Functional Analysis course notes

Second part - a.y. 2015-16

G. Orlandi

1 Introduction

In this second part of the course we analyse the basic theory of linear operators and equations in Banach and Hilbert spaces, in particular the theory of compact operators in Banach and Hilbert spaces and the Lax-Milgram theory, emphasizing the applications to the most common integral and partial differential equations and variational (or optimization) problems. We introduce also appropriate general tools such weak and distributional derivatives, Sobolev spaces and functions of bounded variation. We follow closely the book

[B] Brézis, H.; *Functional Analysis, Sobolev spaces and Partial Differential Equations*, Springer (2010),

referring to the corresponding sections for detailed proofs of the main results presented in this course. Additional content concerning the theory of linear operators is taken from

[K] Kolmogorov, Fomin; *Elements of Theory of Functions and Functional Analysis*, Dover (1999).

A suitable reference for the introductory theory of functions of bounded variations is

[G] Giusti, E.; *Minimal surfaces and functions of bounded variation*, Birkhäuser (1984).

2 Examples of linear problems in Banach and Hilbert spaces

We are interested in solving $Au = f$ for a given $f \in F$, where F is a Banach or Hilbert space, $A : D(A) \subset E \to F$ is a linear operator between two Banach spaces E, F. Some interesting cases are also expressed in the form of the fixed point-type equation $\lambda u = Tu + f$, where $T : E \to E$ is linear (i.e. $A = \lambda I - T$).

In general we say that the problem $Au = f$ is *well-posed* if it enjoys existence and uniqueness of solution, and stability (or continuous dependence) with respect to (perturbation of) the data. This last property is crucial for approximation and discretization of both problem and solution. In case one of the above properties fails, we say that the problem is *ill-posed*. We briefly describe next some motivating examples.

2.1 Initial and boundary value problems for parabolic and elliptic partial differential equations.

Given an electric charge density f in a region Ω having the boundary $\partial\Omega$ connected with earth, i.e. with zero electrostatic potential, the corresponding electrostatic potential u solves the Dirichlet problem (an elliptic PDE) $-\Delta u = f$ on Ω with boundary condition $u = 0$ on $\partial\Omega$, where $\Delta u = \mathrm{div}\ \mathrm{grad}\, u$, the trace of the Hessian of u, is a linear differential operator called the Laplacian of u (which enters in general in the description of equilibrium phenomena). Under certain fairly general assumptions the solution u is unique and is expressed through the integral formula $u(p) = \int_\Omega f(q)G(p,q)dq$, where dq is the Lebesgue measure on Ω and $G(p,q)$ is the so-called Green function for the Laplacian on the domain Ω. In the case $\Omega = B_1(0) \subset \mathbb{R}^2$ we have $G(p,q) = \frac{1}{2\pi}\log|p-q|$, if $\Omega = B_1(0) \subset \mathbb{R}^3$ then $G(p,q) = \frac{1}{4\pi}|p-q|^{-1}$ (i.e. $G(p,q)$ is the Coulomb potential describing the interaction of two unit point charges located at p and q). The Dirichlet problem for the Laplacian is well-posed.

The (homogeneous) heat equation $u_t - \Delta u = 0$ for $u(t,x) \in (0,+\infty) \times \Omega$ with initial condition $u(0,x) = u_0(x)$ (and possibly further boundary conditions on $\partial\Omega$) describes the evolution of the temperature u due to heat diffusion in the domain Ω. If $\Omega = \mathbb{R}^N$, under some suitable fairly general assumptions, the solution is unique (and stable with respect to perturbations of the initial datum, so that the problem is well-posed), and is explicitely given by $u(t,x) = G_t * u_0(x) = \int_{\mathbb{R}^N} u_0(y)G_t(x-y)\,dy$, where $G_t(x) = \frac{1}{\sqrt{2\pi t}^N}\exp(-\frac{|x|^2}{4t})$ is the heat kernel. The heat kernel, being a normal distribution, enters in the description of diffusion phenomena or processes: it describes the transition probability density at time t of a brownian motion (i.e. a random walk with infinitesimal space-time steps) of a particle starting at time $t = 0$ from $0 \in \mathbb{R}^N$. The heat kernel is also referred to as a gaussian convolution kernel or a gaussian filter in the signal processing literature, where $G_t * u_0$ represent a regularization (denoising) at scale t of a given signal u_0 (possibly affected by noise).

In both preceding examples, the solution operator $u = Kf$ is an integral operator: in the case of the heat equation, if we set $v(x) = u(t,x)$ for fixed $t > 0$, we have in particular $v = K_t u_0$ with $K_t u_0 = G_t * u_0$ having the form of a convolution (a weighted average) of the initial datum with the heat kernel G_t.

2.2 Integral equations and integral transforms.

The Volterra equations $\int_\Omega K(x,y)u(y)\,dy = f(x)$ and $u(x) = \int_\Omega K(x,y)u(y)\,dy + f(x)$ in the unknown u for a given kernel $K(x,y)$ and datum f on a domain Ω are examples of integral equations of the type $Ku = f$ or $u = Ku + f$, with K the corresponding integral operator. The problem to prescribe a load u on an elastic horizontal rod (identified with a segment $[a, b]$) in order that it reaches a desired given profile $f(x)$ for $x \in [a, b]$ can be modelized through a Volterra equation for a suitable kernel $K(x, y)$ defined for $x, y \in [a, b]$ that encodes the action of gravity on the load.

The Fourier transform $\mathcal{F}u(\omega) = \int_{\mathbb{R}^N} u(x)e^{-2\pi i x\cdot\omega}dx$, is suitable to define Fourier coefficients or modes relative to the (vector) frequency $\omega \in \mathbb{R}^N$ for a (non-periodic) function u defined on the whole space, and is used to analyse signals in the frequency (or Fourier) space. Under some assumptions (for instance, on $L^2(\mathbb{R}^N)$) it is an invertible integral operator, and it is used to solve certain linear PDE's on the whole space $\mathbb{R}^N$, such as diffusion equations, Maxwell equations of electromagnetism, Schrödinger equation in quantum mechanics: applying the Fourier transform to those PDEs, the resulting equations in the frequency space are of algebraic type.

Further examples of integral transforms are given by the Radon and the Hough transforms, which are used for example to solve inverse problems in tomography. For a given mass density u supported on a planar region Ω, the total mass contained in a line L is given by $\int_L u(z)d\ell(z)$, with $d\ell$ the line element. By scanning the domain Ω with all possible lines in the plane we obtain the output $f = Ru$, the Radon transform of u, which is given by

$$f(r, \theta) = \int_{L_{r,\theta}} u(z)d\ell(z) \qquad L_{r,\theta} = \{(x, y) \in \mathbb{R}^2,\ x\cos\theta + y\sin\theta = r\}.$$

The Radon transform is invertible, so that in principle one can recover the exact mass distribution u given f, but in fact, for instance in the concrete situation of a computer axial tomography, the output f is affected by systematic errors inherent the measurement device and by noise, so that the real inverse problem one is faced to solve is $Ru = f + \epsilon$, where ϵ is an unknown probability distribution encoding error measurements and noise, i.e. it is an ill-posed problem.

2.3 Regularization of ill-posed problems

Typical situations where ill-posed problems appear are inverse (or reconstruction) problems in signal and (biomedical) image processing is a datum f (a function) given by a measured signal (e.g. by a biomedical apparatus such as a magnetic resonance device), which corresponds to a linear transform (e.g. Fourier or Radon transform) of the original datum/signal u, and the task is to recover the original datum u from the measured signal f. In general, besides the linear transform one has to take into account error measurements or noise, so that the effective operator R cannot be expected to be invertible. One thus seeks for approximate solutions according to optimality criteria that penalize both noise and errors (e.g. a least squares optimization criterion among a suitable admissible set $\mathcal{H}$ (the hypothesis space) of signals having an expected structure), so that the resulting well-posed regularized problem can be put in the form $\lambda Bu + Ru = f$, where Bu, the regularization term, is expressed through a suitable invertible operator B, and $\lambda > 0$ is a tuning (or scaling) parameter suited for the specific situation under analysis.

Another similar ill-posed situation, where one has to invert a "low rank" operator R (i.e. with a nontrivial, maybe huge kernel) appears in the approximation of sparse data, a particular instance of *machine learning*. One is given with the values $f(x_i)$ of a function at very few points $x_i \in S$ (the "training set") of a domain D (e.g. a dataset), and wants to infer f on the whole domain D: there are of course many possible arbitrary extensions of f from S to D, moreover the values $f(x_i)$ could be affected by errors, so that the problem is ill-posed. One is led to regularize the problem by asking that the extension u be regular in the sense that it belongs to a certain space $\mathcal{H}$ which is suited for that particular problem, and be close to the data $f(x_i)$ in a least square sense: one is thus led to solve the problem

$$\lambda B u + \sum_i [u(x_i) - f(x_i)] = 0\,,$$

where B is a suitable invertible operator acting on $\mathcal{H}$ and $\lambda > 0$ a tuning parameter. This problem stems equivalently from the minimization of the functional

$$\mathcal{F}(u) = \lambda \|u\|^2_{\mathcal{H}} + \sum_{x_i \in S} |u(x_i) - f(x_i)|^2$$

which is referred to as the *Tychonoff regularization* of the original problem.

3 Linear operators in Banach and Hilbert spaces

Let E, F be two Banach (or Hilbert) spaces, consider the linear operator $T : D(T) \subset E \to F$, where the domain of definition $D(T)$ is a dense subspace of E (it may well be $D(T) = E$).

Recall that T is continuous if and only if it is bounded, i.e. there exists $C > 0$ such that $\|Tv\|_F \leq C\|v\|_E$ for any $v \in D(T)$, and in particular T extends by density to a bounded operator on the whole of E. Here are some examples of bounded operators.

Diagonal operators. The operator $T : \ell^1 \to \ell^1$ given by $(x_i) \mapsto (c_i x_i)$, with $\|(c_i)\|_{\ell^\infty} = \sup_i |c_i| < +\infty$ is bounded, since

$$\|T(x_i)\|_{\ell^1} = \sum_i |c_i x_i| \leq \sup_i |c_i| \sum_i |x_i| = \|(c_i)\|_{\ell^\infty} \|(x_i)\|_{\ell^1}$$

Hilbert-Schmidt operators. Let (a_{ij}) a sequence of numbers such that $\sum_{i,j} a_{ij}^2 = C < +\infty$. Then the operator $T : \ell^2 \to \ell^2$ given by $(x_i) \mapsto (\sum_j a_{ij} x_j)$ is bounded, since

$$\|T(x_i)\|^2_{\ell^2} = \sum_i \left(\sum_j a_{ij} x_j \right)^2 \leq \left(\sum_{i,j} a_{ij}^2 \right) \left(\sum_j x_j^2 \right) = C \|(x_i)\|^2_{\ell^2}\,.$$

By essentially the same calculations one shows that the integral operator $Tu(x) = \int_\Omega K(x,y)u(y)\,dy$ is bounded on the Hilbert space $L^2(\Omega)$ provided $K \in L^2(\Omega \times \Omega)$, and that $\|Tu\|_2 \leq \|K\|_2 \|u\|_2$.

Fourier transform. From the inequality $|\hat{u}(\omega)| \leq \int_{\mathbb{R}} |u(x)|\,|e^{2\pi i \omega x}|\,dx$ it follows that $\hat{u} \in C^0 \cap L^\infty(\mathbb{R})$ whenever $u \in L^1(\mathbb{R})$, and $\|\hat{u}\|_\infty \leq \|u\|_1$, i.e. $u \mapsto \hat{u}$ is a bounded operator from $L^1(\mathbb{R})$ to $C^0 \cap L^\infty(\mathbb{R})$.

Convolution operators. Let $T : u \mapsto g * u$, for $g \in L^1(\mathbb{R}^n)$. By the properties of the convolution product we deduce the estimate

$$\|Tu\|_p = \|g * u\|_p \leq \|g\|_1 \|u\|_p, \qquad \text{i.e. } T \in \mathcal{L}(L^p(\mathbb{R}^n)) \text{ for any } 1 \leq p \leq +\infty\,.$$

3.1 Properties of bounded linear operators.

([B], chap. 2 and 6 and [K], ch. IV, sect. 30)

The space $\mathcal{L}(E,F)$ of bounded linear operators between two Banach spaces E, F is a Banach space when endowed with the operator norm $||T||_{\mathcal{L}} = \sup\{||Tv||_F,\ ||v||_E \leq 1\}$, that measures the maximal elongation factor for unit vectors in E. In particular, we have $\|Tv\|_F \leq \|T\|_{\mathcal{L}} \|v\|_E$ and $\|T\|_{\mathcal{L}}$ is the smallest constant such that this inequality is verified for any $v \in E$. Observe that $\|T \circ S\| \leq \|T\|\|S\|$, i.e. the operator norm is compatible with the composition of operators.

There are various notions of convergence for a sequence of operators $T_n \in \mathcal{L}(E,F)$ to $T \in \mathcal{L}(E,F)$: uniform ($||T_n - T||_{\mathcal{L}} \to 0$), strong ($T_n v \to Tv$ in F $\forall\ v \in F$), weak ($< \phi, T_n v > \to < \phi, Tv > \forall\, v \in F, \forall\, \phi \in F'$).

The Weierstrass criterion for the uniform convergence of series of operators: if $\sum_n \|T_n\|_{\mathcal{L}} < +\infty$, then $T = \sum_n T_n \in \mathcal{L}(E,F)$.

Example (Neumann series). For $T \in \mathcal{L}(E)$ and $||T||_{\mathcal{L}} < 1$, $(I-T)$ is invertible (and $(I-T)^{-1} \in \mathcal{L}(E)$). Moreover, $(I-T)^{-1} = \sum_{n=0}^{+\infty} T^n$ (observe that $\|T^n\| \leq \|T\|^n$ and use W. criterion). In particular, the fixed point equation $u = Tu + f$ has a unique solution $u = (\sum_n T^n)f$, that can be approximated by iterating the operator T on an arbitrary initial guess (for instance, 0 or f itself).

Recall that $T \in \mathcal{L}(E) = \mathcal{L}(E,E)$ is injective and surjective, then $T^{-1} \in \mathcal{L}(E)$ by the open mapping theorem. Reasoning as in the previous example, it follows that the subset of invertible operators is *open* in $\mathcal{L}(E)$: if T is invertible then for any $S \in \mathcal{L}(E)$ such that $||S|| < ||T^{-1}||^{-1}$, the operator $T + S = S \circ (S^{-1}T + I)$ is invertible. This can be rephrased by saying that the invertibility property of an operator is stable under small (or lower order) perturbations.

The adjoint operator $T^* \in \mathcal{L}(F', E')$ is defined through the identity $< T^*\phi, v > = < \phi, Tv >$ for any $v \in F$, $\phi \in F'$. It holds $\|T^*\| = \|T\|$, as a consequence of Hahn-Banach. In case $E = F = H$ a Hilbert space, from the identification $H \equiv H'$ given by the Riesz representation theorem, one considers $T, T^* \in \mathcal{L}(H)$. If $T = T^*$ the operator is called *self-adjoint* or *symmetric*.

Example: Hilbert-Schmidt integral operators on ℓ^2 (respectively $L^2(\Omega)$), induced by symmetric kernel $(a_{ij}) = (\overline{a_{ji}})$ (resp. $K(x,y) = \overline{K(y,x)}$) are self-adjoint operators.

3.2 Compact operators.

([B], sect. 6.1 and [K], chap. IV, sect. 31)

Definition of compact operator: $T \in \mathcal{L}(E,F)$ is compact if $T(B)$ is totally bounded (hence relatively compact) in F whenever B is bounded in E (or equivalently, if the image of the unit ball B_E is relatively compact): in particular, if $\|v_n\|_E \leq M$ is a bounded sequence in E, then Tv_n is convergent in F up to a subsequence.

Uniform limits of compact operators in $\mathcal{L}(E,F)$ are compact, i.e. the space $\mathcal{K}(E,F)$ of compact operators is closed in $\mathcal{L}(E,F)$. Bounded operators whose range is finite dimensional are compact (since in finite dimension boundedness and total boundedness are equivalent, or equivalently, since the closed unit ball is compact in finite dimensions): they are called *finite rank* operators. For example, the orthogonal projection on a finite dimensional subspace of a given Hilbert space is a compact operator.

By the closure property of compact operators it follows that limits of sequences of finite rank bounded operators are compact.

Viceversa, if $F = H$ is a Hilbert space the finite rank approximation property holds true for any $T \in \mathcal{K}(E,H)$: given $v_1, ..., v_N \in H$ a ϵ-net for $T(B_E)$, set $V_N = \text{span}\langle v_1, ..., v_N\rangle$ and $T_N = P_N \cdot T$, where P_N is the orthogonal projection on V_N. We have that T_N has finite rank and $||T_N - T||_{\mathcal{L}(E,H)} \leq 2\epsilon$.

Examples. The diagonal operator $T : (x_n)_n \mapsto (c_n x_n)_n$, where $c_n \to 0$, is compact on ℓ^1 (resp. ℓ^2) as uniform limit of the finite rank operators $T_N : (x_n)_n \mapsto (\sigma_N c_n x_n)_n$, where $\sigma_N = 1$ for $n \leq N$ and $\sigma_N = 0$ for $n > N$. Observe that in case $c_n \neq 0$ for any $n \in \mathbb{N}$, we have that T is injective on ℓ^1, but never surjective: the inverse operator $S : (y_n)_n \mapsto (c_n^{-1} y_n)_n$ is unbounded and its domain of definition is the subspace of ℓ^1 given by those sequences (y_n) such that $\sum_n |c_n^{-1} y_n|$ is finite (in particular, we have necessarily $y_n = o(c_n)$ as $n \to +\infty$).

Hilbert-Schmidt operators are compact: for $(Tu)(x) = \int_\Omega K(x,y)u(y)\,dy$ with $K \in L^2(\Omega \times \Omega)$, we have $||T||_{\mathcal{L}} \leq ||K||_{L^2}$ and given a Hilbert basis (i.e. a complete orthonormal system) $\{\phi_n\}$ of $L^2(\Omega)$, set $\psi_{nm}(x,y) = \phi_n(x)\phi_m(y)$: the elements ψ_{nm} are a Hilbert basis of $L^2(\Omega \times \Omega)$. Expanding $K(x,y) = \sum_{n=1}^\infty \sum_{m=1}^\infty k_{nm}\psi_{nm}(x,y)$, and setting respectively

$$K_N(x,y) = \sum_{n=1}^N \sum_{m=1}^N k_{nm}\psi_{nm}(x,y) \quad \text{and } (T_N u)(x) = \int_a^b K_N(x,y)u(y)\,dy,$$

we have $||T_N - T||_{\mathcal{L}} \leq ||K_N - K||_2 \to 0$, hence $T \in \mathcal{K}(L^2(\Omega))$ as limit of finite rank operators.

Along the same lines one proves that the Hilbert-Schmidt operator on ℓ^2, $T : (x_i) \mapsto (\sum_j a_{ij} x_j)$ with $\sum_{i,j} a_{ij}^2 < +\infty$, belongs to $\mathcal{K}(\ell^2)$.

Fredholm-Volterra operators: if the kernel $K \in C^0(\Omega\times\Omega)$ (possibly except for suitable discontinuity sets of zero measure) and Ω is a compact domain endowed with a finite measure μ, then the integral operator $Tu(x) = \int_\Omega K(x,y)u(y)\,d\mu(y)$ is compact on $C^0(\Omega)$: from the uniform continuity of K on the compact set $\Omega\times\Omega$ we deduce $\|Tu\|_\infty \le \mu(\Omega)|\cdot\|K\|_\infty\|u\|_\infty$, and moreover

$$|Tu(x) - Tu(x')| \le \int_\Omega |K(x,y) - K(x',y)|\,|u(y)|\,dy\ \le \epsilon\mu(\Omega)\cdot\|u\|_\infty \quad \forall\,|x-x'|\le\delta,$$

obtaining the equi-uniformly continuous property of Tu under a uniform bound $\|u\|_\infty \le M$, hence compactness in $C^0(\Omega)$ by Ascoli - Arzelà.

A compact operator (right- or left-) composed with a bounded operator is compact. In particular, $\mathcal{K}(E) \equiv \mathcal{K}(E,E)$ is a bilateral ideal of $\mathcal{L}(E)$. The identity map is compact if and only if E is finite dimensional. Any injective $T \in \mathcal{K}(E)$ doesn't admit a bounded inverse, unless E is finite dimensional.

If E is reflexive (e.g. a Hilbert space), T is compact if and only if for any $u_n \rightharpoonup u$ weakly in E it holds $Tu_n \to Tu$ strongly in E: otherwise said, if and only if weakly converging sequences are transformed in strongly converging sequences, and in particular, $T(\overline{B_E}) = \overline{T(B_E)}$, since the closed convex ball $\overline{B_E}$ is closed for the weak topology.

Proof: given $u_n \rightharpoonup u \in H$, by Banach-Steinhaus we have $\|u_n\| \le M$, hence for a subsequence u_{n_k} it holds $Tu_{n_k} \to v$ in H. Observe that

$$\langle Tu_n, w\rangle = \langle u_n, T^*w\rangle \to \langle u, T^*w\rangle = \langle Tu, w\rangle \quad \forall w \in H,$$

by weak convergence of the sequence u_n, i.e. T is continuous w.r.t. the weak topology. Hence in particular $Tu_{n_k} \rightharpoonup Tu$ and we have necessarily $v = Tu$ by uniqueness of the weak limit (recall that by Hahn-Banach the weak topology separates points, i.e. is Hausdorff), hence the whole sequence Tu_n has Tu as unique (strong) limit point.

Application: **constrained optimization of a quadratic form** involving a compact operator on a Hilbert space. Let $Q(v) = \langle Tv, v\rangle$ for $v \in H$ and $T \in \mathcal{K}(H)$. The quadratic form Q is weakly continuous, i.e. $Q(v) \to Q(v_0)$ if $v \rightharpoonup v_0$, since

$$|\langle Tv, v\rangle - \langle Tv_0, v_0\rangle| \le |\langle Tv - Tv_0, v\rangle| + |\langle Tv_0, v - v_0\rangle| \to 0$$

since $|\langle Tv - Tv_0, v\rangle| \le \|Tv - Tv_0\|\cdot M \to 0$ by strong convergence and $\langle Tv_0, v - v_0\rangle \to 0$ by weak convergence. Hence the constrained optimization problems

$$\max\{Q(v),\ \|v\| \le 1\}, \qquad \min\{Q(v),\ \|v\| \le 1\}$$

have a solution by Weierstrass Theorem, since the set $\overline{B_H} = \{v \in H,\ \|v\| \le 1\}$ is bounded and weakly closed, hence weakly compact in H.

The solution can actually be obtained by the Lagrange multipliers methods, by finding the critical points of the quadratic form $\psi(v,\lambda) = Q(v) + \lambda\|v\|^2$ associated to the operator $T - \lambda I$ (see e.g. the proof of the spectral theorem for compact self-adjoint operators on Hilbert spaces).

Adjoint of a compact operator: if $T \in \mathcal{K}(H)$ then $T^* \in \mathcal{K}(H)$ and conversely.
Example: the adjoint of a Hilbert-Schmidt operator is associated to the conjugate kernel $K^*(x,y) = \overline{K(y,x)}$ (respectively to the sequence $a^*_{ij} = \bar{a}_{ji}$ in the discrete ℓ^2 case)

3.3 Self-adjoint compact operators in Hilbert spaces

([B], sect. 6.4, [K], chap. IX, sect. 59)

Spectral theory for self-adjoint compact operators in Hilbert spaces: the eigenvalues are real and there exists a Hilbert basis made of eigenvectors, which "diagonalizes" the operator. In particular, for $T \in \mathcal{K}(H)$, $T^* = T$, and e_n a orthonormal basis of eigenvectors, i.e. $Te_n = \lambda_n e_n$ (with $\lambda_n \to 0$), we have the diagonal representation $Tv = T(\sum_n c_n e_n) = \sum_n \lambda_n c_n e_n$, i.e. the operator can be identified with the operator $\tilde{T} \in \mathcal{K}(\ell^2)$ given by $\tilde{T}(c_n) = (\lambda_n c_n)$. Moreover, $\|T\|_{\mathcal{L}} = \max_n |\lambda_n|$.

Proof of the spectral theorem: we consider a (iterated) contrained optimization problem on the unit closed ball $\overline{B} = \{\|v\| \leq 1\}$ of H for the quadratic form $Q(v) = \langle Tv, v\rangle$ associated to $T \in \mathcal{K}(H)$. Recall first that $Q(v)$ is weakly continuous, since $v_n \rightharpoonup v_0$ implies $Tv_n \to Tv_0$, and moreover $\|v_n\| \leq M$ (weakly convergent sequences are bounded), whence

$$|\langle Tv_n, v_n\rangle - \langle Tv_0, v_0\rangle| \leq |Tv_n - Tv_0| \cdot |v_n| + |\langle Tv_0, v_n - v_0\rangle| \to 0.$$

By Weierstrass Theorem, $|Q(v)|$ reaches its maximum on the unit closed ball $\overline{B}$, which is weakly compact. Let e_1 be a maximum point. We have necessarily $\|e_1\| = 1$ because $Q(\lambda v) = \lambda^2 Q(v)$ for $\lambda \in \mathbb{R}$. Moreover, for any $e \in H$ such that $\|e\| = 1$ and $\langle e, e_1\rangle = 0$, one has $\langle e, Te_1\rangle = 0$. This follows for instance by applying the Lagrange multipliers theorem to the function

$$\psi(x,y,\lambda) = Q(v) + \lambda\|v\|^2 = x^2\langle Te_1,\, e_1\rangle + 2xy\langle Te_1,\, e\rangle + y^2\langle Te,\, e\rangle + \lambda(x^2+y^2)\,,$$

where $v = xe_1 + ye$ belongs to the 2-dimensional space spanned by e_1 and e. Since e_1 is a critical point of ψ, we deduce $0 = \frac{\partial\psi}{\partial y}(1,0,\lambda) = 2\langle Te_1,\, e\rangle$. In particular, one has $Te_1 = \langle Te_1, e_1\rangle \cdot e_1 = \lambda_1 e_1$, i.e. e_1 is an eigenvector of T and moreover $|Q(e_1)| = |\langle Te_1, e_1\rangle| = |\lambda_1|$, i.e. the eigenvalue λ_1 has maximum modulus among all the eigenvalues of T (actually we have $|\lambda_1| = \|T\|_{\mathcal{L}}$).

Iterating this procedure, one obtains, for $n \geq 1$, an eigenvector e_n of T, with $\|e_n\| = 1$, and such that $< e_n, e_m >= 0$ for any $m < n$, corresponding to the maximum point of $|Q(v)|$ on $(\text{span}\{e_1, ..., e_{n-1}\})^\perp \cap \overline{B}$, with $\lambda_n = Q(e_n)$ the corresponding eigenvalue. Moreover, it holds $|\lambda_{n-1}| \geq |\lambda_n|$.

If for some $n_0 \in \mathbb{N}$ one has $\lambda_{n_0} = Q(e_{n_0}) = 0$, then $(\text{span}\{e_1, ..., e_{n_0-1}\})^\perp = \ker T$. Indeed, $Q(w) = 0$ for any $w \in (\text{span}\{e_1, ..., e_{n_0-1}\})^\perp$, and if $\langle w, e_i\rangle = 0\ \forall\, i < n_0$, then

$\langle Tw, e_i \rangle = \langle v, Te_i \rangle = 0$, i.e. also $Tw \in (\text{span}\{e_1, ..., e_{n_0-1}\})^{\perp}$. The polarization identity $4\langle Tv, u \rangle = Q(u+v) - Q(u-v)$ hence implies that $4\langle Tw, Tw \rangle = Q(w+Tw) - Q(w-Tw) = 0$ for any $w \in (\text{span}\{e_1, ..., e_{n_0-1}\})^{\perp}$, i.e. $Tw = 0$.

We deduce in this case that the set $\{e_1, ..., e_{n_0}\}$, completed with a (complete) orthonormal system $\{e'_j\}$ of $\ker T$ yields a Hilbert basis of eigenvectors of T.

Otherwise, we are left with a orthonormal sequence $\{e_n\}_n$, so that in particular $e_n \rightharpoonup 0$ by Bessel inequality (for any $w \in H$, $\sum_n \langle e_n, w \rangle^2 \leq ||w||^2 \Rightarrow \langle e_n, w \rangle \to 0$ as $n \to +\infty$), and hence $|\lambda_n| = |Q(e_n)| \searrow 0$ by weak continuity of Q. Let $N = \overline{\text{span}\{e_1, ..., e_n, ...\}}^{\perp}$. For any $w \in N$ one necessarily has $|Q(w)| \leq |Q(e_n)|$ for any $n \in \mathbb{N}$, hence $Q(w) = 0$ and $N = \ker T$.

In this case, the set $\{e_n\}_{n \in \mathbb{N}}$, completed with a (complete) orthonormal system $\{e'_j\}$ of $\ker T$ yields a Hilbert basis of eigenvectors of T. □

Using the above Hilbert basis to decompose $u = \sum_n \langle u, e_n \rangle e_n + \sum_j \langle u, e'_j \rangle e'_j$, we may express

$$Tu = \sum_n \lambda_n \langle u,\, e_n \rangle e_n = \left(\sum_n \lambda_n e_n \otimes e_n \right) u \,,$$

where, for $v, w \in H$, the tensor product $v \otimes w$ denotes the rank one operator $u \mapsto \langle u, w \rangle v$.

Thanks to the spectral theorem, under analogous assumptions as in the finite dimensional case, one can prove a singular value decomposition theorem for compact operators on Hilbert spaces, namely the existence of two orthonormal systems $\{e_n\}$ and $\{f_n\}$ and $\lambda_n > 0$ such that

$$Tu = \left(\sum_n \lambda_n e_n \otimes f_n \right) u = \sum_n \lambda_n \langle u,\, f_n \rangle e_n \,.$$

This decomposition is widely used in applications.

3.4 The Fredholm Alternative

([B], sect. 6.3, [K], chap. IV, sect. 31, 32).

The Fredholm alternative gives a procedure to solve equations of the type $Au = u - Tu = f$, with $T \in \mathcal{K}(H)$, H a Hilbert space. We have

1. $\ker A$ is finite dimensional,

2. The range $R(A)$ is closed, hence there holds the orthogonal direct sum decomposition $H = R(A) \oplus \ker A^* = R(A^*) \oplus \ker A$, where $A^* = I - T^*$.

3. $\ker A = 0 \Leftrightarrow R(A) = H$, (in case $T = T^*$, this follows from 2.)

4. $\dim \ker A = \dim \ker A^* < +\infty$.

Proof: we have proved point 1., 2. and (partially) 3., following closely the arguments given in the cited references [B], sect. 6.3 and [K], sect. 31,32.

The Fredholm alternative holds more generally for operators of the type $A = I - T$ with $T \in \mathcal{K}(E)$, E a Banach space.

The procedure to implement the Fredholm alternative goes as follows: first of all, solve the associated adjoint homogeneous equation, i.e. the fixed point equation $v = T^*v$. If the solution is trivial, then $Au = f$ admits a unique solution for any datum $f \in H$ (that can be possibly found in an iterative way via contraction mapping principle, as in the case of Fredholm-Volterra integral operators). Otherwise, call $v_1, ...v_k$ a basis of $\ker A^*$ (i.e. a maximal independent set of fixed points of T^*); then there are solutions of $Au = f$ provided f verifies the orthogonality conditions $< f, v_i >= 0$ for any $i = 1, ..., k$.

3.5 Spectral theory for $\mathbf{T} \in \mathcal{L}(E)$.

([B], sect. 6.2, [K], chap. IV, sect 31, 32)

Definition of resolvent set $\rho(T) \subset \mathbb{C}$: we have $\lambda \in \rho(T)$ if $(\lambda I - T)^{-1} \in \mathcal{L}(E)$. The resolvent set is open in $\mathbb{C}$. Moreover, if $|\lambda| > \|T\|$ then $\{\lambda \in \mathbb{C}\,,\ |\lambda| > ||T||\} \subset \rho(T)$. Actually, denoting $r(T) = \limsup_n (||T^n||)^{1/n} \leq ||T||$ the spectral radius of T, we have $\{\lambda \in \mathbb{C}\,,\ |\lambda| > r\} \subset \rho(T)$. Spectrum $\sigma(T) = \mathbb{C} \setminus \rho(T)$ of $T \in \mathcal{L}(E)$: it is a closed set contained in $B(0, ||T||) \subset \mathbb{C}$. Let $\lambda \in \sigma(T)$: If $\ker(T - \lambda I) \neq 0$ then λ is an eigenvalue of T, and belongs to the point spectrum. Otherwise, λ belongs to the continuous spectrum ($\ker(T - \lambda I) = 0$ but $(T - \lambda I)$ is not surjective). In particular, the map $(T - \lambda I)^{-1}$ may be defined either in a dense or in a proper closed subspace of E, and may be either bounded or unbounded.

Examples. The right shift τ_r in ℓ^1 (or ℓ^2), or the diagonal operator $T_\alpha : \{x_n\} \mapsto \{\alpha_n x_n\}$ where $0 \neq \alpha_n \to 0$. In both cases 0 belongs to the continuous spectrum. Moreover, $\{\alpha_n\} \subset \sigma(T_\alpha)$ is the point spectrum of T_α, while the point spectrum of τ_r is empty.

The multiplication operator $Tu(x) = x \cdot u(x)$ on $C^0([a, b])$. For any $\lambda \in \mathbb{R}$, $\ker(T - \lambda I) = 0$, hence there are no eigenvalues. Moreover, for $\lambda \notin [a, b]$, $(T - \lambda I)^{-1}v(x) = (x - \lambda)^{-1}v(x)$ is well-defined for any $v \in C^0([a, b])$ and is bounded, i.e. $\lambda \in \rho(T)$, while for $a \leq \lambda \leq b$ $(T - \lambda I)^{-1}$ is defined on the dense subspace $\{v \in C^0([a, b]),\ \ v(\lambda) = 0\}$, and it is unbounded. In particular, $\sigma(T) = [a, b]$ is the continuous spectrum of T.

The resolvent operator $R_\lambda = (T - \lambda I)^{-1}$ of $T \in \mathcal{L}(E)$, with $\lambda \in \rho(T)$. Resolvent equation $R_\lambda - R_\mu = (\lambda - \mu)R_\lambda R_\mu$: it yields $\frac{dR_\lambda}{d\lambda} = R_\lambda^2$, that is $\lambda \mapsto R_\lambda$ is a holomorphic function, whose singularities are in $\sigma(T)$. In particular, the Cauchy integral formula (and the calculus of residues) involving R_λ and a given holomorphic function $f(z)$ allows to consistently define $f(T)$ (in particular, if $f(z) = e^z$, we obtain a formula for $\exp(T)$, while if $f(z) = 1$ we derive some information on the Jordan blocks of T).

3.6 Spectrum of a compact operator

([B], sect. 6.3, [K], chap. IV, sect. 31, 32)

Structure of the spectrum of a compact operator: $0 \in \sigma(T)$ and $\sigma(T) \setminus \{0\}$, if non empty, is made of at most countably many eigenvalues λ_n, and in the case of infinite eigenvalues, then $\lambda_n \to 0$ as $n \to +\infty$. The fact that $0 \neq \lambda_n \in \sigma(T)$ is an eigenvalue follows from the Fredholm Alternative, point 3. The corresponding eigenspaces $\ker(\lambda_n I - T) \neq 0$ are finite-dimensional (point 1. of Fredholm Alternative). If the operator is self-adjoint on a Hilbert space, then the eigenvalues are real, and $\max |\lambda_n| = \|T\|_{\mathcal{L}}$ (as a consequence of the Spectral Theorem).

4 Lax-Milgram Theory

4.1 Lax-Milgram and Stampacchia theorem

([B], sect. 5.3)

The Lax-Milgram Lemma: given a bilinear form $a(u,v)$, continuous ($a(u,v) \leq M\|u\|\|v\|$) and coercive ($0 < \alpha\|u\|^2 \leq a(u,u)\ \forall\ u \neq 0$) on a closed subspace S of a Hilbert space H, for any bounded linear form $\phi \in H'$ there exists a unique $u \in S$ such that $a(u,v) = \phi(v)$ for any $v \in S$. In particular, $\|u\| \leq \alpha^{-1}\|\phi\|_*$.

If moreover a is symmetric (i.e. $a(u,v) = a(v,u)$), we have the characterization $u = \arg\min\{\frac{1}{2}a(v,v) - \phi(v)\ ,\ v \in S\}$.

Proof: without loss of generality, the proof is carried out in the case $F = H$. By Riesz representation theorem, the equation to be solved can be rewritten as $\langle Au, v\rangle = \langle f, v\rangle$ for any $v \in H$, i.e. $Au = f$, where $A \in \mathcal{L}(\mathcal{H})$ verifies the estimates $0 < \alpha\|u\| \leq \|Au\| \leq M\|u\|\ \forall\ u \neq 0$.

From $\alpha\|u\| \leq \|Au\|$ (which is called an *a priori estimate*) it follows that $\ker A = 0$. Moreover, $\alpha\|u_n - u_m\| \leq \|Au_n - Au_m\|$ implies that if $y_n = Au_n \to y$ in H, i.e. Au_n is a Cauchy sequence in H, then also u_n is a Cauchy sequence, hence $u_n \to u$ in H by completeness, thus yielding $y = Au$. One concludes that A has a closed range $R(A)$ in H. Finally, if $v \perp R(A)$, then $\langle v, Au\rangle = 0\ \forall\ u \in H$. In particular, choosing $u = v$, we have $0 = \langle v, Av\rangle \geq \alpha\|v\|^2$, thus $v = 0$ and $R(A) = H$. We just proved that A is both injective and surjective, and the conclusion of the Lemma follows.

In case of a symmetric a, since $\alpha\|u\|^2 \leq a(u,u) \leq M\|u\|^2$, the scalar product $((u,v)) := a(u,v)$ is equivalent to $\langle\cdot,\cdot\rangle$, hence by Riesz representation theorem applied to H endowed with $((\cdot,\cdot))$, one has $\phi(v) = a(g,v)$ for a certain $g \in H$, whence u verifies $a(u-g,v) = 0\ \forall\ v \in H$, i.e. u is the orthogonal projection (with respect to the scalar product induced by a) of g on H, in other words u minimizes the (squared) distance (induced by a) $a(v-g, v-g)$, or, equivalently, the quadratic functional $F(v) = \frac{1}{2}a(v,v) - \phi(v)$, for $v \in H$, whose Euler-Lagrange equation $\partial_v F(u) \equiv \langle F'(u), v\rangle = 0$ for any direction $v \in H$ is precisely given by $a(u,v) = \phi(v)$ for any $v \in H$.

$\square$

A generalization of Lax-Milgram lemma is given by Stampacchia theorem (see the corresponding statement in [B], sect. 5.3).

4.2 Galerkin approximation

The Galerkin approximation method: if $V_h \subset H$, $\dim V_h < +\infty$, one considers the solution u_h of the system $a(u,v) = \phi(v)\ \forall\ v \in V_h$. We have the uniform bound $\|u_h\| \le \alpha^{-1}\|\phi\|_*$, which gives weak compactness of the sequence $\{u_h\}$. Moreover, the Lemma of Céa guarantees that $\|u - u_h\| \le \frac{M}{\alpha}\mathrm{dist}\,(u, V_h)$ (in other words, u_h is comparable to the orthogonal projection of u on V_h): indeed, $a(u-u_h, u-u_h) = a(u-u_h, u-v)$ for any $v \in V_h$ since $a(u, v-u_h) = a(u_h, v-u_h) = \phi(v-u_h)$, whence $\alpha\|u-u_h\|^2 \le M\|u-u_h\|\|u-v\|$ for any $v \in V_h$ and the conclusion follows.

Hence, considering a sequence of finite-dimensional spaces $V_h \subset V_{h+1}$ such that $H = \overline{\cup_h V_h}$, one has the convergence $u_h \to u$ in H as $h \to +\infty$.

Remark that the approximating finite-dimensional problem is a linear system with a positive definite coefficients matrix, called *stiffness matrix*, which is given by $[a(f_i, f_j)]$, with $\{f_i\}$ a basis for V_h.

The choice of the sequence V_h invading H and of a basis $\{f_i\}$ for V_h is aimed to efficiently solve the approximating linear system, and also to have the best possible convergence rate for the error estimate $||u_h - u||$. Here are some examples in case $H = L^2(\Omega)$, $\Omega \subset \mathbb{R}^n$:

1) if a is represented by a compact self-adjoint operator, then considering a Hilbert basis $\{e_n\}_{n\in\mathbb{N}}$ of $L^2(\Omega)$ made of eigenvectors, and setting V_h =span$< e_1, ..., e_h >$, the corresponding system is diagonal.

2) considering a basis $\{f_i\}$ of V_h made of *finite elements* (piecewise linear or polynomial function insisting on a fixed triangulation of the domain) yields a sparse stiffness matrix. Finite elements are used in numerical fluid dynamics, material science, elasticity,...

3) Haar basis, wavelets, radial basis functions: these Hilbert basis of $L^2(\Omega)$ are used in signal and image processing and statistical analysis, being not computationally expensive, and also since they are able to take into account localized oscillation phenomena at any scale in physical and in frequency space.

4) if the original problem admits a smooth solution (for example, $u \in C^\infty(\Omega)$ as for Laplace equation), it may be convenient to use *spectral methods* for its approximation, i.e. to consider a Hilbert basis of $L^2(\Omega)$ made of orthogonal polynomials (e.g. the trigonometric system, the Legendre polynomials, the Hermite polynomials, the Tchebycheff polynomials): since the Lemma of Céa states that the error estimate $\|u - u_h\|$ is comparable to the distance of u to its orthogonal projection on V_h, hence the convergence rate will be better according to the regularity of u (for instance, the more regular u, the more rapidly its Fourier coefficients decay to 0).

5 Boundary value problems in dimension one

([B], sect. 8.1, 8.3, 8.4, 8.5, 8.6)

Weak / variational formulation of elliptic boundary value problems in dimension 1. Classical vs weak solutions. Natural ambient spaces for weak solutions of boundary value problems are Sobolev spaces. The space $W_0^{1,p}([a,b])$ is defined as the closure of $C_c^1([a,b])$ w.r.t. the $W^{1,p}$ norm, it is a closed (and hence complete) subspace of $W^{1,p}([a,b])$ and can be characterized as follows: $W_0^{1,p}([a,b]) = \{v \in W^{1,p}([a,b]),\ v(a) = v(b) = 0\}$. Recall the compact embedding of $W^{1,p}([a,b])$ (hence of $W_0^{1,p}([a,b])$) in $C^0([a,b])$. By the Poincaré inequality $\|u\|_p \leq C_p|b-a|\|u'\|_p$ valid on $W_0^{1,p}([a,b])$, the L^p norm of the derivative gives an equivalent norm on $W_0^{1,p}([a,b])$. The Hilbert space $H_0^1([a,b]) = W_0^{1,2}([a,b])$.

The **homogeneous Dirichlet problem** $-u'' + u = f$ on $[a,b]$, $u(a) = u(b) = 0$: weak formulation in $H_0^1([a,b])$ in case $f \in L^2([a,b])$ as

$$a(u,v) = \langle u, v\rangle_{H^1} = \int_a^b u'v' + uv = \int_a^b fv = \phi(v)$$

for any $v \in H_0^1([a,b])$. Application of Lax-Milgram theory for existence, uniqueness, a priori estimates, variational characterization of the weak solution as the minimizer on $H_0^1([a,b])$ of the Dirichlet energy $F(v) = \frac{1}{2}\|v'\|_2^2 + \frac{1}{2}\|v-f\|_2^2$. Hilbert regularity: the weak solution belongs to $H_2^2 \cap H_0^1([a,b])$, i.e. its second (weak) derivative exists and belongs to $L^2([a,b])$. Classical regularity of the weak solution in case $f \in C^0([a,b])$.

Remark. The energy functional $F(v)$ can be used for the time discretization of the gradient flow of the energy $E(v) = \frac{1}{2}\|v'\|_2^2$, starting from a given function $u_0 \in L^2([a,b])$. That gradient flow corresponds in turn to the initial/boundary value problem for the homogeneous heat equation $u_t - u_{xx} = 0$ with initial datum $u_0(x)$ and homogenous boundary data $u(t,a) = u(t,b) = 0$ for any $t > 0$.

Indeed, consider the following Euler-type discretization scheme of the heat equation: fix the discretization step $\Delta t \equiv \lambda^{-1}$ and denote $t_n = n\Delta t$, $u(t_n, \cdot) = v_n(\cdot)$. To solve discretized equation $\frac{1}{\Delta t}(v_{n+1} - v_n) - v''_{n+1} = 0$ with $v_0 = u_0$ one may iteratively obtain v_{n+1} as the minimizer in of the Dirichlet energy functional $F_\lambda(v) = \frac{1}{2}\|v'\|_2^2 + \frac{\lambda}{2}\|v - v_n\|_2^2$ on $H_0^1([a,b])$.

The functional $F_\lambda(v) = \|v'\|_2^2 + \lambda\|v-f\|_2^2$ gives also an example of Tychonoff regularization for ill-posed problems in signal (or image) processing: it contains the fidelity (least square) term $\|v-f\|_2^2$ tuned by the parameter $\lambda > 0$ and the regularization term $\|v'\|_2^2$ penalizing oscillations of v.

Analysis of the (homogeneous) Sturm-Liouville problem (it is an elliptic equation in divergence form) $-(pu')' + qu = f$. Compactness and symmetry of the solution operator $u = Tf$ as an operator on $L^2([a,b])$ thanks to the compact embedding $H_0^1([a,b]) \subset L^2([a,b])$. The **spectral decomposition of Sturm-Liouville operators** ([B], sect. 8.6) produces Hilbert basis of $H_0^1([a,b])$ made by special functions used in spectral (Galerkin) approximation methods (trigonometric polynomials, Legendre, Tchebycheff, Hermite polynomials,...)

Integral representation of the solution operator T (compact and symmetric on $H_0^1([a,b])$): let $e_n(x)$ a Hilbert basis of $H_0^1([a,b])$ made of eigenvectors of T related to the eigenvalues $\mu_n \in \mathbb{R}$, with $\mu_n \to 0$. Then $f = \sum\langle f, e_n\rangle e_n(x)$ and $u = Tf = \sum_n \mu_n\langle f, e_n\rangle e_n(x)$, where $\langle f, e_n \rangle = \int_a^b f(y) e_n(y)\, dy$. Hence, by dominated convergence, we may interchange the sum with the integral, obtaining $u(x) = \int_a^b (\sum_n \mu_n e_n(x) e_n(y)) f(y)\, dy$. We recover the fact that T is the Hilbert-Schmidt integral operator on $L^2([a,b])$ given by $Tf(x) = \int_a^b G(x,y) f(y)\, dy$ with the symmetric Green kernel $G \in L^2$ given by $G(x,y) = \sum_n \mu_n e_n(x) e_n(y)$.

Remark. The corresponding parabolic problem $u_t - (pu_x)_x + qu = f$, $u(0,x) = u_0(x) \in L^2([a,b])$, $u(t,a) = u(t,b) = 0$ can be solved by taking advantage of a Hilbert basis of eigenvectors of the spatial operator through the method of separation of variables: let $f(x) = \sum_n f_n e_n(x)$ and seek for a solution $u(t,x) = \sum_n c_n(t) e_n(x)$ with $u(0,x) = u_0(x) = \sum_n c_{0,n} e_n(x)$, plug it into the equation, obtaining formally

$$\sum_n [\dot{c}_n(t) + \mu_n^{-1} c_n(t) - f_n] e_n(x) = 0 \qquad \text{for any } t > 0,\ x \in [a,b]\,,$$

whence $c_n(t)$ can be explicitely determined as the solution of the Cauchy problem $\dot{c}_n + \mu_n^{-1} c_n = f_n$, $c_n(0) = c_{0,n}$. A rigorous justification of the above representation formula for the solution is obtained by passing to the limit in the corresponding finite dimensional Galerkin approximation, by virtue of uniform bounds in the a priori estimates on the approximate solution. For instance, in the case $f \equiv 0$ we obtain $c_n(t) = c_{0,n} \exp(-t/\mu_n)$ whence we deduce the L^2 bound (uniform in t) for $u(\cdot, t)$

$$\begin{aligned}
\int_a^b |u(t,x)|^2 dx\, dt &= \int_a^b \left(\sum_n c_n(t) e_n(x) \right)^2 dx \\
&= \int_a^b \sum_n [c_n(t) e_n(x)]^2 dx = \sum_n c_{0,n}^2 \exp(-2t/\mu_n) \\
&\le \exp(-2t/\mu_1) \sum_n c_{0,n}^2 = \exp(-2t/\mu_1) \|u_0\|_2^2\,.
\end{aligned}$$

Analogously, the hyperbolic problem (e.g. vibrating string, wave equations) $u_{tt} - (pu_x)_x + qu = 0$ can be transformed in the family of second order o.d.e's $\ddot{c}_n + \mu_n^{-1} c_n = 0$, with the initial conditions $c_n(0)$ and $\dot{c}_n(0)$ given by the Fourier coefficients of the initial data $u(0,x)$ and $u_t(0,x)$.

Reaction-diffusion-transport equation in dimension one: the bilinear form $a(u,v)$ on $H_0^1([a,b])$ associated to the equation $-(pu')' + ru' + qu = f$ is not symmetric, and is coercive in some cases, for instance if $p(x) > \alpha > 0$, $q(x) > \beta > 0$ and $\|r\|_\infty < 2\sqrt{\alpha\beta}$, since in this case there exists $\gamma > 0$ such that

$$a(u,u) \ge \alpha \|u'\|_2^2 + \beta \|u\|_2^2 - \|r\|_\infty \|u'\|_2 \|u\|_2 > \gamma \|u'\|_2^2,$$

and coercivity on $H_0^1([a,b])$ is granted by Poincaré inequality.

Example of weak formulation of a non homogeneous Dirichlet problem within the framework of Stampacchia Theorem and a homogeneous Neumann problem.

Maximum principle, stability, well-posedness of elliptic and parabolic initial/boundary value problems: for the Dirichlet model problem $-u''+u=f$ on $[a,b]$, $u(a)=\alpha$, $u(b)=\beta$ we have the pointwise bound $\min\{\alpha,\beta,\inf_{[a,b]} f\}\le u(x)\le \max\{\alpha,\beta,\sup_{[a,b]} f\}$ for any $x\in[a,b]$, that gives a control of the sup norm of the solution in term of the sup norm of the data. As a consequence one obtains uniqueness and stability (i.e. continuous dependence) of the solution with respect to the data, i.e. well-posedness of the Dirichlet problem. For a classical solution the maximum principle follows by observing that a maximum (resp. minimum) point x_0 for u is either on the boundary (hence controlled by α,β), or in the interior, in which case it is a critical point (i.e. $u'(x_0)=0$) and $u''(x_0)\le 0$ (resp. $u''(x_0)\ge 0$), whence $u(x_0)\le f(x_0)$ (resp. $u(x_0)\ge f(x_0)$).

To prove the maximum principle in the general framework of weak formulations one uses Stampacchia truncation method ([B], sect. 8.5).

6 Boundary value problems in dimension N

6.1 Reaction - diffusion - transport equations

Let $u(x)$ denote the concentration of a chemical diffusing in a fluid medium (e.g. air, water) occupying a region of space Ω, which in turn is transported by a velocity vector field V which is supposed to be divergence-free (i.e. the fluid has no sources nor sinks, and every streamline of the vector field is either closed or starts and ends on the boundary of any fluid subregion). The evolution of the concentration of the chemical is governed by a balance law that relates the rate of change of the concentration in a given region with the flux Φ_u accross the boundary and the mass production inside the region, originated possibly by some source term having density $f(x)$ or some reaction term $g(u)$ which in the simplest can be assumed to be linear, i.e. $g(u)=\alpha u$ (Malthusian model) but could be also nonlinear, e.g. $g(u)=\alpha u(\beta-u)$ (logistic model). Thus we may write, for any $A\subset\Omega$ having smooth boundary $\partial A\subset\Omega$, denoting by n its exterior unit normal,

$$-\frac{d}{dt}\int_A u(x)\,dx=\int_{\partial A}\Phi_u\cdot n\,d\sigma+\int_A g(u)+f(x)\,dx$$

whence, by the divergence theorem,

$$\int_A (u_t-\operatorname{div}\Phi_u)\,dx=\int_A g(u)+f(x)\,dx\,,\qquad \text{for any } A\subset\Omega\,.$$

Passing to the limit as A concentrates around a point $x\in\Omega$, by the mean-value theorem we may deduce the differential equation $u_t-\nabla\cdot\Phi_u=g(u)+f$. The flux due to the drift by the velocity field V is uV, and assuming that the flux due to diffusion is proportional to ∇u (Fourier law), we have $\Phi_u=k\nabla u+uV$, so that we recover the following reaction - drift - diffusion equation (recall we assume $\operatorname{div}V=0$)

$$u_t-k\Delta u+V\cdot\nabla u=g(u)+f\qquad \text{in }\Omega\,,$$

that is complemented by an initial condition u_0 and suitable boundary conditions on $\partial\Omega$ of Dirichlet (e.g. fixing the concentration on $\partial\Omega$) or Neumann type (e.g. fixing the flux accross $\partial\Omega$: the homogeneous Neumann condition $\frac{\partial u}{\partial n} = 0$ means an insulation condition).

In case $g(u) = 0$ and $V = 0$ (pure diffusion) we get the heat equation, that models the diffusion of heat (here u is the temperature distribution).

Drift-diffusion equations arise also to compute the transition probability u underlying a stochastic process (e.g. a brownian motion with drift) to pass from a given initial state to the state $x \in \Omega$ at time t (Fokker-Planck equation).

They also model, in their nonlinear version, some evolution phenomena in population dynamics (e.g. in epidemiology).

6.2 Weak formulation of the Dirichlet problem in $\mathbf{H}^1_0(\boldsymbol{\Omega})$

([B], sect. 9.4, 9.5)

Consider the model Dirichlet problem $-\Delta u + u = f$ in a bounded open set $\Omega \subset \mathbb{R}^N$ with $\partial\Omega$ of class C^1 (or piecewise C^1), with homogeneous Dirichlet boundary condition $u = 0$ on $\partial\Omega$. If $f \in C^0(\Omega)$ then u is a classical solution if $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$. Recall first the integration by parts formula

$$\int_{\partial\Omega} \phi \frac{\partial u}{\partial n} = \int_\Omega \operatorname{div}(\phi \nabla u) = \int_\Omega \nabla\phi \cdot \nabla u + \int_\Omega \phi \Delta u \,.$$

Multiplying the equation $-\Delta u + u = f$ by a test function $\phi \in C^\infty_c(\Omega)$ (it is enough $\phi \in C^1_c(\Omega)$) and then integrating by parts, taking into acccount that ϕ vanishes on the boundary, we obtain that a classical solution verifies the equivalent formulation

$$\int_\Omega \nabla u \cdot \nabla \phi + \int_\Omega u\phi = \int_\Omega f\phi \qquad \forall\, \phi \in C^1_c(\Omega)\,,$$

that has the form

$$a(u,v) = \langle u, v\rangle_{H^1} := \langle u, v\rangle_{L^2} + \sum_{j=1}^N \langle \frac{\partial u}{\partial x_j}, \frac{\partial v}{\partial x_j}\rangle_{L^2} = \langle f, v\rangle_{L^2} = \varphi(v) \quad \forall v \in C^1_c(\Omega)\,.$$

It is hence natural to look for a weak formulation of the dirichlet problem in the Hilbert space $H^1_0(\Omega) = \overline{C^1_c(\Omega)}^{H^1}$, i.e. the closure of $C^1_c(\Omega)$ with respect to the euclidean H^1 norm, which, as defined above, is the sum of the L^2 norms of the function and of its partial derivatives.

Observe that for $u \in H^1_0(\Omega)$ one may define weak derivatives, since for $u_n \in C^1_c(\Omega)$, $u_n \to u$ in the H^1 metric, we have that $\frac{\partial u}{\partial x_i} \to g_i$ in $L^2(\Omega)$ for any $i = 1, ..., N$, and hence the integration by parts formula gives

$$\int_\Omega u \frac{\partial \phi}{\partial x_i} = \lim_{n\to+\infty} \int_\Omega u_n \frac{\partial \phi}{\partial x_i} = -\lim_{n\to+\infty} \int_\Omega \frac{\partial u}{\partial x_i}\phi = -\int_\Omega g_i \phi\,,$$

i.e. $g_i \in L^2(\Omega)$, for $i = 1, ..., N$ are the weak partial derivatives of u, and we still denote them by $\frac{\partial u}{\partial x_i}$.

Hence, the bilinear form $a(u, v)$ (i.e. the H^1 inner product), for $u, v \in H_0^1(\Omega)$ has the same representation as for C_c^1 functions, just interpreting partial derivatives in the weak sense.

It follows by Lax-Milgram (or Riesz representation theorem in $H_0^1(\Omega)$), that the weak formulation of the Dirichlet problem

$$\langle u, v\rangle_{H^1} \equiv \langle u, v\rangle_{L^2} + \sum_{j=1}^{N} \langle \frac{\partial u}{\partial x_j}, \frac{\partial v}{\partial x_j}\rangle_{L^2} = \langle f, v\rangle_{L^2} \qquad \forall v \in H_0^1(\Omega)$$

has a unique solution $u \in H_0^1(\Omega)$ for any given $f \in L^2(\Omega)$ (observe that since $|\varphi(v)| = |\int_\Omega fv| \leq \|f\|_2 \|v\|_2 \leq \|f\|_2 \|v\|_{H^1}$ we have $\varphi \in (H_0^1)'$).

Hilbert regularity of the weak solution: we have $u \in H^2 \cap H_0^1(\Omega)$, i.e. u has weak second partial derivatives in $L^2(\Omega)$. The proof of this fact is not immediate at all and makes use of the so-called technique of difference quotients.

Classical regularity of the weak solution: in case $f \in C^0(\Omega)$ it holds $u \in C^2(\Omega) \cap C^0(\overline{\Omega})$, hence u is a classical solution.

Variational characterization of the weak solution:

$$u = \arg\min \left\{ F(v) = \frac{1}{2} \int_\Omega |\nabla v|^2 + \frac{1}{2} \int_\Omega |v - f|^2 \, , \quad v \in H_0^1(\Omega) \right\} \, .$$

Remark. As remarked in the one dimensional case, the functional $F_\lambda(v) = \|\nabla u\|_2^2 + \lambda \|v - f\|_2^2$ is used for the time discretization of the heat flow of the energy $E(v) = \|\nabla u\|_2^2$ with time step $\Delta t = \lambda^{-1}$ and also for the regularization of ill-posed problems (arising e.g. in image processing).

Remark. In general $H_0^1(\Omega)$ is not contained in $C^0(\Omega)$, so that in particular the boundary condition $u = 0$ on $\partial\Omega$, differently from the one dimensional situation, cannot be given a pointwise sense (the boundary has null measure and u is defined up to sets of null measure!). One can nevertheless give a meaning to the boundary value of u in $L^2(\partial\Omega)$ (in the so-called trace sense) by taking suitable limits of averages of u near the boundary.

Remark. A dense class in $H_0^1(\Omega)$, suitable for Galerkin approximation of the Dirichlet problems, is given by (linear) finite elements, compactly supported in a triangulated mesh of the domain. The corresponding stiffnes matrix is sparse.

Poincaré inequality on $H_0^1(\Omega)$. As in the one dimensional case, the L^2 norm of the gradient gives an equivalent norm to the H^1 norm on $H_0^1(\Omega)$, by virtue of Poincaré inequality, valid for bounded domains. Indeed, for $u \in C_c^1(\Omega)$ we have

$$u(x) = \int_{-M}^{x_1} \frac{\partial u}{\partial x_1}(t, x_2, ..., x_N) \, dt \, ,$$

where $\Omega \subset B_M(0)$. By Cauchy-Schwarz we obtain $|u(x)|^2 \le 2M \int_M^M |\partial_1 u|^2 \, dt$, and integrating on Ω we obtain, by Fubini,

$$\int_\Omega |u(x)|^2 \, dx \le 4M^2 \int_\Omega \left| \frac{\partial u}{\partial x_1} \right|^2 \, dx \le 4M^2 \int_\Omega |\nabla u(x)|^2 \, dx \, ,$$

i.e. $\|u\|_{L^2(\Omega)} \le 2M \|\nabla u\|_{L^2(\Omega)}$.

6.3 Elliptic problems in divergence form

([B], Theorem 9.23)

For $a_{ij}(x) \in L^\infty(\Omega)$, $i, j = 1, ..., N$, $a_i(x) \in L^\infty(\Omega)$ for $i = 0, ..., N$, $f(x) \in L^2(\Omega)$ consider the weak formulation in $H_0^1(\Omega)$ of the following problem in divergence form (anisotropic diffusion - transport - reaction equation)

$$-\sum_{i,j=1}^N \frac{\partial}{\partial x_i} \left(a_{ij}(x) \frac{\partial u}{\partial x_j} \right) + \sum_{i=1}^N a_i(x) \frac{\partial u}{\partial x_i} + a_0(x) u = f(x) \quad \text{in } \Omega$$

with $u = 0$ on $\partial\Omega$. The weak formulation reads

$$\int_\Omega \sum_{i,j=1}^N a_{ij}(x) \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} + \sum_{i=1}^N a_i(x) \frac{\partial u}{\partial x_i} v + a_0(x) u v = \int_\Omega f v \quad \forall\, v \in H_0^1(\Omega) \, ,$$

Hence it is of the form $a(u, v) = \varphi(v)$ (with a non symmetric). If $\sum_{i,j=1}^N a_{ij}(x) \xi_i \xi_j \ge \alpha |\xi|^2$ for some $\alpha > 0$ then, for some sufficiently large $\lambda > 0$ the bilinear form $a(u, v) + \lambda \langle u, v \rangle_{L^2}$ is coercive on $H_0^1(\Omega)$ (compare the argument in the one dimensional case), hence by Lax-Milgram, for any $\Phi \in H_0^1(\Omega)'$ there exists a unique solution of $a(u, v) + \lambda \langle u, v \rangle_{L^2} = \varphi(v)$ for any $v \in H_0^1(\Omega)$.

Denoting by $f \mapsto Tf$ the solution operator, we have $T : L^2(\Omega) \to H_0^1(\Omega) \subset L^2(\Omega)$, with compact embedding (cfr. Rellich theorem), hence $T \in \mathcal{K}(L^2(\Omega))$ and u solves the fixed point equation $u = T(f + \lambda u)$. Setting $w = f + \lambda u$ we have that w solves $w = \lambda T w + f$, i.e. $(I - \lambda T) w = f$. Suppose $f = 0$ (hence $w = \lambda u$). Then, either the homogeneous equation $u = \lambda T u$ is uniquely solvable (if and only if the original equation has zero as unique solution, and this is precisely the case for instance if $a_0(x) \ge \beta > 0$, by the maximum principle (see [B], Remark 27 after Theorem 9.27), and hence there exists a unique solution $u \in H_0^1(\Omega)$ for any $f \in L^2(\Omega)$, or the homogeneous equation has a (finite) d-dimensional set of solutions, whence the original problem is solvable if f is orthogonal to a suitable d-dimensional subspace of $L^2(\Omega)$.

7 Distributions, Sobolev spaces and functions of bounded variation

7.1 Distributions and distributional derivatives

Given $u \in L^1_{loc}(\Omega)$, we may identify u with the distribution (linear functional on $C^\infty_c(\Omega)$) $T_u(\phi) = \int_\Omega u(x)\phi(x)\,dx$. One advantage of the distributional point of view is to embed functions in dual spaces, where the Banach-Alaoglu compactness theorem holds true.

Recall **Banach-Alaoglu theorem**: given a Banach space X, bounded sequences in X' are relatively compact with respect to the weak* topology, i.e. for $T_j \in X'$, $\|T_j\|_* \le C$, there exists $T \in X'$ such that, up to a subsequence, $T_j(\phi) \to T(\phi)$ for any $\phi \in X$. Remark that the weak* topology is a priori weaker than the weak topology, where one tests convergence against any element of the dual space, i.e. in this case $(X')' = X''$, which contains strictly X in the non reflexive case. Of course, if X is reflexive, weak and weak* convergence do coincide.

Supposing for simplicity that Ω is bounded and $u \in L^1(\Omega)$, we have $|T_u(\phi)| \le \|u\|_1\|\phi\|_\infty$. i.e. T_u extends to a bounded functional on $C^0_c(\Omega) = \overline{C^\infty_c(\Omega)}^{L^\infty}$, the space of continuous functions vanishing on $\partial\Omega$. Moreover,

$$\|T_u\|_* = \sup\{T_u(\phi),\ \|\phi\|_\infty \le 1\} = \|u\|_1$$

as it follows by considering test functions ϕ converging to the function $\mathbf{1}_{\{u>0\}} - \mathbf{1}_{\{u<0\}}$, which in turn gives that $u\phi$ converges to $|u| = u^+ + u^-$. We thus have an isometric injection $L^1(\Omega) \to [C^0_c(\Omega)]^*$.

Measures as distributions. The distribution T_μ associated to a (possibly σ-finite) Radon measure μ on Ω is defined in the same way: $T_\mu(\phi) \equiv \langle T_\mu, \phi\rangle = \int_\Omega \phi(x)d\mu(x)$. For any $\mu \in \mathcal{M}(\Omega)$ (the space of finite Radon measures on Ω) it actually holds $T_\mu \in [C^0_c(\Omega)]^*$ and

$$\|T_\mu\|_* = \sup\{\langle T_\mu, \phi\rangle\,,\ \phi \in C^0_c(\Omega)\,,\ \|\phi\|_\infty \le 1\} = |\mu|(\Omega),$$

where $|\mu| = \mu^+ + \mu^-$ is the total variation measure of $\mu = \mu^+ - \mu^-$ (Hahn decomposition) and $\|\mu\| = |\mu|(\Omega)$ is the *total variation* of the measure μ on Ω (as for L^1 function, one tests with a function $\phi \simeq \mathbf{1}_{spt\mu^+} - \mathbf{1}_{spt\mu^-}$).

For $\mu, \nu \in \mathcal{M}(\Omega)$ two Radon measures on Ω, $T_\mu = T_\nu$ implies $\int_\Omega \phi(x)\,d\mu(x) = \int_\Omega \phi(x)\,d\nu(x)$, for any $\phi \in C^\infty_c(\Omega)$, so that by considering ϕ_j converging to the characteristic function of an open subset $A \subset \Omega$ se obtain, by the dominated convergence theorem, that $\mu(A) = \nu(A)$ for any $A \subset \Omega$ open, hence $\mu = \nu$ by the regularity of Radon measures. Hence $\mu \mapsto T_\mu$ gives an injection $\mathcal{M}(\Omega) \to [C^0_c(\Omega)]^*$. Hence (finite and σ-finite) Radon measures in Ω are distributions (of order zero).

Example (Dirac mass as a distribution): let δ_0 be the Dirac mass concentrated at $0 \in \mathbb{R}^N$. We have

$$T_{\delta_0}(\phi) = \int_{\mathbb{R}^N} \phi d\delta_0 = \phi(0)\,,$$

since

$$\int_{\mathbb{R}^N} \phi d\delta_0 = \sup\{c_0 \in \mathbb{R},\ \phi(x) \geq \sum c_i \mathbf{1}_{E_i}(x),\ E_0 \ni 0\} = \sup\{c_0 \leq \phi(0)\} = \phi(0)\,.$$

in particular

$$\|T_{\delta_0}\|_* = \sup\{\phi(0),\ \phi \in C_c^0(\mathbb{R}^N), \|\phi\|_\infty \leq 1\} = 1.$$

Recall **Riesz representation theorem**: $\mu \in \mathcal{M}(\Omega) \mapsto T_\mu \in [C_c^0(\Omega)]^*$ is an isomorphism of Banach spaces, where the norm of a measure $\mu \in \mathcal{M}(\Omega)$ is given by its *total variation* $\|\mu\| = |\mu|(\Omega)$. In particular, it holds $|\mu|(\Omega) = \|\mu\| = \|T_\mu\|_*$. Weak* compactness and convergence in the sense of measures: by the Banach-Alaoglu theorem (bounded sets in the dual of a Banach space are relatively compact with respect to the weak* topology), a sequence of equibounded Radon measures μ_n on Ω (i.e. $|\mu_n|(\Omega) \leq C$) is weakly* compact, i.e. there exists a subsequence μ_{n_k} and a measure μ such that $\mu_{n_k} \stackrel{*}{\rightharpoonup} \mu$ in $\mathcal{M}(\Omega)$, or in other words $\int_\Omega \phi(x)\, d\mu_{n_k}(x) \to \int_\Omega \phi(x)\, d\mu(x)$ for any $\phi \in C_c^0(\Omega)$.

The space of Radon measures is suited to solve optimization problems involving L^1 or total variation norms. For example, for $Y \subset \mathcal{M}(\Omega)$, Y closed convex and bounded, consider the minimization problem

$$\inf\{\|\mu\| = \mu(\Omega) = \|T_\mu\|_*,\ \mu \in Y\}\,.$$

Then, since Y is closed and convex it is weakly* closed by Hahn-Banach, and since it is also bounded it is weakly* compact by Banach-Alaoglu. The total variation norm is weakly* lowersemicontinuous being characterized as a supremum, as a dual norm. Hence the existence of a minimizer in Y is guaranteed by Weierstrass theorem (or equivalently by the direct method of the calculus of variations).

Example: the **Monge-Kantorovich optimal mass transport problem** consists in minimizing the transportation cost of displacing a given amount M of mass (or economical goods) whose initial distribution in a (metric measure) space X is described by a measure μ_1 of total mass $\mu_1(X) = M$, in order to reach a final destination described by a measure μ_2 in a (metric measure) space Y (with total mass $\mu_2(Y) = M$). Observe that, for $M = 1$, μ_1 and μ_2 are probability measures. Denote by $c(x, y)$ the transportation cost function for displacing a unit of mass (or good) from location $x \in X$ to location $y \in Y$ (if X and Y are subset of a euclidean space, say $\mathbb{R}^N$, typical cost functions are $c(x, y) = |x - y|$ or $c(x, y) = |x - y|^2$), and denote by $T \subset \mathcal{M}(X \times Y)$ the set of transport plans from μ_1 to μ_2, i.e. the measures μ on $X \times Y$ having marginals μ_1 and μ_2 (i.e. $\int_Y d\mu(x, y) = \mu_1(x)$ and $\int_X d\mu(x, y) = \mu_2(y)$). Observe that in case μ_1 and μ_2 are discrete probability measures this is a linear programming problem. The set T is an affine (hence convex) closed subset of $\mathcal{M}(X \times Y)$ (fixing the marginals is a continuous constraint with respect to the convergence of measures), hence the constrained optimization problem

$$\min_{\mu \in T} F(\mu) = \int_{X\times Y} c(x,y) d\mu(x,y)\,.$$

admit a solution by Banach-Alaoglu theorem, assuming lower semi-continuity for the cost function.

Derivative of a distribution: for u and T_u as above, we can define the (i-th partial) derivative of T_u as the linear functional $\partial_i T_u(\phi) = -\int_\Omega u \frac{\partial \phi}{\partial x_i}$. A priori, since $|\partial_i T_u(\phi)| \leq \|u\|_1 \|\partial_i \phi\|_\infty$, we have $\partial_i T_u \in [C_c^1(\Omega)]^*$. This definition is coherent with the integration by parts formula, since for smooth functions u one has $\partial_i T_u(\phi) = T_{\partial_i u}(\phi)$ for any test function ϕ. In particular, since integration by parts holds true for functions in $W^{1,p}(\Omega)$, we have that Sobolev functions may be defined as those functions in $L^p(\Omega)$ such that their distributional derivatives are (associated to) functions in $L^p(\Omega)$.

Example: for $u(x) = |x|$, $x \in \mathbb{R}$, we have

$$-\int_{\mathbb{R}} u\phi' = -\int_{-\infty}^{0} x\phi'(x)\,dx + \int_0^{+\infty} x\phi'(x)\,dx = -\int_{-\infty}^{0} \phi(x)\,dx + \int_0^{+\infty} \phi(x)\,dx\,,$$

hence it follows that $(T_u)' = T_v$ where $v(x) = 1$ if $x > 0$ and $v(x) = -1$ if $x < 0$.

Example (distributional derivative of the Heaviside function): consider the Heaviside function H on $\mathbb{R}$, i.e $H(x) = 1$ for $x \geq 0$, and $H(x) = 0$ for $x < 0$, and consider the following approximating sequence for H, given by $u_j(x) = 1$ if $x \geq j^{-1}$ and $u_j(x) = j \cdot x$ for $0 \leq x \leq j^{-1}$, $u_j(x) = 0$ for $x < 0$. We have $(T_{u_j})' = -T_{v_j}$ where $v_j(x) = j$ for $0 \leq x \leq j^{-1}$, and $v_j(x) = 0$ elsewhere in $\mathbb{R}$. Notice that $\|T_{v_j}\|_* = \int_{\mathbb{R}} v_j(x)\,dx = 1$ for any $j \in \mathbb{N}$, hence, up to a subsequence, $T_{v_j} \rightharpoonup T$ weakly* in $[C_c^0(\mathbb{R})]^*$ by Banach-Alaloglu theorem, i.e.

$$T(\phi) = \lim_j T_{v_j}(\phi) = j\int_0^{1/j} \phi(x)\,dx = \phi(0) = T_{\delta_0}(\phi) \qquad \forall\, \phi \in C_c^0(\mathbb{R})\,.$$

On the other hand, we have $-T(\phi) = \lim_j T_{u_j}(\phi') = T_H(\phi')$, hence $T = (T_H)'$. Hence, in $\mathbb{R}$ we have $T_{\delta_0} = (T_H)'$, i.e. the Dirac mass is the distributional derivative of the Heaviside function. Observe that the Heaviside function has a classical derivative a.e. (actually everywhere except in the origin), which is identically zero, while the distributional derivative of H keeps the information on the unit jump of H at the origin.

7.2 The Sobolev spaces $\mathbf{W^{1,p}(\Omega)}$

([B], sect. 9.1) Sobolev spaces arise naturally when solving PDE's through weak formulations, or in order to deal with optimization problems, such as for example the variational (minimum) problem

$$\min\left\{F(v)=\int_\Omega|\nabla v|^p+\lambda\int_\Omega|v-f|^q,\quad v=0\;\text{ on }\partial\Omega\right\}$$

that arises in many contexts, especially in the cases $q=2$ (least square) and $p=1,2$, or in the case $1<p=q$. Existence of minimizers is found in suitable Hilbert and Banach spaces called Sobolev spaces or Functions of Bounded Variation, where the dual space type norms are controlled by F, ensuring at a glance lowersemicontinuity of F and compactness of minimizing sequences with respect to weak (or weak*) convergences.

We refer to [B], sect. 9.1 for details. Definition of $W^{1,p}(\Omega)$ as the space of L^p functions having weak partial derivatives in $L^p(\Omega)$. The $W^{1,p}$ norm. Some elementary properties of $W^{1,p}(\Omega)$: completeness, reflexivity, separability, according to the exponent p. The Hilbert space $H^1(\Omega)=W^{1,2}(\Omega)$. The space $W^{1,\infty}(\Omega)$ coincides with the space of Lipschitz functions on Ω, when Ω is a regular domain (with boundary of class C^1 or lipschitz).

In case $\Omega=\mathbb{R}^N$ Sobolev norms may be defined through suitable weighted norms on $L^2(\mathbb{R}^N_\omega)$, where ω is the frequency (or Fourier) variable. For example, by Plancherel theorem, we have $\|u\|^2_{H^1}=\int_{\mathbb{R}^N}(1+4\pi|\omega|^2)|\hat{u}(\omega)|^2d\omega$.

The closed subspace $W^{1,p}_0(\Omega)\subset W^{1,p}(\Omega)$ is defined as the closure with respect of the $W^{1,p}$ norm of $C^1_c(\Omega)$, and is suited to deal with homogeneous Dirichlet boundary conditions. On $W^{1,p}_0(\Omega)$ there holds the Poincaré inequality $\|v\|_p\le C\|\nabla v\|_p$.

Density of smooth functions in $W^{1,p}(\Omega)$: one extends a function $u\in W^{1,p}(\Omega)$ by zero outside Ω, obtaining a function $\tilde{u}$, and then regularize $\tilde{u}$ by convolution with a family of smoothing kernels (or symmetric mollifiers) $\rho_n\in C^\infty_c(\mathbb{R}^N)$ such that $\rho_n=n^N\rho_1(nx)$, $\rho_1\ge 0$ supported in the unit ball and $\int_{B_1}\rho_1=1$. Denoting by $u_n=\tilde{u}*\rho_n$, we have $u_n\in C^\infty_c(\mathbb{R}^N)$, and one deduces $u_{n|\Omega}\to u$ in $L^p(\Omega)$, while due to the discontinuity of $\tilde{u}$ on the boundary of Ω, we deduce $\nabla u_{n|\omega}\to\nabla u$ in $[L^p(\omega)]^N$ for any $\omega\subset\subset\Omega$. If the boudary is regular, say piecewise C^1 then one may prove by slightly modifying the construction of u_n that actually there exists $u_n\in C^\infty_c(\mathbb{R}^N)$ such that $\nabla u_{n|\Omega}\to\nabla u$ in $[L^p(\Omega)]^N$.

Recall the notion of approximation of identity: for ρ_n as above, consider the (positive) measures $\mu_n=\rho_n(x)\,dx$. We have $\mu_n\rightharpoonup\delta_0$ in the sense of measures, where δ_0 is the Dirac mass concentrated at zero. Indeed, since the support of ρ_n is $B_{1/n}$ and the total integral is 1, we have $\mu_n(A)\to 0=\delta_0(A)$ if the open set A does not contain 0, and $\mu_n(A)\to 1=\delta_0(A)$ otherwise. One concludes by using the regularity of both (Radon) measures: if they coincide on open sets, they coincide on any Borel set.

Recall further that $\rho_n\,dx\rightharpoonup\delta_0$ implies $\rho_n*u\to u$ in $L^p(\mathbb{R}^N)$ and $\|\rho_n*u\|_p\le\|\rho_n\|_1\|u\|_p=\|u\|_p$. Similarly, $\partial_i(\rho_n*u)\to\partial_i u$ in $L^p(\mathbb{R}^N)$ for any $\in W^{1,p}(\mathbb{R}^N)\equiv W^{1,p}_0(\mathbb{R}^N)$.

Characterization of maps in $W^{1,p}(\Omega)$, $\Omega\subset\mathbb{R}^N$ bounded, for $p>1$ (see [B], Proposition 9.3 for a clean statement):

1) L^p functions whose (weak) derivatives are identified (in the distributional sense) to bounded linear functionals on $L^{p'}(\Omega)$,
2) L^p functions having uniformly bounded differential quotients w.r.t the L^p norm.

(for the notion of distribution and distributional derivative see next section).

In case $p = 1$ the property corresponding to 1) is expressed by

1') L^1 functions whose distributional derivatives are identified to bounded linear functionals on $C^0_c(\Omega)$.

In case $p = 1$ properties 1') and 2) are still equivalent, and they characterize the space $BV(\Omega) \subset L^1(\Omega)$ of functions of bounded variation, which is strictly larger than $W^{1,1}(\Omega)$.

Some comments on properties 1) (resp. 1')) and 2). Property 1) emphasize the fact that the weak derivative, seen a distribution, belongs to a dual space, yielding by Banach-Alaoglu weak (resp. weak*) compactness for sequences of gradients uniformly bounded with respect to the dual norm, which by the Riesz representation theorem corresponds, in case $p > 1$, to the L^p norm, while in case $p = 1$, to the so-called total variation norm, which is the norm on the space of Radon measures on Ω, $\mathcal{M}(\Omega) \equiv [C^0_c(\Omega)]^*$ by Riesz representation theorem.

Remark. (Variational problems in $\mathbf{W^{1,p}}$) Minimizing sequences for the functional $F(v) = \|\nabla v\|_p^p + \|v - f\|_p^p$ are uniformly bounded with respect to the $W^{1,p}$ norm, hence they converge weakly (up to a subsequence) in $W^{1,p}$ for $p > 1$. Hence, by lower semicontinuity of F with respect to weak convergence (L^p norms, being equivalent to the dual norm in $(L^{p'})'$ are lower semicontinuous by Banach - Steinhaus) we obtain existence of minimizers of F in $W^{1,p}$.

In the case $p = 1$ there exist sequences uniformly bounded with respect to the $W^{1,1}$ norm that do not converge in $W^{1,1}$ (for example functions approximating the Heaviside function, see the example in the previous section), so that in particular $W^{1,1}$ is not a suitable ambient space for variational problems. On the other hand, property 1') ensures weak* compactness in $[C^0_c(\Omega)]^*$ of sequence of gradients uniformly bounded in L^1, so that in particular they converge in $BV(\Omega)$ (see section 7.3), which is therefore suited to handle variational problems for functionals involving the L^1 norm of the gradient, such as the so-called Rudin - Osher - Fatemi model for image processing $F_\lambda(v) = \|\nabla v\|_1 + \lambda \|v - f\|_2^2$.

Property 2) implies that the hypothesis of Fréchét-Kolmogorov theorem (which gives a criterion for compactness in the strong topology of $L^p(\Omega)$ for $p \geq 1$) is satisfied. In particular, the injection $W^{1,p}(\Omega) \to L^p(\Omega)$ for $p > 1$ (resp. the injection $BV(\Omega) \to L^1(\Omega)$) is compact (Rellich-Kondrachov compact embedding theorem).

Proof of the compact injection of $W^{1,p}(\Omega)$ in $L^p(\Omega)$ for bounded $\Omega \subset \mathbb{R}^N$: if $\mathcal{F}$ is a bounded family of $W^{1,p}(\Omega)$, and $\omega \subset\subset \Omega$, then $\rho_n * \mathcal{F}_{|\omega}$ is ϵ-close to $\mathcal{F}_{|\omega}$ for large n, and uniformly bounded in L^∞ and equi-uniformly continuous, hence relatively compact with respect to the $\|\cdot\|_\infty$ norm by Ascoli-Arzelà, so that in particular it is relatively compact in $L^p(\omega)$. An ϵ-net for $\rho_n * \mathcal{F}_{|\omega}$ in $L^p(\omega)$ is then a 2ϵ-net for $\mathcal{F}_{|\omega}$ and a 3ϵ-net for $\mathcal{F}$ in $L^p(\Omega)$ if ω is sufficiently close in measure to Ω.

7.3 BV functions and variational problems

([G])

Distributional gradient of a function $u \in L^1(\Omega)$: it is a vector distribution defined by $\langle DT_u, \vec{\varphi}\rangle = -\langle T_u, \operatorname{div}\vec{\varphi}\rangle$ for $\vec{\varphi} \in [C_c^\infty(\Omega)]^N$. Remark that for distributions associated to functions u of class C^1 previous formula coincides with Gauss-Green (or integration by parts) formula.

Definition of the space $BV(\Omega)$ (functions of bounded variation): $u \in BV(\Omega)$ if $u \in L^1(\Omega)$ and the distributional gradient $Du = (\partial_1 u, ..., \partial_N u)$ is a (vector) Radon measure $\vec{\mu} = (\mu_1, ..., \mu_N)$, which satisfies the integration by part formula (Gauss-Green)

$$\int_\Omega u \operatorname{div}\vec{\phi} = -\int_\Omega \vec{\phi}\cdot d\vec{\mu} \qquad \text{for any } \vec{\phi} \in [C_c^0(\Omega)]^N .$$

In other words, $u \in BV(\Omega)$ if $u \in L^1(\Omega)$ and $\partial_i T_u(\varphi) = \int_\Omega \varphi d\mu_i = T_{\mu_i}(\phi)$.

Total variation of a vector Radon measure: for $\vec{\mu} = (\mu_1, ..., \mu_n)$ with $\mu_i \in \mathcal{M}(\Omega) = (C_c^0(\Omega))'$ we have the decomposition $\vec{\mu} = \vec{\nu}|\vec{\mu}|$, where $|\vec{\mu}|$ is a positive measure (called the total variation measure) and $|\vec{\nu}(x)| = 1$ for $|\vec{\mu}|$ a.e. $x \in \Omega$. The total variation of $\vec{\mu}$ is defined as the dual norm in $[C_c^0(\Omega)]^*$

$$\|\vec{\mu}\|_* = \sup\left\{\int_\Omega \vec{\phi}\cdot d\vec{\mu} = \int_\Omega \vec{\phi}\cdot\vec{\nu}\, d|\vec{\mu}|\,, \quad \vec{\phi} \in [C_c^0(\Omega)]^n, \quad \|\vec{\phi}\|_\infty \le 1\right\} = |\vec{\mu}|(\Omega).$$

For $u \in BV(\Omega)$ we define its BV norm

$$\|u\|_{BV(\Omega)} = \|u\|_{L^1(\Omega)} + |Du|(\Omega)\,.$$

Remark that for $u \in W^{1,1}(\Omega)$ the BV norm coincides with the $W^{1,1}$ norm. Endowed with such a norm, $BV(\Omega)$ is a Banach space.

Characterization of $BV(\Omega)$: properties 1') and 2) in previous section are equivalent characterizations of $BV(\Omega)$. In particular, 2) implies compactness in the strong L^1 topology of sequences of functions equibounded with respect to the BV norm.

Example: the characteristic function $\mathbf{1}_E$ of an open bounded set $E \subset \mathbb{R}^n$ with $\partial E \cap \Omega$ of class C^1 belongs to $BV(\Omega)$, since by Gauss-Green formula

$$D\mathbf{1}_E(\vec{\phi}) = -\int_E \operatorname{div}\vec{\phi}\, dx = -\int_{\partial E} \vec{\phi}\cdot\vec{n}\, d\sigma\,,$$

where $\vec{n}$ is the unit outer normal to ∂E and $d\sigma$ is the surface measure on ∂E, so that $|D\mathbf{1}_E(\vec{\phi})| \le \|\vec{\phi}\|_\infty \cdot \text{Area}(\partial E \cap \Omega)$, i.e. $D\mathbf{1}_E$ is a vector Radon measure, and in particular $D\mathbf{1}_E = \vec{\nu}|D\mathbf{1}_E|$, where $\vec{\nu}(x) = -\vec{n}(x)$ is the inner unit normal to $\partial E \cap \Omega$ and $|D\mathbf{1}_E| = d\sigma$. By a suitable choice of the test function $\vec{\phi}$ in such a way that $|\vec{\phi}(x)| \le 1$ and $\vec{\phi} = -\vec{n}$ on $\partial E \cap \Omega$ one gets $|D\mathbf{1}_E|(\Omega) = \text{Area}(\partial E \cap \Omega)$.

An example of **variational problem in BV: the Rudin-Osher-Fatemi model** $F(u) = |Du|(\Omega) + \frac{1}{2}\|u - f\|^2_{L^2(\Omega)}$, for $u \in BV(\Omega)$ is suitable in image processing (u denotes for instance the grey level of an image), since it has the ability to regularize (denoise) a given image f (identified as its grey level function) by preserving at the same time edges and boundaries, since characteristic functions are members of BV. Existence of a minimizer in $BV(\Omega)$ on a bounded Ω is guaranteed because F controls the full BV norm and F is lower semicontinuous with respect to the weak* convergence of the total variation and weak L^2 convergence. Its Euler-Lagrange equation results in a nonlinear PDE.

An example of **geometric variational problem: the isoperimetric and the isovolumetric problem** within the class of finite perimeter sets. The two problem are equivalent (and dual each other), since they amount to maximize the (dilation invariant) isoperimetric ratio $|\Omega|^{\frac{n-1}{n}}/|\partial\Omega|$.

Definition of finite perimeter (or Caccioppoli) sets in $\Omega \subset \mathbb{R}^n$: they are Lebesgue measurable sets $E \subset \mathbb{R}^n$ such that

$$P_\Omega(E) \equiv |D\mathbf{1}_E|(\Omega) \equiv \sup\left\{\int_{E\cap\Omega} \text{div}\vec{\phi},\ \|\vec{\phi}\|_\infty \le 1,\quad \vec{\phi} \in [C_c^\infty(\Omega)]^n\right\} < +\infty,$$

in other words $\mathbf{1}_E \in BV(\Omega)$.

Observe that for finite perimeter sets $E \subset \mathbb{R}^n$, the Sobolev embedding theorem $BV(\mathbb{R}^n) \to L^{\frac{n}{n-1}}(\mathbb{R}^n)$ applied to $\mathbf{1}_E$ yields the isoperimetric inequality

$$|E|^{\frac{n-1}{n}} \le C|D\mathbf{1}_E|(\mathbb{R}^n) = CP_{\mathbb{R}^n}(E).$$

The minimal constant $C > 0$ in the above inequality corresponds to the isoperimetric ratio of a round ball in $\mathbb{R}^n$.

Weak formulation of the isovolumetric problem in the class of finite perimeter sets in $\mathbb{R}^n$: fix $R > 1$ (sufficiently large) and set

$$\mathcal{P} = \left\{E \subset \overline{B_R(0)},\ \mathcal{L}^n(E) = \int_{\mathbb{R}^n} \mathbf{1}_E\, d\mathcal{L}^n = 1,\ \mathbf{1}_E \in BV(\mathbb{R}^n)\right\},$$

i.e. $\mathcal{P}$ contains sets $E \subset B_R(0)$ having unit volume and finite perimeter $\|D\mathbf{1}_E\| \equiv |D\mathbf{1}_E|(\mathbb{R}^n)$. Consider the isovolumetric problem

$$\min_{E\in\mathcal{P}} \|D\mathbf{1}_E\|\,.$$

If $E_n \in \mathcal{P}$ is a minimizing sequence, i.e. $\|D\mathbf{1}_{E_n}\| \to \inf_{F\in\mathcal{P}} \|D\mathbf{1}_F\|$, we have

$$\|\mathbf{1}_{E_n}\|_{BV} = 1 + \|D\mathbf{1}_{E_n}\| \le C,$$

so that, up to a subsequence, $\mathbf{1}_{E_n} \to \mathbf{1}_E$ in $L^1(B_{2R}(0))$ since the supports of E_n are contained in $B_{2R}(0)$ and hence there holds the compact embedding of $BV(B_{2R}(0))$ in $L^1(B_{2R}(0))$. We immediately deduce $E \subset \overline{B_R(0)}$ and $\mathcal{L}^n(E) = 1$. Moreover, we have $D\mathbf{1}_{E_n}(\vec{\phi}) \to D\mathbf{1}_E(\vec{\phi})$ for any $\vec{\phi} \in [C_c^\infty(\mathbb{R}^n)]^n$ (i.e. convergence in the sense of distributions) and

$$\|\nabla \mathbf{1}_E\| \leq \liminf_{n\to+\infty} \|\nabla \mathbf{1}_{E_n}\| = \inf_{F\in\mathcal{P}} \|\nabla \mathbf{1}_F\|$$

by lower semicontinuity of the total variation norm. Hence E has minimum perimeter in the class $\mathcal{P}$.

The regularity theory (based for example on Steiner symmetrization) allows to conclude that the optimal set E is the unit volume round ball in $\mathbb{R}^n$.