z(t))_H - \langle A(t)z(t) + \dot{z}(t), p(t)\rangle_V] \, \mathrm{d}t \\ & + (y(T) - y_D(T), z(T))_H - (q, z(0))_H. \end{aligned} \tag{220}$$

We now choose to take the costate p in $W(0,T)$. Then we have the integration by parts formula

$$-\int_0^T \langle \dot{z}(t), p(t)\rangle_V \mathrm{d}t = (p(0), z(0))_H - (p(T), z(T))_H + \int_0^T \langle \dot{p}(t), z(t)\rangle_V \mathrm{d}t. \tag{221}$$

Choosing first $z \in \mathcal{D}(0,T;V)$ we deduce from (220) that

$$-\dot{p}(t) + A^*(t)p(t) = y(t) - y_d(t) \quad \text{in } L^2(0,T;V^*). \tag{222}$$

Now, choosing z arbitrary in $W(0,T)$ in (220), and combining with (221), we obtain after cancellation of the integral terms that, for all $z \in W(0,T)$:

$$((p(0) - q), z(0))_H + (y(T) - y_D(T) - p(T), z(T))_H = 0. \tag{223}$$

As a byproduct of the theory of parabolic equations, the mapping $z \mapsto (z(0), z(T))$ is onto: $W(0,T) \to H \times H$ (exercise). It follows that

$$p(T) = y(T) - y_D(T); \quad p(0) = q. \tag{224}$$

In particular, p is the unique solution (exercise) of the **backwards parabolic equation**, that called the costate equation:

$$\begin{cases} -\dot{p}(t) + A^*(t)p(t) & = y(t) - y_d(t) \ \text{ in } L^2(0,T;V^*); \\ p(T) & = y(T) - y_D(T). \end{cases} \tag{225}$$

The derivative of the reduced cost $J_R(u) := J(u, y[u])$ is therefore, identifying U with its dual:

$$DJ_R(u)v = \mathcal{L}_u v = \int_0^T (B^*p(t) + u(t), v(t))_U \mathrm{d}t. \tag{226}$$

We deduce the following necessary and sufficient optimality condition:

Proposition 6.10. *Problem* (217) *has a unique solution* $u \in K_U$*, with associated state* y *and costate* p *resp. solution of* (212) *and* (225)*, characterized by the conditions*

$$\int_0^T (B^*p(t) + u(t), v(t) - u(t))_U \mathrm{d}t \geq 0, \quad \textit{for all } v \in K_U. \tag{227}$$

6.2.2. *Smoothness of the optimal control and state.* Note that (227) is equivalent to

$$u = P_{K_U}(-B^*p). \tag{228}$$

In the case of constraints that are local in time, more precisely if

$$K_U := \{u \in L^2(0,T;U);\ \ u(t) \in k_U, \quad \text{for a.a. } t \in (0,T), \tag{229}$$

where k_U is a nonempty, closed convex subset of U, then (227) is equivalent to

$$(B^*p(t) + u(t), v(t) - u(t))_U \geq 0, \quad \text{for all } v \in k_U, \quad \text{for a.a. } t \in (0,T), \tag{230}$$

or equivalently

$$u(t) = P_{k_U}(-B^*p(t)), \quad \text{for a.a. } t \in (0,T). \tag{231}$$

This allows to deduce some regularity of the optimal control in specific cases, such as the one below.

Lemma 6.11. *Let* (229) *hold, and assume that* $B \in L(U,H)$*. Then the optimal control* u *belongs to* $C([0,T];U)$*.*

6.2.3. *Optimization of the initial state.* Consider the case when the control constraint (216) holds, and the initial state y_0 must also be optimized, under the constraint that

$$y_0 \in K_0, \tag{232}$$

with K_0 convex, nonempty subset of H. Keeping the same cost function, we write now the state as $y[u,y^0]$ and the reduced cost as $J_R(u,y^0) := J(u,y[u,y^0])$. Again applying the reduction Lagrangian technique, we obtain in the same way the sensitivity with respect to the initial state, using that $q = p(0)$:

$$D_{y^0} J_R(u,y_0)z_0 = \mathcal{L}_{y_0} z_0 = (p(0), z_0)_H, \quad \text{for all } z_0 \in H. \tag{233}$$

Note that the costate is the same as for the original optimal control problem (217). So, if $(\bar{u}, \bar{y}_0)$ is optimal iff (227) holds, as well as the optimality condition

$$(p(0), k_0 - y(0))_H \geq 0, \quad \text{for all } k_0 \in K_0. \tag{234}$$

6.2.4. *Control by the coefficients.* Controlling by the coefficients of the state equation is useful for **identifying the parameters** of the state equation.

We continue the study of the optimal control problem (217) except that we optimize also w.r.t. the operator $A \in L^2(0,T;L(V,V^*))$, with (local in time) constraint

$$A(t) \in K_A, \quad \text{for a.a. } t, \tag{235}$$

where K_A is a bounded, closed, convex and nonempty subset of $L(V,V^*)$, so that we have the uniform continuity property

$$\|A(t)\| \leq M \quad \text{for a.a. } t. \tag{236}$$

We also assume that we have a uniform semi-coercivity property: for some $\alpha > 0$ and $\lambda \in \mathbb{R}$, the associated bilinear form $a(t,y,z) := \langle A(t)y, z\rangle_V$ satisfies for a.a. t:

$$a(t;y,y) \geq \alpha\|y\|_V^2 - \lambda\|y\|_H^2, \quad \text{for all } y \in V, \text{ whenever } A(t) \in K_A. \tag{237}$$

We choose $\mathcal{A} := L^\infty(0,T;L(V,V^*))$ as operator space. The well-posedness of the state equation and the sensitivity analysis of the mapping $(u,A) \mapsto y = y[u,A]$ follow from the application of the implicit function theorem to the state equation with the bilinear term $A(t)y(t)$. The costate equation is again (225). Denoting now the reduced cost by $J_R(u,A)$, we get that

$$D_u J_R(u,A) = B^*p + u. \tag{238}$$

The **new information** is the expression of the directional derivative of the reduced cost in direction $A' \in \mathcal{A}$, thanks to the reduction Lagrangian. This reduction

Lagrangian has the same expression as before, with explicit additional argument A. We obtain

$$D_A J_R(u,A)A' = D_A \mathcal{L} A' = -\int_0^T \langle A'(t)y(t), p(t)\rangle_V \mathrm{d}t. \tag{239}$$

Equivalently, denoting by $a'(t,\cdot,\cdot)$ the bilinear form associated with $A'(t)$:

$$D_A J_R(u,A)A' = -\int_0^T a'(t, y(t), p(t))\mathrm{d}t. \tag{240}$$

Let k_A denote the set of bilinear forms with associated operator in K_A. Assume that the control constraints are

$$u(t) \in k_U, \quad \text{for a.a. } t, \tag{241}$$

where k_U is a nonempty, closed convex subset of U. Since the constraints on the control and the coefficients are local in time, we have the necessary optimality condition for (u,a) to be optimal:

$$-(B^* p(t) + u(t), v - u(t))_U \geq 0, \text{ for all } v \in k_u, \quad \text{for a.a. } t, \tag{242}$$

$$-(a'(y(t), p(t)) - a(t, y(t), p(t)) \geq 0, \quad \text{for all } a' \in k_A, \quad \text{for a.a. } t. \tag{243}$$

6.3. **Application to the Fokker-Planck equation.** We start from the continuity formulation of the FP equation

$$\begin{cases} \dot{y}(x,t) + \operatorname{div}\phi[y](x,t) = f(x,t) \text{ in } Q, \text{ where} \\ \phi_i[y](x,t) := y(x,t)b_i(x,t) - \displaystyle\sum_{j=1}^n \frac{\partial(a_{ij}(x,t)y(x,t))}{\partial x_j}, \; 1 \leq i \leq n. \end{cases} \tag{244}$$

The boundary condition is on the normal flux ($\nu(x)$ denoting the outward normal):

$$-\phi[y](x,t)\cdot\nu(x) = g(x,t) \text{ on } \Sigma. \tag{245}$$

The data are $b \in L^\infty(Q)^n$, $f \in L^2(Q)$, $g \in L^2(\Sigma)$. Multiplying (244) by a smooth enough test function $z(x,t)$ and integrating in space, using Green's formula we see that

$$\int_\Omega (\dot{y}(x,t)z(x,t) - \phi[y](x,t)\cdot\nabla z(x,t))\,\mathrm{d}x = \int_\Omega f(x,t)z(x,t)\mathrm{d}x + \int_\Sigma g(x,t)z(x,t)\mathrm{d}x. \tag{246}$$

We apply the variational Lions-Magenes setting, with $V = H^1(\Omega)$, $H = L^2(\Omega)$. The corresponding bilinear form is, for functions y, z in V:

$$a(t;y,z) := a'(t;y,z) + a''(t;y,z), \tag{247}$$

where

$$\begin{cases} a'(t;y,z) & := -\displaystyle\int_\Omega y(x)\, b(x,t)\cdot\nabla z(x)\mathrm{d}x, \\ a''(t;y,z) & := \displaystyle\sum_{i,j=1}^n \int_\Omega \frac{\partial(a_{ij}(x,t)y(x))}{\partial x_j}\frac{\partial z(x)}{\partial x_i}\mathrm{d}x. \end{cases} \tag{248}$$

We next assume the a_{ij} to be C^1 functions over Ω, so that we can expand a'' as

$$a''(t;y,z) := \sum_{i,j=1}^n \int_\Omega a_{ij}(x,t)\frac{\partial y(x)}{\partial x_j}\frac{\partial z(x)}{\partial x_i}\mathrm{d}x + \sum_{i,j=1}^n \int_\Omega y(x)\frac{\partial a_{ij}(x,t)}{\partial x_j}\frac{\partial z(x)}{\partial x_i}\mathrm{d}x. \tag{249}$$

We need assumptions over the a_{ij}, and b of boundedness:

$$a,\ \nabla a,\ b \text{ are essentially bounded} \tag{250}$$

and uniform ellipticity: for some $\alpha_0 > 0$,

$$\sum_{i,j} a_{ij}(x,t)\xi_i\xi_j \geq \alpha_0|\xi|^2, \text{ for all } \xi \in \mathbb{R}^n, x \text{ and } t. \tag{251}$$

Then we can cast (246) in the abstract form (203), with initial condition $y(0) = y_0 \in H$, and the existence theorem 6.6 applies.

Remark 6.12. On the other hand, in general, due to $a'(t;\cdot,\cdot)$ and the boundary conditions, the semi symmetry condition (208) does not hold, so that the second parabolic estimate of theorem 6.8 does not apply.

Remark 6.13. If $\text{meas}(\Omega)$ is finite, then the constant function $\mathbf{1}(x)$, with value 1 over Ω, belongs to V. Using the integration by parts formula (198) with $y = \rho$ and $z = \mathbf{1}$, we deduce the mass conservation formula:

$$\int_\Omega (y(x,t'')-y(x,t'))\mathrm{d}x = \int_{t'}^{t''} \langle \dot{y}(\cdot,t), \mathbf{1}\rangle_V = \int_{t'}^{t''} \left(\int_\Omega f(x,t)\mathrm{d}x + \int_{\partial\Omega} g(x,t)\mathrm{d}x \right) \mathrm{d}t. \tag{252}$$

6.3.1. *Maximum principle.* We assume that Ω is the unit torus and discuss the equation

$$\dot{y} - \Delta y + \text{div}(yb) = f; \quad y(0) = y_0, \tag{253}$$

with y_0 and f nonnegative.

Lemma 6.14. *Let b, $\text{div}\, b$, f and y_0 be continuous, and for some $\varepsilon_0 > 0$, $y_0(x) \geq \varepsilon_0$ for all $x \in \Omega$. Assume that $y \in C^{2,1}(\bar{Q})$ (set of function over $\bar{Q}$, that have continuous first-order time derivative, and second-order space derivatives). Then*

$$y(x,t) \geq \varepsilon_0 \exp\big(-T\|\text{div}\, b\|_{L^\infty(Q)}\big), \quad \textit{for all } (x,t) \in Q. \tag{254}$$

Lemma 6.15. *Let $b \in L^\infty(Q)$, $y_0 \in L^2(\Omega)$, $f \in L^2(Q)$, y_0 and f nonegative. Then* (254) *holds.*

6.4. **Existence and optimality conditions.** We now state a general compactness result in the parabolic framework.

6.4.1. *The Aubin-Lions theorem.* Let X, Y be Banach spaces, and $X \subset Y$ with continuous and dense inclusion. For $p \in [1,\infty]$ and $-\infty \leq \tau < \tau' \leq +\infty$, we set

$$W^p(\tau,\tau';X,Y) := \{y \in L^p(\tau,\tau';X);\ \dot{y} \in L^p(\tau,\tau';Y)\}. \tag{255}$$

The following holds, see Aubin [1], and Lions [5, Thm 5.1, ch. 1].

Theorem 6.16 (Aubin-Lions)**.** *Assume that $-\infty < \tau < \tau' < +\infty$, X and Y are reflexive, and let X_1 be another Banach space such that $X \subset X_1 \subset Y$, the first inclusion being compact.*
Let $p \in (1,\infty)$. Then the inclusion of $W^p(\tau,\tau';X,Y)$ into $L^p(\tau,\tau';X_1)$ is compact.

6.4.2. *Existence.* The standard arguments apply, excepting for passing to the limit in the bilinear term of the state equation. We detail this point. Let $y[b,f,g]$ denote the solution of the Fokker-Planck equation (244)-(245).

Lemma 6.17. *Let b_k *weakly converge to $\bar{b}$ in $L^\infty(Q)^n$, and (f_k,g_k) weakly converge to $(\bar{f},\bar{g})$ in $L^2(Q)^n \times L^2(\Sigma)$. Set $y_k := y[b_k,f_k,g_k]$, $\bar{y} := y[\bar{b},\bar{f},\bar{g}]$. Then* (i) *$y_k b_k$ weakly converges in $L^2(Q)^n$ to $\bar{y}\bar{b}$, and* (ii) *y_k weakly converges in $W(0,T)$ to $\bar{y}$.*

6.4.3. *Optimality conditions.* Set $u = (b, f, g)$, element of the Banach space

$$U := L^\infty(Q)^n \times L^2(Q) \times L^2(\Sigma). \tag{256}$$

Consider a cost function of the form $J(u, y) = J_1(u) + J_2(y)$, with

$$J_1(u) := \tfrac{1}{2} \int_Q \left(|b(x,t)|^2 + f(x,t)^2 \right) \mathrm{d}x\mathrm{d}t + \tfrac{1}{2} \int_\Sigma g(x,t)^2 \mathrm{d}x\mathrm{d}t, \tag{257}$$

$$J_2(y) := \tfrac{1}{2} \int_0^T \|y(t) - y_d(t)\|_H^2 \mathrm{d}t + \tfrac{1}{2} \|y(T) - y_D\|_H^2 \mathrm{d}t, \tag{258}$$

where

$$y_d \in L^2(Q); \quad y_D \in L^2(\Omega). \tag{259}$$

The variational formulation for the costate equation (222) is

$$\begin{cases} -\langle \dot{p}, z \rangle_V + a(t; z, p) & = (y(t) - y_d(t), z(t))_H, \text{ for all } z \in V, \text{ for a.a. } t \in (0,T); \\ p(T) & = y(T) - y_D(T). \end{cases} \tag{260}$$

The corresponding PDE has (compare to the elliptic case, see (187)), a boundary condition of Neumann type:

$$\begin{cases} -\dot{p}(x,t) - b(x,t) \cdot \nabla p(x,t) - \displaystyle\sum_{i,j=1}^n a_{ij}(x,t) \frac{\partial^2 p(x,t)}{\partial x_i \partial x_j} = y - y_d \text{ in } Q; \\ \dfrac{\partial p}{\partial \nu_{A^\top}} = \gamma y \text{ on } \Sigma. \end{cases} \tag{261}$$

The derivative of the reduced cost function is, denoting by $v = (v_b, v_f, v_g) \in U$ the direction of variation:

$$DJ_R(u)v = \int_Q (y \nabla p + b) \cdot v_b + \int_Q (f + p) v_f + \int_\Sigma (g + p) v_g. \tag{262}$$

6.5. **Complements.** Analysis of Dirichlet boundary control problems.

References

[1] J.-P. Aubin. Un théorème de compacité. *C. R. Acad. Sci. Paris*, 256:5042–5044, 1963.

[2] H. Brézis. *Functional analysis, Sobolev spaces and partial differential equations.* Springer, New York, 2011.

[3] J. Duoandikoetxea. *Fourier analysis*, volume 29 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2001. Translated and revised from the 1995 Spanish original by D. Cruz-Uribe.

[4] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 69 of *Classics in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011. Reprint of the 1985 original.

[5] J.-L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires.* Dunod, Paris, 1969.

[6] J.-L. Lions and E. Magenes. *Problèmes aux limites non homogènes et applications. Vol. 2.* Travaux et Recherches Mathématiques, No. 18. Dunod, Paris, 1968.

[7] J.-L. Lions and E. Magenes. *Non-homogeneous boundary value problems and applications. Vol. I.* Springer-Verlag, New York, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.

[8] J.-L. Lions and J. Peetre. Sur une classe d'espaces d'interpolation. *Inst. Hautes Études Sci. Publ. Math.*, (19):5–68, 1964.

[9] P. Malliavin. *Integration and probability.* Springer-Verlag, New York, 1995. French edition: Masson, Paris, 1982.