Operator theory for electronic structure calculation

Part I: Introduction to functional analysis

Eric CANCES

Ecole des Ponts and INRIA Paris, France

Mini-school on mathematics in electronic structure theory of GDR CORREL
Paris, January 9-11, 2017

1 - Metric spaces

Definition (metric space). **Let $\mathcal{M}$ be a set. A function $d : \mathcal{M} \times \mathcal{M} \to \mathbb{R}_+$ satisfying for all $(u, v, w) \in \mathcal{M}^3$,**

1. $d(u, v) = 0 \quad \Leftrightarrow \quad u = v$;
2. $d(v, u) = d(u, v)$;
3. $d(u, w) \leq d(u, v) + d(v, w)$ **(triangular inequality),**

is called a metric (or a distance) on $\mathcal{M}$. A set $\mathcal{M}$ endowed with a metric is called a metric space.

Example 1: **some useful metrics on $\mathcal{M} = \mathbb{R}^n$**

- **Euclidean distance:** $d_2(\mathbf{x}, \mathbf{y}) := \left(\sum_{i=1}^N |x_i - y_i|^2 \right)^{1/2}$;
- **Taxi (or Manhattan) distance:** $d_1(\mathbf{x}, \mathbf{y}) := \sum_{i=1}^N |x_i - y_i|^2$;
- **SNCF (or British Rail) distance:** $d_{\mathrm{SNCF}}(\mathbf{x}, \mathbf{y}) = d_2(\mathbf{x}, \mathbf{0}) + d_2(\mathbf{0}, \mathbf{y})$ **for** $\mathbf{x} \neq \mathbf{y}$.

None of these metrics is *a priori* better than the other ones. In practice, one has to choose an appropriate metric for the specific problem (QOI, numerical method) under consideration.

Example 2: **some metrics on** $\mathcal{M} = \left\{\rho : [0,1] \to \mathbb{R},\ \rho \textbf{ cont.},\ \rho \geq 0,\ \int_0^1 \rho = 1\right\}$

- $d_{L^p}(\rho_1, \rho_2) := \left(\int_0^1 |\rho_1(x) - \rho_2(x)|^p \, dx\right)^{1/p}$, $\quad 1 \leq p < +\infty$;
- $d_{L^\infty}(\rho_1, \rho_2) = d_{C^0}(\rho_1, \rho_2); = \max_{x \in [0,1]} |\rho_1(x) - \rho_2(x)|$;
- $d_{\sqrt{H^1}}(\rho_1, \rho_2) := \left(\int_0^1 |\sqrt{\rho_1(x)} - \sqrt{\rho_2(x)}|^2 \, dx + \int_0^1 \left|\frac{d\sqrt{\rho_1}}{dx}(x) - \frac{d\sqrt{\rho_2}}{dx}(x)\right|^2 dx\right)^{1/2}$.

  Note that $d_{\sqrt{H^1}}$ **is an extended metric: it takes values in** $\mathbb{R}_+ \cup \{+\infty\}$**.**

$$\forall (\rho_1, \rho_2) \in \mathcal{M} \times \mathcal{M}, \quad d_{L^1}(\rho_1, \rho_2) \leq d_{L^p}(\rho_1, \rho_2), \quad d_{L1}(\rho_1, \rho_2) \leq 2 d_{\sqrt{H^1}}(\rho_1, \rho_2),$$

$$|q_1(\rho_N^0) - q_1(\rho^0)| \leq d_{L^1}(\rho_N^0, \rho^0), \quad |q_2(\rho_N^0) - q_2(\rho^0)| \leq \left(q_2(\rho_N^0) + q_2(\rho^0)\right)\ d_{\sqrt{H^1}}(\rho_N^0, \rho^0),$$

$$\rho^0(x) = 1, \quad \rho_N^0(x) = 1 + \frac{\sin(N\pi x)}{1 + \sqrt{N}}, \quad d_{L^p}(\rho_N^0, \rho^0) \underset{N \to +\infty}{\longrightarrow} 0, \quad q_2(\rho^0) = 0, \quad q_2(\rho_N^0) \underset{N \to +\infty}{\longrightarrow} +\infty.$$

Topology of metric spaces

Let $(\mathcal{M}, d)$ be a metric space.

Definition (open subsets of $\mathcal{M}$). A subset U of $\mathcal{M}$ is called open if

$$\forall v \in U,\ \exists \varepsilon > 0 \mid B(v, \varepsilon) := \{w \in \mathcal{M} \mid d(v, w) < \varepsilon\} \subset U.$$

Definition (closed subsets of $\mathcal{M}$). A subset F of $\mathcal{M}$ is called closed if $\mathcal{M} \setminus F$ is open.

Proposition.

- **A union of open subsets of $\mathcal{M}$ is an open subset of $\mathcal{M}$.**
- **An intersection of closed subsets of $\mathcal{M}$ is a closed subset of $\mathcal{M}$.**

Definition (closure of a subset of $\mathcal{M}$). Let $A \subset \mathcal{M}$. The closure of A is the subset of $\mathcal{M}$ denoted by $\overline{A}$ and defined as the smallest closed subset of $\mathcal{M}$ containing A.

Definition (dense subsets of $\mathcal{M}$). A subset D of $\mathcal{M}$ is called dense if $\overline{D} = \mathcal{M}$.

Topology of metric spaces (continued)

Definition (bounded subsets of $\mathcal{M}$). A subset B of $\mathcal{M}$ is called bounded if

$$\mathrm{diam}(B) := \sup_{(v,w)\in B\times B} d(v,w) < \infty.$$

Definition (converging sequences). A sequence $(v_n)_{n\in\mathbb{N}}$ of elements of $\mathcal{M}$ converges in $\mathcal{M}$ (endowed with the metric d) if there exists $v \in V$ s.t.

$$d(v_n, v) \underset{n\to\infty}{\longrightarrow} 0.$$

If such a v exists, it is unique and is called the limit of the sequence $(v_n)_{n\in\mathbb{N}}$. We denote $v_n \underset{n\to\infty}{\longrightarrow} v$ in $\mathcal{M}$ (implicitly endowed with the metric d).

Definition (continuous functions). Let $(\mathcal{M}_1, d_1)$ and $(\mathcal{M}_2, d_2)$ be two metric spaces. A function $f : \mathcal{M}_1 \to \mathcal{M}_2$ is continuous if

$$v_n \underset{n\to\infty}{\longrightarrow} v \text{ in } \mathcal{M}_1 \quad \Rightarrow \quad f(v_n) \underset{n\to\infty}{\longrightarrow} f(v) \text{ in } \mathcal{M}_2.$$

Topology of metric spaces (continued)

Theorem (characterization of closed subsets and dense subsets in metric spaces).

- **A subset F of $\mathcal{M}$ is closed if and only if the limit of any sequence of elements of F which converges in $\mathcal{M}$ is in F.**
- **A subset D of $\mathcal{M}$ is dense if any point of $\mathcal{M}$ is the limit of a sequence of elements of D.**

Examples

- **$\mathbb{Q}$ is not closed in $\mathbb{R}$ but is dense in $\mathbb{R}$;**
- **let us endow $C^0([0,1],\mathbb{R}) := \{v : [0,1] \to \mathbb{R},\ v$ continuous$\}$ with the metric**

$$d_{C^0}(v,w) := \max_{x \in [0,1]} |v(x) - w(x)|.$$

Let $P := \{v : [0,1] \to \mathbb{R},\ v$ polynomial function$\}$.

Then, P is an (infinite dimensional) vector subspace of $C^0([0,1],\mathbb{R})$, which is not closed, but is dense (Weierstrass approximation theorem).

Notion of completeness

Let $(\mathcal{M}, d)$ be a metric space.

Definition (Cauchy sequences). **A sequence $(v_n)_{n\in\mathbb{N}}$ of elements of $\mathcal{M}$ is called Cauchy if**

$$\forall \varepsilon > 0, \quad \exists N \in \mathbb{N} \quad \textbf{s.t.} \quad \forall q \geq p \geq N, \quad d(u_p, u_q) \leq \varepsilon.$$

Any converging sequence is Cauchy, but the converse is not always true!

Definition (completeness). **The metric space $(\mathcal{M}, d)$ is called complete if any Cauchy sequence of elements of $\mathcal{M}$ converges (to an element of $\mathcal{M}$).**

Important remark:

1. **if a metric space $(\mathcal{M}, d)$ is not complete, it usually means that $\mathcal{M}$ is not the right set to consider;**
2. **any metric space $(\mathcal{M}, d)$ can be embedded in a "canonical" way into a complete metric space $(\widetilde{\mathcal{M}}, \widetilde{d})$, called the completed space, such that $\mathcal{M}$ is dense in $\widetilde{\mathcal{M}}$ and $\widetilde{d} = d$ on $\mathcal{M}\times\mathcal{M}$ (ex: $(\mathbb{R}, d_2)$ is the completed of $(\mathbb{Q}, d_2)$).**

2 - Normed vector spaces

Normed vector spaces are special, very important, instances of metric spaces.

In this section, $\mathbb{K}$ denotes $\mathbb{R}$ or $\mathbb{C}$.

Definition ($\mathbb{K}$-vector space). A $\mathbb{K}$-vector space is a set V endowed with

1. **an addition law $V \times V \ni (u,v) \mapsto u+v \in V$ such that $(V,+)$ is an abelian group;**
2. **a scalar multiplication law $\mathbb{K} \times V \ni (\lambda, v) \mapsto \lambda v \in V$,**

such that for all $(\lambda, \mu, v, w) \in \mathbb{K} \times \mathbb{K} \times V \times V$,

$$\lambda(\mu v) = (\lambda\mu)v, \quad 1v = v, \quad \lambda(v+w) = \lambda v + \lambda w, \quad (\lambda+\mu)v = \lambda v + \mu v.$$

The elements of a vector space are called vectors.

Definition (basis and dimension). Let V be a $\mathbb{K}$-vector space. A family $(e_1, \cdots, e_d)$ of d elements of V is called a basis of V if for any $v \in V$, there exists a unique $(\lambda_1, \cdots, \lambda_d) \in \mathbb{K}^d$ such that

$$v = \sum_{i=1}^{d} \lambda_i v_i.$$

If V admits a basis, then all the bases of V have the same cardinality. This number is called the dimension of V.

If V does not admit a basis, it is called an infinite dimensional vector space.

Examples of vector spaces:

- $\mathbb{K}^d$ **is a** d**-dimensional** $\mathbb{K}$**-vector space;**
- **the set** $\mathbb{K}^{m\times n}$ **of matrices of size** $m \times n$ **is a** $\mathbb{K}$**-vector space of dimension** mn**;**
- **the set of real square matrices of size** $n \times n$ **is an** $\mathbb{R}$**-vector space of dim.** $\frac{n(n+1)}{2}$**;**
- **the set**

$$P_n := \{v : [0,1] \to \mathbb{K},\ v \textbf{ polynomial function of degree} \leq n\}$$

  is an $\mathbb{K}$**-vector space of dimension** $n+1$**;**
- **the set**

$$P_\infty := \{v : [0,1] \to \mathbb{K},\ v \textbf{ polynomial function}\}$$

  is an infinite dimensional $\mathbb{K}$**-vector space;**
- **the set**

$$C^0([0,1], \mathbb{K}) := \{v : [0,1] \to \mathbb{K},\ v \textbf{ continuous}\}$$

  is an infinite dimensional $\mathbb{K}$**-vector space.**

Definition (norm). **Let V be a $\mathbb{K}$-vector space ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$). A norm on V is a mapping $\|\cdot\| : V \to \mathbb{R}_+$ satisfying for all $(\lambda, v, w) \in \mathbb{K} \times V \times V$,**

$$\|\lambda v\| = |\lambda|\, \|v\|, \qquad \|v + w\| \leq \|v\| + \|w\|, \qquad (\|v\| = 0 \Leftrightarrow v = 0)\,.$$

A norm $\|\cdot\|$ on V defines a metric on V

The distance between two elements v and w of V (for the norm $\|\cdot\|$) is defined as

$$d(v, w) = \|v - w\|.$$

Example **(finite dimensional case): $V = \mathbb{K}^n$**

- **l^p-norm:** $\|\mathbf{x}\|_p := \left(\sum_{1 \leq i \leq d} |x_i|^p \right)^{1/p}$, $\quad 1 \leq p < \infty$;
- **l^∞-norm:** $\|\mathbf{x}\|_\infty := \max_{1 \leq i \leq n} |x_i|$

Definition (norm). **Let** V **be a** $\mathbb{K}$**-vector space (**$\mathbb{K} = \mathbb{R}$ **or** $\mathbb{C}$**). A norm on** V **is a mapping** $\|\cdot\| : V \to \mathbb{R}_+$ **satisfying for all** $(\lambda, v, w) \in \mathbb{K} \times V \times V$**,**

$$\|\lambda v\| = |\lambda| \, \|v\|, \qquad \|v + w\| \leq \|v\| + \|w\|, \qquad (\|v\| = 0 \Leftrightarrow v = 0) \, .$$

A norm $\|\cdot\|$ **on** V **defines a metric on** V

The distance between two elements v **and** w **of** V **(for the norm** $\|\cdot\|$**) is defined as**

$$d(v, w) = \|v - w\|.$$

Example (infinite dimensional case): $V = C^0([0, 1], \mathbb{K})$**.**

- L^p**-norm (**$1 \leq p < \infty$**):** $\|v\|_{L^p} := \left(\int_0^1 |v|^p \right)^{1/p}$**;**
- L^∞**-norm (or** C^0**-norm):** $\|v\|_{L^\infty} = \|v\|_{C^0} := \sup_{x \in [0,1]} |v(x)|$**.**

Definition (equivalence of norms). Let V be a vector space and $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on V. The norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent if there exist $0 < c \leq C < \infty$ such that

$$\forall v \in V, \quad c\|v\|_1 \leq \|v\|_2 \leq C\|v\|_1.$$

Importance of the notion of equivalent norms. Two equivalent norms give rise to the same topology. In particular,

- they lead to the same set of open (resp. closed, dense, bounded, ...) sets;
- if a sequence converges for one of the norms, it converges for the other one, and the limits are the same.

Theorem. All the norms of a finite dimensional vector space are equivalent.

Example of non-equivalent norms: let $V = C^0([0,1],\mathbb{K})$ (dim$(V) = \infty$) and $v_n(x) = \sqrt{2n+1}\,x^n \in V$. The sequence $(v_n)_{n\in\mathbb{N}}$

- converges to 0 for any L^p-norm with $1 \leq p < 2$;
- is bounded, but does not converge, for the L^2-norm ($\|v_n\|_{L^2} = 1$ for all $n \in \mathbb{N}$);
- goes to infinity in any L^p-norm with $2 < p \leq \infty$.

Let V and W be two normed $\mathbb{K}$-vector spaces.

Definition (linear maps). **A mapping $A : V \to W$ is called linear if**

$$\forall(\lambda, \mu, u, v) \in \mathbb{K} \times \mathbb{K} \times V \times V, \quad A(\lambda u + \mu v) = \lambda A u + \mu A v.$$

Definition (kernel and range of a linear map). **The vector subspaces**

$$\textbf{Ker}(A) = \{v \in V \mid Av = 0\} \quad \textbf{and} \quad \textbf{Ran}(A) = \{w \in W \mid \exists v \in V \text{ s.t. } Av = w\}$$

are respectively called the kernel and the range (or the image) of A.

Theorem. **A linear map $A : V \to W$ is continuous if and only if**

$$\|A\|_{\mathcal{B}(V;W)} := \sup_{v \in V \setminus \{0\}} \frac{\|Av\|_W}{\|v\|_V} < \infty.$$

A continuous linear map therefore maps bounded sets of V into bounded sets of W. For this reason, it is often called a bounded linear operator. The set $\mathcal{B}(V;W)$ of bounded linear operators from V to W is a vector space and $\|\cdot\|_{\mathcal{B}(V;W)}$ is a norm on $\mathcal{B}(V;W)$.

Definition (dual of a normed vector space). **A linear map $L : V \to \mathbb{K}$ is called a linear form. The set of continuous linear forms on V is a normed vector space called the (topological) dual of V and denoted by V'.**

In infinite dimensional vector spaces, not all linear maps are continuous!

Example: **let $V = C^0([0,1],\mathbb{K})$ and $L : V \to \mathbb{K}$ be defined by $Lv = v(0)$.**

- **the linear form L is continuous, and is therefore an element of V', if V is endowed with the norm $\|\cdot\|_{C^0}$ since**

$$|Lv| = |v(0)| \leq \|v\|_{C^0};$$

- **on the other hand, it is not continuous if V is endowed with the L^1-norm. Consider for instance the sequence $(v_n)_{n\in\mathbb{N}}$ of elements of C^0 defined by $v_n(x) = e^{-nx}$ and $v(x) = 0$. We have**

$$\|v_n - v\|_{L^1} = \frac{1-e^{-n}}{n} \underset{n\to\infty}{\longrightarrow} 0 \qquad \textbf{while} \qquad Lv_n = 1 \textbf{ and } Lv = 0.$$

Banach spaces

Definition (Banach space). A complete normed vector space is called a Banach space.

Examples:

- Any finite-dimensional vector space, endowed with any norm, is a Banach space.
- If K is a closed bounded subset of $\mathbb{R}^d$ (that is a compact subset of $\mathbb{R}^d$), then $C^0(X, \mathbb{K})$, endowed with the norm defined by

$$\|u\|_{C^0} = \max_{\mathbf{x} \in X} |u(\mathbf{x})|$$

  is a Banach space.

Theorem. If V and W are two normed vector spaces, and if W is a Banach space, then $\mathcal{B}(V, W)$, endowed with the norm $\| \cdot \|_{\mathcal{B}(V,W)}$, is a Banach space.

In particular, the dual V' of a normed vector space V is always a Banach space.

L^p spaces

- **For any $1 \leq p < \infty$, the mapping**

$$v \mapsto \|v\|_{L^p} := \left(\int_{\mathbb{R}^d} |v|^p \right)^{1/p}$$

defines a norm on

$$C^0_{\mathrm{c}}(\mathbb{R}^d, \mathbb{K}) := \left\{ v \in C^0(\mathbb{R}^d; \mathbb{K}) \mid \ v = 0 \text{ outside some bounded set} \right\},$$

but $C^0_{\mathrm{c}}(\mathbb{R}^d, \mathbb{K})$, endowed with the norm $\| \cdot \|_{L^p}$, is not complete.

- **Its completed space is the vector space**

$$L^p(\mathbb{R}^d, \mathbb{K}) := \left\{ u : \mathbb{R}^d \to \mathbb{K} \mid \int_{\mathbb{R}^d} |u|^p < \infty \right\},$$

which, endowed with the norm $\| \cdot \|_{L^p}$, is a Banach space.

- **Technical details:**
  - **one must use the Lebesgue integral (doesn't work with Riemann integral);**
  - **the elements of $L^p(\mathbb{R}^d)$ are in fact equivalence classes of measurable functions (for the Lebesgue measure) for the equivalence relation $u \sim v$ iff $u = v$ everywhere except possibly on a set of Lebesgue measure equal to zero.**

3 - Hilbert spaces

"Each quantum system is associated with a separable complex Hilbert space $\mathcal{H}$."

A separable complex Hilbert space is a set $\mathcal{H}$

- **endowed with a $\mathbb{C}$-vector space structure;**
- **endowed with a scalar product $\langle \cdot | \cdot \rangle$;**
- **complete for the norm $\| \cdot \|$ associated with the scalar product $\langle \cdot | \cdot \rangle$;**
- **admitting a countable dense subset.**

In this section, $\mathbb{K}$ denotes $\mathbb{R}$ or $\mathbb{C}$.

Definition (scalar product on real vector spaces). **Let V be an $\mathbb{R}$-vector space. A scalar product on V is a symmetric, positive definite, bilinear form on $V \times V$.**

In other words, a scalar product on V is a mapping

$$\begin{aligned} V \times V &\to \mathbb{R} \\ (v, w) &\mapsto (v, w)_V \end{aligned}$$

satisfying the following properties:

- **bilinearity: for all** $(\lambda, v, v', w, w') \in \mathbb{R} \times V \times V \times V \times V$**,**

$$\begin{aligned} (v + v', w)_V &= (v, w)_V + (v', w)_V, \quad (\lambda v, w)_V = \lambda(v, w)_V, \\ (v, w + w')_V &= (v, w)_V + (v, w')_V, \quad (v, \lambda w)_V = \lambda(v, w)_V; \end{aligned}$$

- **symmetry: for all** $(v, w) \in \times V \times V,\ (w, v)_V = (v, w)_V$**;**
- **positive definiteness:** $\forall v \in V,\ (v, v)_V \geq 0$ **and** $(v, v)_V = 0 \Leftrightarrow v = 0$**.**

Theorem (norm associated with a scalar product). **The map** $\|\cdot\|_V : V \to \mathbb{R}_+$ **defined by** $\|v\|_V = (v, v)_V^{1/2}$ **is a norm on V and it holds**

$$\forall (v, w) \in V, \quad |(v, w)_V| \leq \|v\|_V \, \|w\|_V \qquad \textbf{(Cauchy-Schwarz inequality)}.$$

Definition (scalar product on complex vector spaces). **Let V be a $\mathbb{C}$-vector space. A scalar product on V is a skew-symmetric, positive definite, sesquilinear form on $V \times V$.**

In other words, a scalar product on V is a mapping

$$\begin{aligned} V \times V &\to \mathbb{R} \\ (v, w) &\mapsto (v, w)_V \end{aligned}$$

satisfying the following properties:

- **sesquilinearity: for all** $(\lambda, v, v', w, w') \in \mathbb{C} \times V \times V \times V \times V$**,**

$$\begin{aligned} (v + v', w)_V &= (v, w)_V + (v', w)_V, \quad &(\lambda v, w)_V = \overline{\lambda}(v, w)_V, \\ (v, w + w')_V &= (v, w)_V + (v, w')_V, \quad &(v, \lambda w)_V = \lambda(v, w)_V; \end{aligned}$$

- **skew-symmetry: for all** $(v, w) \in \times V \times V,\ (w, v)_V = \overline{(v, w)_V}$**;**
- **positive definiteness:** $\forall v \in V,\ (v, v)_V \geq 0$ **and** $(v, v)_V = 0 \Leftrightarrow v = 0$**.**

Theorem (norm associated with a scalar product). **The map** $\|\cdot\|_V : V \to \mathbb{R}_+$ **defined by** $\|v\|_V = (v, v)_V^{1/2}$ **is a norm on** V **and it holds**

$$\forall (v, w) \in V, \quad |(v, w)_V| \leq \|v\|_V \, \|w\|_V \qquad \textbf{(Cauchy-Schwarz inequality)}.$$

Definition (Hilbert space). **A Hilbert space is a vector space V endowed with a scalar product $(\cdot,\cdot)_V$ and complete for the associated norm $\|\cdot\|_V$.**

Example: **all finite dimensional $\mathbb{K}$-vector spaces equipped with a scalar product are Hilbert spaces.**

- **Endowed with the Euclidean scalar product, $\mathbb{R}^n$ is a Hilbert space:**

$$(\mathbf{x},\mathbf{y})_2 = \sum_{i=1}^{n} x_i y_i, \qquad \|\mathbf{x}\|_2 = (\mathbf{x},\mathbf{x})_2^{1/2} = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}.$$

- **Let $S \in \mathbb{R}^{n\times n}$ be a positive definite symmetric matrix ($S_{ji} = S_{ij}$ for all $1 \leq i,j \leq n$ and $\mathbf{x}^T S\mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n \setminus \{0\}$).**

  Then $(\mathbf{x},\mathbf{y})_S = \mathbf{x}^T S\mathbf{y}$ defines a scalar product on $\mathbb{R}^n$ and

$$\forall \mathbf{x} \in \mathbb{R}^n, \quad \lambda_1(S)\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_S \leq \lambda_d(S)\|\mathbf{x}\|_2,$$

  where $\lambda_1(S) \leq \lambda_2(S) \leq \cdots \leq \lambda_n(S)$ are the eigenvalues of S.

Bra-ket notation and Riesz representation theorem

In QM, vectors of the Hilbert space $\mathcal{H}$ are denoted by kets: $|\psi\rangle$

To each ket $|\psi\rangle \in \mathcal{H}$ is associated the bra $\langle\psi| \in \mathcal{H}'$ defined by

$$\forall|\phi\rangle \in \mathcal{H}, \quad \langle\psi|(|\phi\rangle) := \langle\psi|\phi\rangle.$$

Theorem (Riesz representation theorem). The antilinear map

$$\begin{array}{rcl} \mathcal{H} & \rightarrow & \mathcal{H}' \\ |\psi\rangle & \mapsto & \langle\psi| \end{array}$$

is antiunitary (it is a bijective isometry), and therefore allows one to identify $\mathcal{H}$ (the set of "kets") with $\mathcal{H}'$ (the set of "bras").

Definition-Theorem (adjoint of a bounded linear operator). Let V and W be Hilbert spaces and $A \in \mathcal{B}(V;W)$. The map $A^* \in \mathcal{B}(W;V)$ defined by

$$\forall(v,w) \in V \times W, \quad \langle v|A^*w\rangle_V = \langle Av|w\rangle_W,$$

is called the adjoint of A. We have Ker$(A^*) =$ Ran$(A)^\perp$ and $\overline{\text{Ran}(A^*)} =$ Ker$(A)^\perp$.

Separability and orthonormal bases

Let $\mathcal{H}$ be a Hilbert space.

Definition (separable Hilbert space). **$\mathcal{H}$ is called separable if it admits a countable dense subset.**

Definition (orthonormal basis). **A family $(e_n)_{n\in\mathcal{N}}$ with $\mathcal{N}$ finite or countable is called an orthonormal basis of $\mathcal{H}$ if**

- $\langle e_n | e_m \rangle = \delta_{mn}$ **for all** $m, n \in \mathcal{N}$**;**
- **for all** $v \in \mathcal{H}$**, we have** $v = \sum_{n\in\mathcal{N}} \langle e_n | v \rangle \, e_n$ **and** $\|v\|^2 = \sum_{n\in\mathcal{N}} |\langle e_n | v \rangle|^2$ **(Parseval).**

Theorem. **$\mathcal{H}$ is separable if and only if it admits an orthonormal basis.**

Example: **the space $L^2(\mathbb{R}^d, \mathbb{K})$.**

- **The map**

$$(u, v) \mapsto (u, v)_{L^2} := \int_{\mathbb{R}^d} \overline{u(\mathbf{x})}\, v(\mathbf{x})\, d\mathbf{x}$$

  defines a scalar product on

$$C_c^\infty(\mathbb{R}^d, \mathbb{K}) := \left\{ v \in C^\infty(\mathbb{R}^d, \mathbb{K}) \mid \; v = 0 \text{ outside some bounded set} \right\},$$

  but $C_c^\infty(\mathbb{R}^d, \mathbb{K})$, endowed with the scalar product $(\cdot, \cdot)_{L^2}$, is not complete.

- **Its completed space is the vector space**

$$L^2(\mathbb{R}^d, \mathbb{K}) := \left\{ u : \mathbb{R}^d \to \mathbb{K} \mid \int_{\mathbb{R}^d} |u|^2 < \infty \right\}.$$

  Endowed with the scalar product $(u, v)_{L^2}$, it is a separable Hilbert space.

- **Technical details:**
  - **one must use the Lebesgue integral (doesn't work with Riemann integral);**
  - **the elements of $L^2(\mathbb{R}^d, \mathbb{K})$ are in fact equivalence classes of measurable functions (for the Lebesgue measure) for the equivalence relation $u \sim v$ iff $u = v$ everywhere except possibly on a set of Lebesgue measure equal to zero.**

Example: **the Sobolev spaces $H^1(\mathbb{R}^d, \mathbb{K})$ and $H^2(\mathbb{R}^d, \mathbb{K})$.**

- **The sets**

$$H^1(\mathbb{R}^d, \mathbb{K}) := \left\{u \in L^2(\mathbb{R}^d, \mathbb{K}) \mid \nabla u \in (L^2(\mathbb{R}^d, \mathbb{K}))^d\right\},$$
$$H^2(\mathbb{R}^d, \mathbb{K}) := \left\{u \in L^2(\mathbb{R}^d, \mathbb{K}) \mid \nabla u \in (L^2(\mathbb{R}^d), \mathbb{K})^d \text{ and } D^2 u \in (L^2(\mathbb{R}^d), \mathbb{K})^{d\times d}\right\}$$

are vector spaces. Respectively endowed with the scalar products

$$(u,v)_{H^1} := \int_{\mathbb{R}^d} \overline{u} v + \int_{\mathbb{R}^d} \overline{\nabla u} \cdot \nabla v,$$
$$(u,v)_{H^2} := \int_{\mathbb{R}^d} \overline{u} v + \int_{\mathbb{R}^d} \overline{\nabla u} \cdot \nabla v + \int_{\mathbb{R}^d} \overline{D^2 u} : D^2 v,$$

they are separable Hilbert spaces.

- **Technical detail: the gradient and the second derivatives are defined by means of distribution theory.**

Remark. **Let $u \in H^1(\mathbb{R}^d, \mathbb{K})$. A function $\widetilde{u} \in H^1(\mathbb{R}^d, \mathbb{K})$ can be a very accurate approximation of u in L^2 and a terrible approximation of u in H^1.**

For instance, let $u(x) = \frac{1}{1+x^2}$ and $u_n(x) = \left(1 + \frac{\sin(n^2 x^2)}{n}\right) u(x)$. The sequence $(u_n)_{n\in\mathbb{N}^*}$ converges to u in $L^2(\mathbb{R}, \mathbb{R})$ and goes to infinity in $H^1(\mathbb{R}, \mathbb{R})$.

Theorem (orthogonal projection on closed vector subspaces). **Let F be a closed vector subspace of a Hilbert space V. Then,**

1. **for any $v \in V$, there exists a unique point of F, denoted by $\Pi_F(v)$ and called the orthogonal projection of v on F such that**

$$\|v - \Pi_F(v)\|_V = \min_{w \in F} \|v - w\|_V;$$

2. **the vector $\Pi_F(v)$ is characterized by**

$$\Pi_F(v) \in F \quad \textbf{and} \quad v - \Pi_F(v) \in F^\perp := \{w \in V \mid (w, z)_V = 0 \text{ for all } z \in F\};$$

3. **the mapping $v \mapsto \Pi_F(v)$ is a linear orthogonal projector, i.e. $\Pi_F : V \to V$ is a bounded linear operator on V satisfying**

$$\Pi_F^2 = \Pi_F = \Pi_F^*.$$

Besides, Ran$(\Pi_F) = F$ and Ker$(\Pi_F) = F^\perp$.

Application: **best approximation of a function $f \in H^2(\mathbb{R}^3)$ by a gaussian-polynomial of total degree $\leq n$ centered at zero, for the L^2, H^1 and H^2 norms.**

4 - Differential calculus in Hilbert spaces

For brevity, we will limit ourselves to the setting of real Hilbert spaces.

All the results presented here can be extended to complex Hilbert spaces.

All these results, except the ones involving the notion of gradient, can be extended to normed $\mathbb{K}$-vector spaces.

Definition (derivative of a function at a point). Let V, W be Hilbert spaces, $F : V \to W$, and $v \in V$. The function F is called differentiable at v, if there exists a continuous linear map $d_vF : V \to W$ such that in the vicinity of v,

$$F(v+h) = F(v) + d_vF(h) + o(h),$$

which means

$$\forall \varepsilon > 0,\ \exists \eta > 0 \text{ s.t. } \forall h \in V \text{ s.t. } \|h\|_V \leq \eta, \quad \|F(v+h) - F(v) - d_vF(h)\|_W \leq \varepsilon \|h\|_V.$$

If such a linear map d_vF exists, it is unique. It is called the derivative of F at v.

Definition (differentiable and C^1 functions). F is called differentiable if F is differentiable at each point of V. In this case, the mapping

$$\begin{aligned} dF : V &\longrightarrow \mathcal{B}(V;W) \\ v &\mapsto d_vF \end{aligned}$$

is called the derivative of F. F is called of class C^1 on V if dF is continuous.

Remark. One can similarly define the derivative of a function $F : U \to W$, where U is an open subset of V (that is $U = V \setminus F$, where F is a closed subset of V).

Theorem (Riesz). **Let V be a Hilbert space and $l : V \to \mathbb{R}$ a continuous linear map. Then there exists a unique vector $w \in V$ such that**

$$\forall v \in V, \quad l(v) = (w, v)_V.$$

Definition (gradient). **Let V be a Hilbert space endowed with the scalar product $(\cdot, \cdot)_V$, U an open subset of V and $E : U \to \mathbb{R}$ a function differentiable at $v \in U$.**

The unique vector of V denoted by $\nabla E(v)$ and defined by

$$\forall h \in V, \qquad d_v E(h) = (\nabla E(v), h)_V \qquad \text{(by means of Riesz theorem)}$$

is called the gradient of E at v.

Gradient of a function $E : \mathbb{R}^d \to \mathbb{R}$

The above abstract definition of the gradient agrees with the usual one when $V = \mathbb{R}^d$ endowed with the Euclidean scalar product:

$$\forall \mathbf{h} \in \mathbb{R}^d, \quad E(\mathbf{x}+\mathbf{h}) = E(\mathbf{x}) + \sum_{i=1}^{d} \frac{\partial E}{\partial x_i}(\mathbf{x})\, h_i + o(\mathbf{h}) = E(\mathbf{x}) + \nabla E(\mathbf{x}) \cdot \mathbf{h} + o(\mathbf{h})$$

$$\textbf{with} \qquad \nabla E(\mathbf{x}) = \begin{pmatrix} \frac{\partial E}{\partial x_1}(\mathbf{x}) \\ \cdot \\ \cdot \\ \cdot \\ \frac{\partial E}{\partial x_d}(\mathbf{x}) \end{pmatrix}.$$

If $\mathbb{R}^d$ is endowed with the scalar product $(\mathbf{x}, \mathbf{y})_S := \mathbf{x}^T S \mathbf{y}$, where $S \in \mathbb{R}^{d \times d}$ is a positive definite symmetric matrix, then the gradient of E, which we will denote by $\nabla_S E(\mathbf{x})$, is related to the Euclidean gradient $\nabla E(\mathbf{x})$ by

$$\nabla_S E(\mathbf{x}) = S^{-1} \nabla E(\mathbf{x}).$$

Geometric interpretation of the gradient

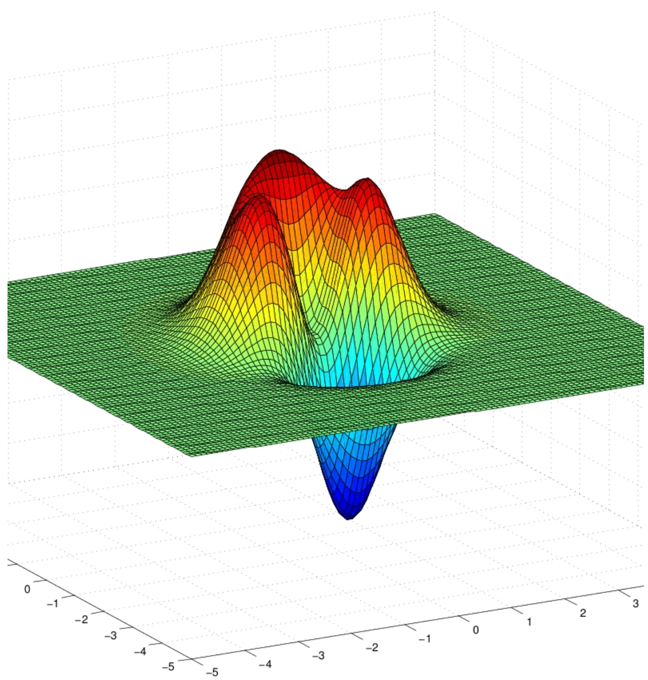
Let $E : V \to \mathbb{R}$ of class C^1, $v \in V$ and $\alpha = E(v)$. If $\nabla E(v) \neq 0$, then

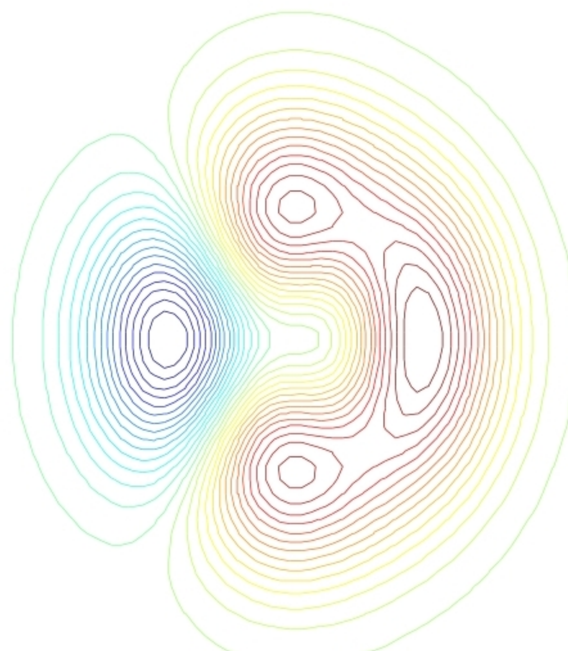
- **in the vicinity of v, the level set**

$$\mathcal{C}_\alpha := \{w \in V \mid \; E(w) = \alpha\}$$

  is a C^1 hypersurface (a codimension 1 C^1 manifold);

- **the vector $\nabla E(v)$ is orthogonal to the affine hyperplane tangent to $\mathcal{C}_\alpha$ at v and points toward the steepest ascent direction.**





Example (Yukawa functional): $V = H^1(\mathbb{R}^d) = \left\{ v \in L^2(\mathbb{R}^d) \mid \nabla v \in (L^2(\mathbb{R}^d))^d \right\}$

$$E(v) = \frac{1}{2}\int_{\mathbb{R}^d} |\nabla v|^2 + \frac{1}{2}\int_{\mathbb{R}^d} v^2 - \int_{\mathbb{R}^d} fv, \quad f \in L^2(\mathbb{R}^d) \text{ given.}$$

Let $v \in V$ **and** $h \in V$**. We have**

$$\begin{aligned} E(v+h) &= \frac{1}{2}\int_{\mathbb{R}^d} |\nabla(v+h)|^2 + \int_{\mathbb{R}^d} (v+h)^2 - \int_{\mathbb{R}^d} f(v+h) \\ &= \frac{1}{2}\int_{\mathbb{R}^d} |\nabla v|^2 + \int_{\mathbb{R}^d} \nabla v \cdot \nabla h + \frac{1}{2}\int_{\mathbb{R}^d} |\nabla h|^2 + \frac{1}{2}\int_{\mathbb{R}^d} v^2 + \int_{\mathbb{R}^d} vh + \frac{1}{2}\int_{\mathbb{R}^d} h^2 - \int_{\mathbb{R}^d} fv - \int_{\mathbb{R}^d} fh \\ &= E(v) + \int_{\mathbb{R}^d} \nabla v \cdot \nabla h + \int_{\mathbb{R}^d} vh - \int_{\mathbb{R}^d} fh + \frac{1}{2}\int_{\mathbb{R}^d} |\nabla h|^2 + \frac{1}{2}\int_{\mathbb{R}^d} h^2 \end{aligned}$$

with

$$\left| \int_{\mathbb{R}^d} \nabla v \cdot \nabla h + \int_{\mathbb{R}^d} vh - \int_{\mathbb{R}^d} fh \right| \le C_{v,f} \|h\|_{H^1}, \qquad \left| \frac{1}{2}\int_{\mathbb{R}^d} |\nabla h|^2 + \frac{1}{2}\int_{\mathbb{R}^d} h^2 \right| = \frac{1}{2}\|h\|_{H^1}^2 = o(h).$$

Therefore, E **is differentiable at** v **and**

$$\forall h \in V, \quad d_v E(h) = \int_{\mathbb{R}^d} \nabla v \cdot \nabla h + \int_{\mathbb{R}^d} vh - \int_{\mathbb{R}^d} fh.$$

Example (Yukawa functional - continued)

The gradient of E at v therefore is the function $w \in H^1(\mathbb{R}^d)$ characterized by

$$\forall h \in V = H^1(\mathbb{R}^3), \quad (w,h)_{H^1} = d_vE(h) = \int_{\mathbb{R}^3} \nabla v \cdot \nabla h + \int_{\mathbb{R}^3} vh - \int_{\mathbb{R}^d} fh.$$

To compute $w = \nabla E(v)$, we have to solve the linear elliptic problem

$$\begin{cases} \textbf{seek } w \in V \textbf{ such that} \\ \forall h \in V, \quad a(w,h) = L(h) \end{cases}$$

with

$$a(w,h) = \int_{\mathbb{R}^3} \nabla w \cdot \nabla h + \int_{\mathbb{R}^3} wh \qquad \textbf{and} \qquad L(h) = \int_{\mathbb{R}^3} \nabla v \cdot \nabla h + \int_{\mathbb{R}^3} vh - \int_{\mathbb{R}^3} fh,$$

or equivalently the PDE

$$\textbf{seek } w \in H^1(\mathbb{R}^3) \textbf{ such that } -\Delta w + w = -\Delta v + v - f.$$

Therefore

$$[\nabla E(v)](\mathbf{x}) = v(\mathbf{x}) - \int_{\mathbb{R}^3} \frac{e^{-|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|} f(\mathbf{y})\, d\mathbf{y}.$$

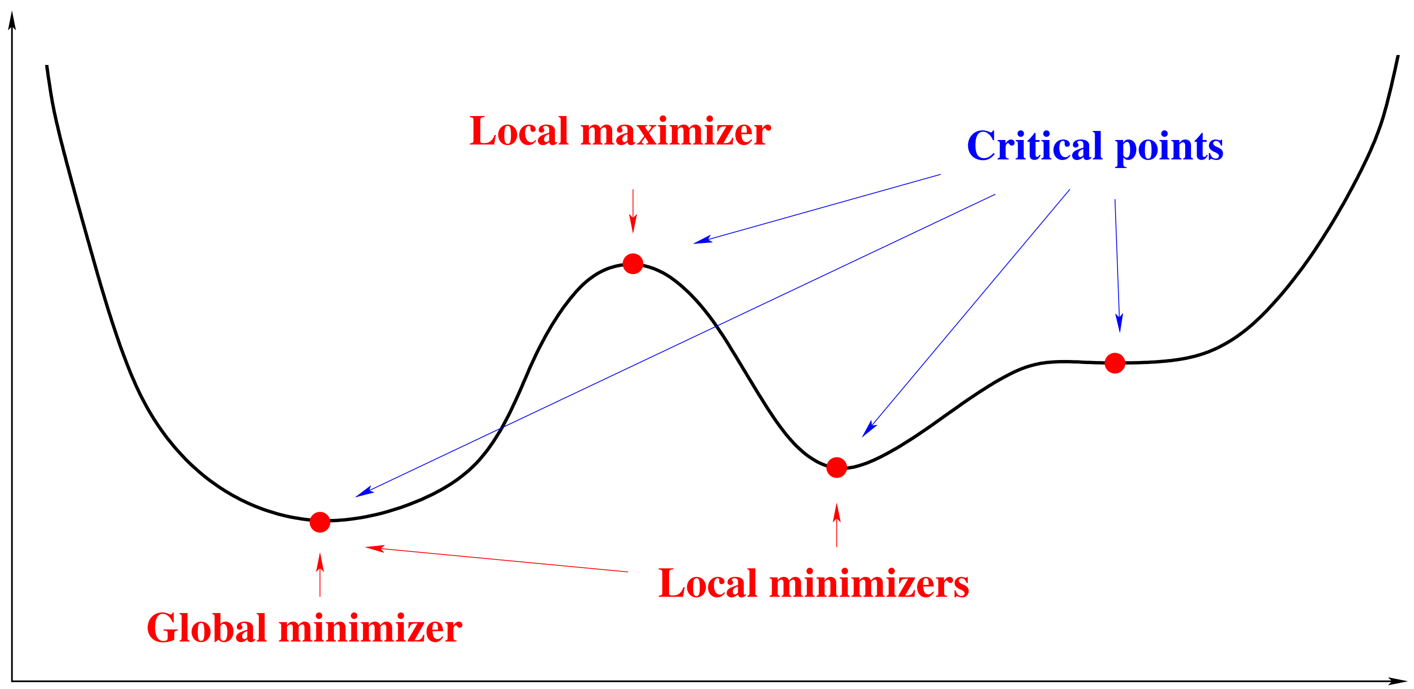
5 - Optimization in Hilbert spaces

For brevity, we will limit ourselves to the setting of real Hilbert spaces.

All the results presented here can be extended to complex Hilbert spaces.

A simple result: 1D unconstrained optimization

If $E : \mathbb{R} \to \mathbb{R}$ is differentiable, the set of the local minimizers of E is included in the set $\mathcal{C} = \{x \in \mathbb{R} \mid E'(x) = 0\}$ of the critical points of E.



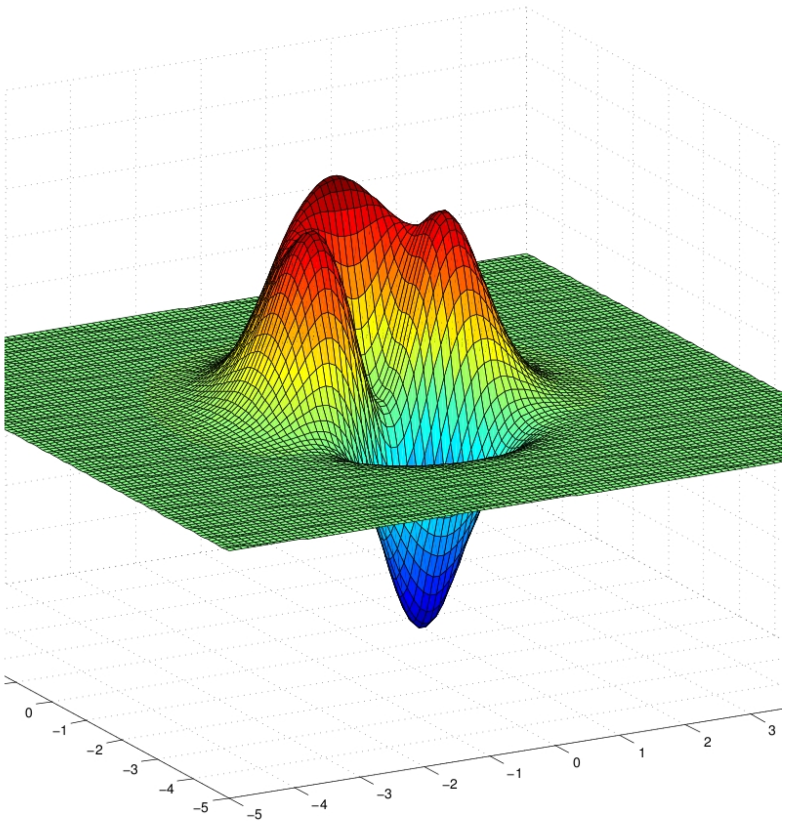
Goal: **extend this result to unconstrained and constrained optimization in finite or infinite dimensional Hilbert spaces.**

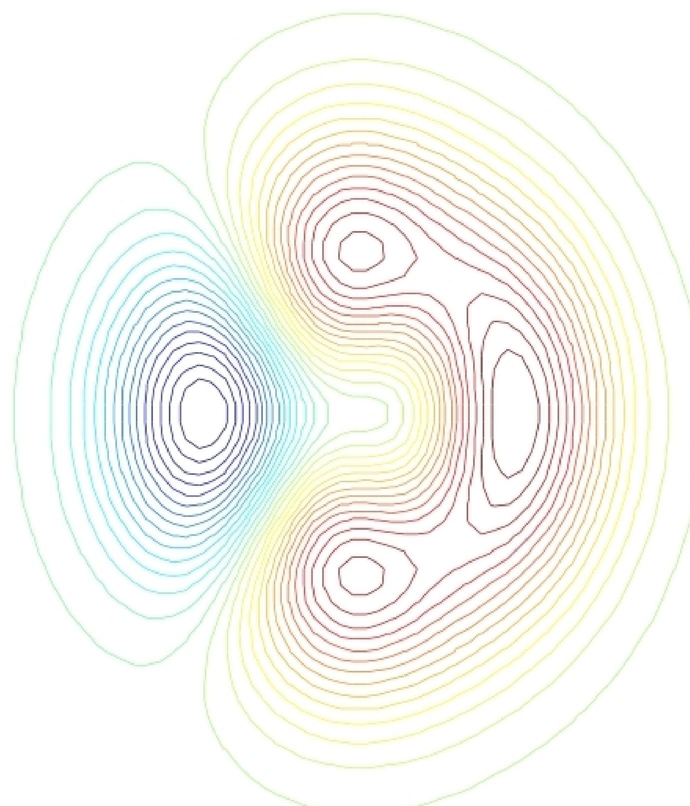
Unconstrained optimization in Hilbert spaces

Theorem. **Let V be a Hilbert space and $E : V \to \mathbb{R}$ a differentiable function. The set of the local minima of E is included in the set**

$$\mathcal{C} = \{v \in V \mid d_v E = 0\} = \{v \in V \mid \nabla E(v) = 0\}$$

of the critical points of E.

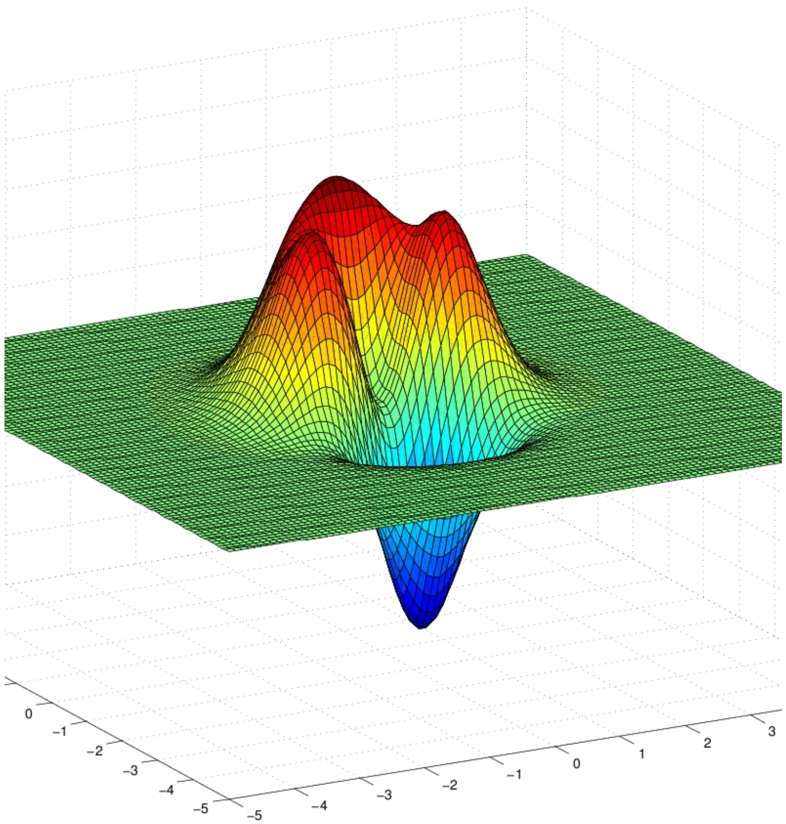


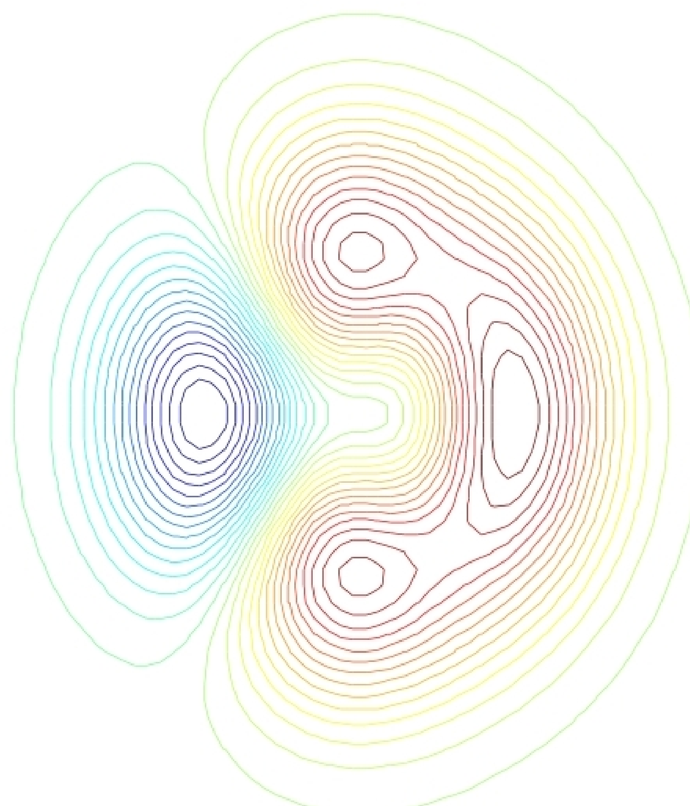


Example: $V = \mathbb{R}^2$ **(endowed with the Euclidean scalar product)**

$$E(x_1, x_2) = (x_1^3 + x_2^2)\ \exp(-(x_1^2 + x_2^2))$$

$$\nabla E(x) = \begin{pmatrix} x_1\left(3x_1 - 2x_1^3 - 2x_2^2\right) e^{-(x_1^2+x_2^2)} \\ 2x_2\left(1 - x_1^3 - x_2^2\right) e^{-(x_1^2+x_2^2)} \end{pmatrix} = 0 \quad \Leftrightarrow \quad (x_1, x_2) = \left| \begin{array}{l} (0,0) \\ (\pm\sqrt{3/2}, 0) \\ (0, \pm 1) \\ (2/3, \pm\sqrt{19/27}) \end{array} \right.$$





Example: back to the Yukawa functional

$$\inf_{v\in H^1(\mathbb{R}^3)} E(v) \quad \textbf{where} \quad E(v)=\frac{1}{2}\int_{\mathbb{R}^d}|\nabla v|^2+\frac{1}{2}\int_{\mathbb{R}^d}v^2-\int_{\mathbb{R}^d}fv, \quad f\in L^2(\mathbb{R}^d) \textbf{ given.}$$

Since $[\nabla E(v)](\mathbf{x})=v(\mathbf{x})-\int_{\mathbb{R}^3}\frac{e^{-|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}f(\mathbf{y})\,d\mathbf{y}$**, the unique critical point of** E **in** $H^1(\mathbb{R}^3)$ **is**

$$u(\mathbf{x})=\int_{\mathbb{R}^3}\frac{e^{-|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}f(\mathbf{y})\,d\mathbf{y},$$

which is a solution to the Yukawa equation $-\Delta u+u=f$ **on** $\mathbb{R}^3$**.**

u **is in fact the unique global minimizer of** E **on** $H^1(\mathbb{R}^3)$ **(strict convexity argument), and the unique tempered distribution solution to** $-\Delta u+u=f$**.**

Remark. The 3D Poisson equation $-\Delta u=f$ **is more complicated to handle:**

- **for** $f\in L^2(\mathbb{R}^3)$**, it may have no solution even in the sense of distributions;**
- **if it has one solution, it has infinitely many;**
- **for e.g.** $f\in C_c^\infty(\mathbb{R}^3)$**, the physical solution is the unique minimizer of**

$$E(v):=\frac{1}{2}\int_{\mathbb{R}^3}|\nabla v|^2-\int_{\mathbb{R}^3}fv \quad \textbf{on} \quad W^1(\mathbb{R}^3):=\left\{v\,\middle|\;\nabla v\in(L^2(\mathbb{R}^3))^3,\ \frac{v(\mathbf{x})}{(1+|\mathbf{x}|^2)^{1/2}}\in L^2(\mathbb{R}^3)\right\}.$$

Equality constrained optimization. **Let V and W be Hilbert spaces s.t. dim$(W) < \infty$, $E : V \to \mathbb{R}$, $g : V \to W$. Consider the optimization problem**

$$\inf_{v \in K} E(v) \quad \textbf{where} \quad K = \{v \in V \mid g(v) = 0\}\,.$$

Definition (qualification of the constraints). **The equality constraints $g = 0$ are called qualified at $u \in K$ if $d_u g : V \to W$ is surjective (i.e. Ran$(d_u g) = W$).**

Theorem (Euler-Lagrange theorem). **Let $u \in K$ be a local minimum of E on**

$$K = \{v \in V \mid g(v) = 0\}\,.$$

Assume that

1. **E is differentiable at u and g is C^1 in the vicinity of u;**
2. **the equality constraint $g(v) = 0$ is qualified at u.**

Then, there exists a unique $\lambda \in W$ such that

$$\forall h \in V,\ d_u E(h) = (\lambda, d_u g(h))_W \quad \textbf{or equivalently} \quad \nabla E(u) = d_u g^*(\lambda),$$

where $d_u g^*$ is the adjoint of $d_u g$. λ is called the Lagrange multiplier of the constraint $g = 0$.

Euler-Lagrange equations

Assume that the constraints are qualified at any point of K. Then solving

$$\begin{cases} \textbf{seek } (u,\lambda) \in V \times W \textbf{ such that} \\ \nabla E(u) - d_u g^*(\lambda) = 0 \\ g(u) = 0 \end{cases} \tag{1}$$

allows one to find all the critical points (among which the local minimizers and the local maximizers) of E on K.

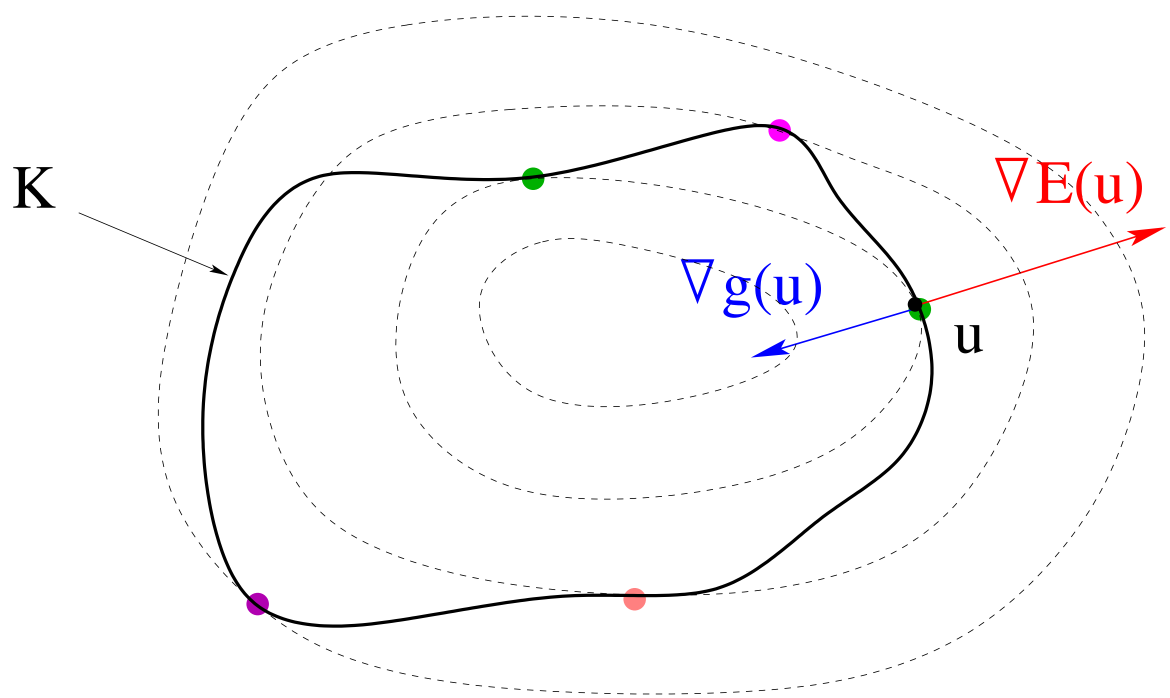
The solutions of the Euler-Lagrange equations (1) are called the critical points of E on K.

Remark : **if dim$(V) = d < \infty$ and dim$(W) = m < \infty$, then the above problem consists of $(d+m)$ scalar equations with $(d+m)$ scalar unknowns.**

Remark. **Equations (1) are equivalent to seeking $(u,\lambda) \in V \times W$ such that**

$$\frac{\partial L}{\partial v}(u,\lambda) = 0, \ \frac{\partial L}{\partial \mu}(u,\lambda) = 0, \quad \textbf{where} \quad L(v,\mu) := E(v) - (\mu, g(v))_W \quad \textbf{(Lagrangian)}.$$

A simple 2D example



On $K = g^{-1}(0) = \{v \in V \mid g(v) = 0\}$**, the function** E **possesses**

- **two local minimizers, all global**
- **two local maximizers, among which the global maximizer**
- **one critical point which is neither a local minimizer not a local maximizer.**

Sketch of the proof

- **Let u be a local minimizer of E on $K = g^{-1}(0) = \{v \in V \mid g(v) = 0\}$ and $\alpha = E(u)$.**

- **If the constraint $g = 0$ is qualified at u (i.e. if $d_u g : \mathcal{H} \to \mathcal{K}$ is surjective), then, in the vicinity of u, K is a C^1 manifold with tangent space**

$$T_u K = \{h \in \mathcal{H} \mid d_u g(h) = 0\} = \textbf{Ker}(d_u g).$$

- **Since u is a minimizer of E on K, the vector $\nabla E(u)$ must be orthogonal to $T_u K$. Indeed, for any $h \in T_u K$, there exists a C^1 curve $\phi : [-1, 1] \to V$ drawn on K such that $\phi(0) = u$ et $\phi'(0) = h$, and we have**

$$0 \leq E(\phi(t)) - E(u) = E(u + th + o(t)) - E(u) = t\nabla E(u) \cdot h + o(t).$$

- **We have**

$$\nabla E(u) \in (T_u K)^\perp = (\textbf{Ker}(d_u g))^\perp = \overline{\textbf{Ran}(d_u g^*)} = \textbf{Ran}(d_u g^*) \text{ \textbf{since dim}}(W) < \infty.$$

- **Therefore, there exists $\lambda \in W$ such that $\nabla E(u) = d_u g^*(\lambda)$.**

Most often, Lagrange multipliers have a "physical" interpretation

- **statistical mechanics, the equilibrium state of a chemical system interacting with its environment is obtained by maximizing the entropy under the constraints that the energy, the volume and the concentration of chemical species are given on average:**

  $\rightarrow$ the Lagrange multipliers are respectively $1/T$, P/T and μ_i/T (T: temperature, P: pressure, μ_i chemical potential of species i)

- **fluid mechanics, the admissible dynamics of an incompressible fluid are the critical points of the action under the constraint that the density of the fluid remains constant ($\mathrm{div}\,(u) = 0$)**

  $\rightarrow$ the Lagrange multiplier of the incompressibility constraint is the pressure field.

Analytical derivatives

$$\forall \mathbf{x} \in \mathbb{R}^d, \quad W(\mathbf{x}) = \inf\left\{E(\mathbf{x}, v),\ v \in V,\ g(\mathbf{x}, v) = 0\right\} \tag{2}$$

with $E : \mathbb{R} \times V \to \mathbb{R}$**,** $g : \mathbb{R} \times V \to W$**,** $\quad V, W$ **Hilbert spaces, dim**$(W) < \infty$**.**

Assume (2) has a unique minimizer $v(\mathbf{x})$ **and** $\mathbf{x} \mapsto v(\mathbf{x})$ **is regular. Then,**

$$W(\mathbf{x}) = E(\mathbf{x}, v(\mathbf{x})) \quad \Rightarrow \quad \frac{\partial W}{\partial x_i}(\mathbf{x}) = \frac{\partial E}{\partial x_i}(\mathbf{x}, v(\mathbf{x})) + \frac{\partial E}{\partial v}(\mathbf{x}, v(\mathbf{x}))\left(\frac{\partial v}{\partial x_i}(\mathbf{x})\right),$$

$$g(\mathbf{x}, v(\mathbf{x})) = 0 \quad \Rightarrow \quad \frac{\partial g}{\partial x_i}(\mathbf{x}, v(\mathbf{x})) + \frac{\partial g}{\partial v}(\mathbf{x}, v(\mathbf{x}))\left(\frac{\partial v}{\partial x_i}(\mathbf{x})\right) = 0.$$

Euler-Lagrange equation: $\quad \forall h \in V, \quad \frac{\partial E}{\partial v}(\mathbf{x}, v(\mathbf{x}))\,(h) = \left(\frac{\partial g}{\partial v}(\mathbf{x}, v(\mathbf{x}))(h), \lambda(\mathbf{x})\right)_W.$

Therefore $$\frac{\partial W}{\partial x_i}(\mathbf{x}) = \frac{\partial E}{\partial x_i}(\mathbf{x}, v(\mathbf{x})) - \left(\frac{\partial g}{\partial x_i}(\mathbf{x}, v(\mathbf{x})), \lambda(\mathbf{x})\right)_W.$$

Warning: weird things can happen in infinite dimension!

- **in an infinite dimensional Hilbert space V, a function $E : V \to \mathbb{R}$ can be continuous and go to $+\infty$ at infinity, and nevertheless have no minimizer**

$$V = \left\{ v = (v_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid \|v\|_2^2 = \sum_{n\in\mathbb{N}} |v_n|^2 < \infty \right\}, \qquad E(v) = \left(\|v\|_2^2 - 1\right)^2 + \sum_{n\in\mathbb{N}} \frac{v_n^2}{1+n};$$

- **this cannot happen under the additional assumption that E is convex;**
- **if V is not a Hilbert space, but simply a Banach space (that is a complete normed vector space), then this can happen even if E is convex;**

- **in an infinite dimensional Hilbert space V, a critical point v of a C^∞ functional $E : V \to \mathbb{R}$ can be such that $\nabla E(v) = 0$ and $D^2E(v)$ positive definite, and nevertheless not be local minimum of E;**
- **in an infinite dimensional Hilbert space V, a global minimizer u of a functional $E : V \to \mathbb{R}$ twice differentiable at u can be such that $\nabla E(u) = 0$ and $[D^2E(u)](h,h) \geq \|h\|_V^2$, and nevertheless not be a strict global minimizer of E (i.e the unique minimizer of E in the vicinity of u).**

Operator theory for electronic structure calculation

Part II: Introduction to spectral theory

Eric CANCES

Ecole des Pontsand INRIA Paris, France

Mini-school on mathematics in electronic structure theory of GDR CORREL
Paris, January 9-11, 2017

The finite dimensional case ($\mathcal{H} = \mathbb{C}^d$)

The spectrum of a matrix $A \in \mathbb{C}^{d\times d}$ is the finite set

$$\sigma(A) = \left\{ z \in \mathbb{C} \,|\, (zI_d - A) \in \mathbb{C}^{d\times d} \textbf{ non-invertible} \right\} .$$

As $\mathbb{C}^d$ is finite dimensional, $(zI_d - A)$ non-invertible $\Leftrightarrow$ $(zI_d - A)$ non-injective:

$$\sigma(A) = \left\{ z \in \mathbb{C} \,|\, \exists \mathbf{x} \in \mathbb{C}^d \setminus \{0\} \textbf{ s.t. } A\mathbf{x} = z\mathbf{x} \right\} = \{\textbf{eigenvalues of } A\} .$$

A matrix $A \in \mathbb{C}^{d\times d}$ is called hermitian if $A^* = A$ (i.e. $\overline{A_{ij}} = A_{ji}$, $\forall 1 \leq i, j \leq d$).

Key properties of hermitian matrices:

- **the spectrum of a hermitian matrix is real: $\sigma(A) \subset \mathbb{R}$;**
- **any hermitian matrix A can be diagonalized in an orthonormal basis:**

$$A = \sum_{i=1}^{d} \lambda_i \mathbf{x}_i \mathbf{x}_i^*, \quad \lambda_i \in \sigma(A) \subset \mathbb{R}, \quad \mathbf{x}_i \in \mathbb{C}^d, \quad \mathbf{x}_i^* \mathbf{x}_j = \delta_{ij}, \quad A\mathbf{x}_i = \lambda_i \mathbf{x}_i;$$

- **there exists a functional calculus for hermitian matrices.**

Functional calculus for hermitian matrices

Let A be a hermitian matrix of $\mathbb{C}^{d\times d}$ such that

$$A = \sum_{i=1}^{d} \lambda_i \mathbf{x}_i \mathbf{x}_i^*, \quad \lambda_i \in \sigma(A) \subset \mathbb{R}, \quad \mathbf{x}_i \in \mathbb{C}^d, \quad \mathbf{x}_i^* \mathbf{x}_j = \delta_{ij}, \quad A\mathbf{x}_i = \lambda_i \mathbf{x}_i.$$

For any $f : \mathbb{R} \to \mathbb{C}$, the matrix

$$f(A) := \sum_{i=1}^{d} f(\lambda_i) \mathbf{x}_i \mathbf{x}_i^*$$

is independent of the choice of the spectral decomposition of A.

This definition agrees with the usual definition of $f(A)$ for $f(\lambda) = \sum_{k=0}^{n} \alpha_k \lambda^k$:

$$f(A) := \sum_{i=1}^{d} f(\lambda_i) \mathbf{x}_i \mathbf{x}_i^* = \sum_{i=1}^{d} \left(\sum_{k=0}^{n} \alpha_k \lambda_i^k \right) x_i x_i^* = \sum_{k=0}^{n} \alpha_k \left(\sum_{i=1}^{d} \lambda_i^k x_i x_i^* \right) = \sum_{k=0}^{n} \alpha_k A^k.$$

Most of the nice properties of hermitian matrices are also valid for self-adjoint operators in (infinite dimensional) Hilbert spaces.

References:

- E.B. Davies, *Linear operators and their spectra*, Cambridge University Press 2007.
- B. Helffer, *Spectral theory and its applications*, Cambridge University Press 2013.
- M. Reed and B. Simon, *Modern methods in mathematical physics*, Vol. 1, 2nd edition, Academic Press 1980.

Notation: from now on, $\mathcal{H}$ denotes a separable complex Hilbert space, $\langle\cdot|\cdot\rangle$ its scalar product, and $\|\cdot\|$ the associated norm.

A Hilbert space $\mathcal{H}$ is called separable if it admits a countable dense subset.

An infinite dimensional Hilbert space $\mathcal{H}$ is separable if and only if it admits an orthonormal basis $(e_n)_{n\in\mathbb{N}}$, that is

- $\langle e_n|e_m\rangle = \delta_{mn}$ for all $m, n \in \mathbb{N}$;
- for all $v \in \mathcal{H}$, we have $v = \sum_{n\in\mathbb{N}} \langle e_n|v\rangle\, e_n$ and $\|v\|^2 = \sum_{n\in\mathbb{N}} |\langle e_n|v\rangle|^2$ (Parseval).

1 - Linear operators on a Hilbert space - Self-adjointness

Reminder: bounded linear operators

Definition-Theorem (bounded linear operator). A bounded operator on $\mathcal{H}$ is a continuous linear map $A : \mathcal{H} \to \mathcal{H}$, that is a linear map satisfying

$$\|A\| := \sup_{u \in \mathcal{H} \setminus \{0\}} \frac{\|Au\|}{\|u\|} < \infty.$$

The set $\mathcal{B}(\mathcal{H})$ of the bounded operators on $\mathcal{H}$ is a non-commutative algebra and $\| \cdot \|$ is a norm on $\mathcal{B}(\mathcal{H})$.

Definition-Theorem (adjoint of a bounded linear operator). Let $A \in \mathcal{B}(\mathcal{H})$. The operator $A^* \in \mathcal{B}(\mathcal{H})$ defined by

$$\forall (u, v) \in \mathcal{H} \times \mathcal{H}, \quad \langle u | A^* v \rangle = \langle Au | v \rangle,$$

is called the adjoint of A. The operator A is called self-adjoint if $A^* = A$.

(Non necessarily bounded) linear operators on Hilbert spaces

Definition (linear operator). A linear operator on $\mathcal{H}$ is a linear map $A : D(A) \to \mathcal{H}$, where $D(A)$ is a subspace of $\mathcal{H}$ called the domain of A. Note that bounded linear operators are particular linear operators.

Definition (extensions of operators). Let A_1 and A_2 be operators on $\mathcal{H}$. A_2 is called an extension of A_1 if $D(A_1) \subset D(A_2)$ and if $\forall u \in D(A_1)$, $A_2 u = A_1 u$.

Definition (unbounded linear operator). An operator A on $\mathcal{H}$ which does not possess a bounded extension is called an unbounded operator on $\mathcal{H}$.

Definition (symmetric operator). A linear operator A on $\mathcal{H}$ with dense domain $D(A)$ is called symmetric if

$$\forall (u,v) \in D(A) \times D(A), \quad \langle Au|v\rangle = \langle u|Av\rangle.$$

Symmetric operators are not very interesting. Only self-adjoint operators represent physical observables and have nice mathematical properties:

- **real spectrum;**
- **spectral decomposition and functional calculus.**

Definition (adjoint of a linear operator with dense domain). Let A be a linear operator on $\mathcal{H}$ with dense domain $D(A)$, and $D(A^*)$ the vector space defined as

$$D(A^*) = \{v \in \mathcal{H} \mid \exists w_v \in \mathcal{H} \textbf{ s.t. } \forall u \in D(A),\ \langle Au|v\rangle = \langle u|w_v\rangle\}\,.$$

The linear operator A^* on $\mathcal{H}$, with domain $D(A^*)$, defined by

$$\forall v \in D(A^*), \quad A^*v = w_v,$$

(if w_v exists, it is unique since $D(A)$ is dense) is called the adjoint of A. This definition agrees with the one for bounded operators given on Slide 6.

Definition (self-adjoint operator). A linear operator A with dense domain is called self-adjoint if $A^* = A$ (that is if A symmetric <u>and</u> $D(A^*) = D(A)$).

Case of bounded operators:

symmetric $\Leftrightarrow$ self-adjoint.

Case of unbounded operators:

symmetric (easy to check) $\overset{\not\Rightarrow}{\Leftarrow}$ self-adjoint (sometimes difficult to check)

Schrödinger operators commonly encountered in quantum physics

- **Free particle Hamiltonian**

$$\mathcal{H} = L^2(\mathbb{R}^d), \quad D(T) = H^2(\mathbb{R}^d), \quad \forall u \in D(T),\ Tu = -\frac{1}{2}\Delta u.$$

- **Schrödinger operators with confining potential** $V \in C^0(\mathbb{R}^d)$ **s.t.** $V(\mathbf{r}) \underset{|\mathbf{r}|\to+\infty}{\longrightarrow} +\infty$

$$\mathcal{H} = L^2(\mathbb{R}^d), \quad D(H) = \left\{ u \in L^2(\mathbb{R}^d) \mid -\frac{1}{2}\Delta u + Vu \in L^2(\mathbb{R}^d) \right\}$$

$$\forall u \in D(H),\ Hu = -\frac{1}{2}\Delta u + Vu.$$

- **Schrödinger operators with uniformly locally** L^2 **potentials in 3D. Let**

$$V \in L^2_{\rm unif}(\mathbb{R}^3, \mathbb{R}) := \left\{ u : \mathbb{R}^3 \to \mathbb{R} \mid \sup_{\mathbf{r}\in\mathbb{R}^3} \int_{\mathbf{r}+[0,1]^3} |u|^2 < \infty \right\}$$

and

$$\mathcal{H} = L^2(\mathbb{R}^3), \quad D(H) = H^2(\mathbb{R}^3), \quad \forall u \in D(H),\ Hu = -\frac{1}{2}\Delta u + Vu.$$

All these operators are self-adjoint.

2 - Spectrum

Definition-Theorem (spectrum of a closed operator). Let A be a closed[1] linear operator on $\mathcal{H}$.

- **The open set $\rho(A) = \{z \in \mathbb{C} \mid (zI - A) : D(A) \to \mathcal{H}$ invertible$\}$ is called the resolvent set of A.**

- **The closed set $\sigma(A) = \mathbb{C} \setminus \rho(A)$ is called the spectrum of A.**

- **If A is self-adjoint, then $\sigma(A) \subset \mathbb{R}$ and it holds**

$$\sigma(A) = \sigma_{\mathrm{p}}(A) \cup \sigma_{\mathrm{c}}(A),$$

  where $\sigma_{\mathrm{p}}(A)$ and $\sigma_{\mathrm{c}}(A)$ are respectively

  - **the point spectrum of A**

$$\sigma_{\mathrm{p}}(A) = \{z \in \mathbb{C} \mid (zI - A) : D(A) \to \mathcal{H} \text{ non-injective}\} = \{\text{eigenvalues of } A\};$$

  - **the continuous spectrum of A**

$$\sigma_{\mathrm{c}}(A) = \overline{\{z \in \mathbb{C} \mid (zI - A) : D(A) \to \mathcal{H} \text{ injective but non surjective}\}}.$$

[1] **The operator A is called closed if its graph $\Gamma(A) := \{(u, Au),\ u \in D(A)\}$ is a closed subspace of $\mathcal{H} \times \mathcal{H}$.**

On the physical meaning of point and continuous spectra

Theorem (RAGE, Ruelle '69, Amrein and Georgescu '73, Enss '78).

Let H be a locally compact self-adjoint operator on $L^2(\mathbb{R}^d)$.
[Ex.: the Hamiltonian of the hydrogen atom satisfies these assumptions.]

Let $\mathcal{H}_p = \overline{\text{Span}\left\{\text{eigenvectors of } H\right\}}$ and $\mathcal{H}_c = \mathcal{H}_p^\perp$.
[Ex.: for the Hamiltonian of the hydrogen atom, $\dim(\mathcal{H}_p) = \dim(\mathcal{H}_c) = \infty$.]

Let χ_{B_R} be the characteristic function of the ball $B_R = \left\{\mathbf{r} \in \mathbb{R}^d \mid |\mathbf{r}| < R\right\}$.

Then

$$(\phi_0 \in \mathcal{H}_p) \Leftrightarrow \forall \varepsilon > 0, \ \exists R > 0, \ \forall t \geq 0, \ \left\| (1 - \chi_{B_R}) e^{-itH/\hbar} \phi_0 \right\|_{L^2}^2 \leq \varepsilon;$$

$$(\phi_0 \in \mathcal{H}_c) \Leftrightarrow \forall R > 0, \ \lim_{T \to +\infty} \frac{1}{T} \int_0^T \left\| \chi_{B_R} e^{-itH/\hbar} \phi_0 \right\|_{L^2}^2 dt = 0.$$

$\mathcal{H}_{\rm p}$: **set of bound states,** $\mathcal{H}_{\rm c}$: **set of scattering states.**

Definition. **Let A be a closed linear operator on $\mathcal{H}$. Then**

$$\sigma(A) = \sigma_{\rm d}(A) \cup \sigma_{\rm ess}(A)$$

(disjoint union) where

$\sigma_{\rm d}(A) = \{$**isolated eigenvalues of A with finite multiplicity**$\}$ **(discrete spectrum)**
$\sigma_{\rm ess}(A) = \sigma(A) \setminus \sigma_{\rm d}(A)$ **(essential spectrum)**

The essential spectrum therefore consists of

- **the continuous spectrum;**
- **the eigenvalues of infinite multiplicities;**
- **the eigenvalues embedded in the continuous spectrum.**

The notions of discrete and essential spectra are extremely important to understand the numerical approximations of the spectra of Schrödinger, Hartree-Fock, Kohn-Sham or Dirac Hamiltonians.

Spectra of Schrödinger operators with confining potentials

$$\mathcal{H} = L^2(\mathbb{R}^d), \qquad V \in C^0(\mathbb{R}^d), \qquad \lim_{|\mathbf{r}|\to+\infty} V(\mathbf{r}) = +\infty \text{ (confining potential)}$$

$$D(H) = \left\{ u \in L^2(\mathbb{R}^d) \mid -\frac{1}{2}\Delta u + Vu \in L^2(\mathbb{R}^d) \right\}, \qquad \forall u \in D(H),\ Hu = -\frac{1}{2}\Delta u + Vu.$$

H is bounded below and its spectrum is purely discrete ($\sigma_{\rm d}(H) = \sigma(H)$, $\sigma_{\rm ess}(H) = \emptyset$).

As a consequence, H is diagonalizable in a orthonormal basis: there exist

- **a non-decreasing sequence $(E_n)_{n\in\mathbb{N}}$ of real numbers going to $+\infty$;**
- **an orthonormal basis $(\psi_n)_{n\in\mathbb{N}}$ of $\mathcal{H}$ composed of vectors of $D(H)$,**

such that

$$\forall n \in \mathbb{N}, \qquad H\psi_n = E_n\psi_n.$$

In addition, the ground state eigenvalue E_0 is non-degenerate and the corresponding eigenvector can be chosen positive on $\mathbb{R}^d$.

Spectra of 3D Schrödinger operators with potentials decaying at infinity

V **such that** $\forall \varepsilon > 0,\ \exists (V_2, V_\infty) \in L^2(\mathbb{R}^3) \times L^\infty(\mathbb{R}^3)$ **s.t.** $V = V_2 + V_\infty$ **and** $\|V_\infty\|_{L^\infty} \leq \varepsilon$,

$$\mathcal{H} = L^2(\mathbb{R}^3), \qquad D(H) = H^2(\mathbb{R}^3), \qquad \forall u \in D(H),\ Hu = -\frac{1}{2}\Delta u + Vu.$$

The operator H is self-adjoint, bounded below, and $\sigma_{\rm ess}(H) = [0, +\infty)$.

Depending on V, the discrete spectrum of H may be

- **the empty set;**
- **a finite number of negative eigenvalues;**
- **a countable infinite number of negative eigenvalues accumulating at 0 (ex: Ridberg states).**

If H has a ground state, then its energy is a non-degenerate eigenvalue and the corresponding eigenvector can be chosen positive on $\mathbb{R}^d$.

The special case of Kohn-Sham LDA Hamiltonians

$$H_\rho = -\frac{1}{2}\Delta + V_\rho^{\mathrm{KS}} \quad \textbf{with} \quad V_\rho^{\mathrm{KS}}(\mathbf{r}) = -\sum_{k=1}^{M} \frac{z_k}{|\mathbf{r}-\mathbf{R}_k|} + \int_{\mathbb{R}^3} \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}' + \frac{de_{\mathrm{xc}}^{\mathrm{LDA}}}{d\rho}(\rho(\mathbf{r}))$$

For any $\rho \in L^1(\mathbb{R}^3) \cap L^3(\mathbb{R}^3)$, the KS potential V_ρ^{KS} satisfies the assumptions of the previous slide. In particular H_ρ is bounded below and $\sigma_{\mathrm{ess}}(H_\rho) = [0, +\infty)$.

Let $Z = \sum_{k=1}^{M} z_k$ be the total nuclear charge of the molecular system and $N = \int_{\mathbb{R}^3} \rho$.

- **If $N < Z$ (positive ion), H_ρ has a countable infinite number of negative eigenvalues accumulating at 0.**
- **If $N = Z$ (neutral molecular system) and if ρ is a ground state density of the system, then H_ρ has at least N non-positive eigenvalues.**

Spectra of Hartree-Fock Hamiltonians

Let $\Phi = (\phi_1, \cdots, \phi_N) \in (H^1(\mathbb{R}^3))^N$ **be such that** $\int_{\mathbb{R}^3} \phi_i \phi_j = \delta_{ij}$**,**

$$\gamma(\mathbf{r}, \mathbf{r}') = \sum_{i=1}^{N} \phi_i(\mathbf{r})\phi_i(\mathbf{r}'), \qquad \rho_\gamma(\mathbf{r}) = \gamma(\mathbf{r}, \mathbf{r}) = \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2.$$

$$\mathcal{H} = L^2(\mathbb{R}^3), \quad D(H) = H^2(\mathbb{R}^3),$$

$$(H\phi)(\mathbf{r}) = -\frac{1}{2}\Delta\phi(\mathbf{r}) - \sum_{k=1}^{M} \frac{z_k}{|\mathbf{r} - \mathbf{R}_k|}\phi(\mathbf{r}) + \left(\int_{\mathbb{R}^3} \frac{\rho_\gamma(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}'\right)\phi(\mathbf{r}) - \int_{\mathbb{R}^3} \frac{\gamma(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\phi(\mathbf{r}')\, d\mathbf{r}'$$

Let $Z := \sum_{k=1}^{M} z_k$**. The operator** H **is self-adjoint, bounded below, and we have:**

- $\sigma_{\text{ess}} = [0, +\infty)$**;**
- **if** $N < Z$ **(positive ion),** H **has a countable infinite number of negative eigenvalues accumulating at** 0**;**
- **if** $N = Z$ **(neutral molecular system) and if** Φ **is a HF minimizer of the system, then** H **has at least** N **negative eigenvalues (counting multiplicities).**

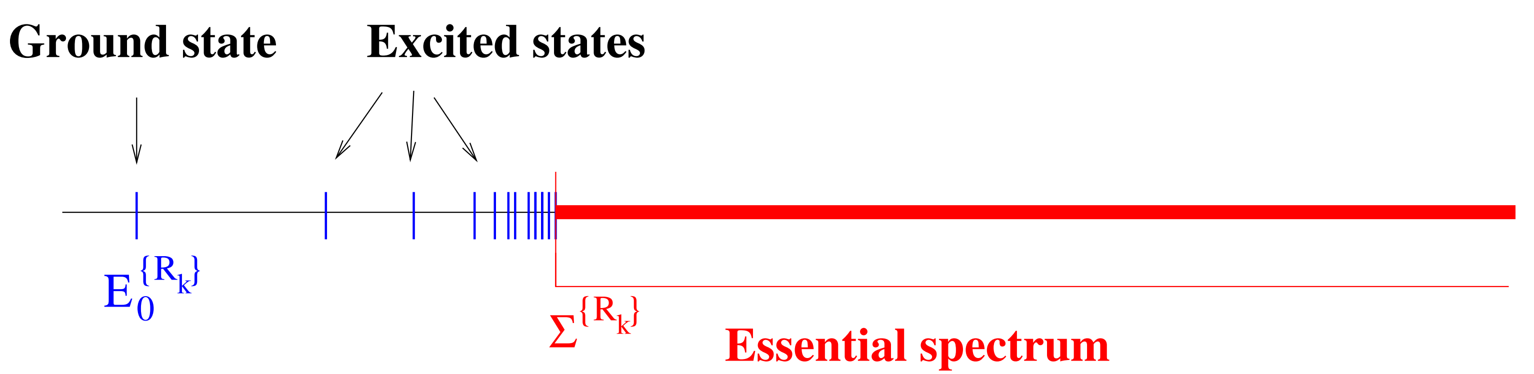
Spectra of N-body electronic Schrödinger Hamiltonians

$$H_N = -\sum_{i=1}^{N} \frac{1}{2}\Delta_{\mathbf{r}_i} - \sum_{i=1}^{N}\sum_{k=1}^{M} \frac{z_k}{|\mathbf{r}_i - \mathbf{R}_k|} + \sum_{1\le i<j\le N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \quad \textbf{on} \quad \mathcal{H} = \bigwedge^N L^2(\mathbb{R}^3, \mathbb{C}^2)$$

(Pauli principle)

Zhislin's theorem: if $N \le \sum_{k=1}^{M} z_k$ (neutral or positively charged system), then

$$\sigma(H_N) = \left\{E_N^0 \le E_N^1 \le E_N^2 \le \cdots\right\} \cup [\Sigma_N, +\infty), \quad \textbf{with } \Sigma_1 = 0 \textbf{ and } \Sigma_N < 0 \textbf{ if } N \ge 2.$$



In addition, $\Sigma_N = E_{N-1}^0$ (HVZ theorem).

Spectra of Dirac Hamiltonians

$$\mathcal{H} = L^2(\mathbb{R}^3; \mathbb{C}^4), \qquad D(D_0) = H^1(\mathbb{R}^3; \mathbb{C}^4), \qquad D_0 = c\vec{p} \cdot \vec{\alpha} + mc^2\beta$$

$$p_j = -\mathrm{i}\hbar\partial_j, \qquad \alpha_j = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix} \in \mathbb{C}^{4\times 4}, \qquad \beta = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix} \in \mathbb{C}^{4\times 4}$$

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$ **(Pauli matrices)**

The free Dirac operator D_0 is self-adjoint and

$$\sigma(D_0) = \sigma_{\mathrm{ac}}(D_0) = (-\infty, -mc^2] \cup [mc^2, +\infty).$$

Theorem. **Let** $\alpha := \frac{e^2}{4\pi\varepsilon_0\hbar c} \simeq 1/137.036$ **be the fine structure constant. Let**

$$D_Z = D_0 - \frac{Z}{|\mathbf{r}|}, \qquad Z \in \mathbb{R} \quad \textbf{(physical cases: } Z = 1, 2, 3, \cdots\textbf{).}$$

- **if** $|Z| < \frac{\sqrt{3}}{2\alpha} \simeq 118.677$**, the Dirac operator** D_Z **is essentially self-adjoint (meaning that there exists a unique domain** $D(D_Z)$ **containing** $C_c^\infty(\mathbb{R}^3;\mathbb{C}^4)$ **for which** D_Z **is self-adjoint);**
- **if** $|Z| > \frac{\sqrt{3}}{2\alpha} \simeq 118.677$**,** D_Z **has many self-adjoint extensions;**
- **if** $|Z| < \frac{1}{\alpha} \simeq 137.036$**,** D_Z **has a special self-adjoint extension, considered as the physical one. The essential spectrum of this self-adjoint extension is** $(-\infty, -mc^2] \cup [mc^2, +\infty)$ **and its discrete spectrum consist of the eigenvalues**

$$E_{nj} = mc^2 \left[1 + \left(\frac{Z\alpha}{n - j - \frac{1}{2} + \sqrt{(j + \frac{1}{2})^2 - Z^2\alpha^2}} \right)^2 \right]^{-1/2}, \quad n \in \mathbb{N}^*, \; j = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \cdots \leq n - \frac{1}{2}.$$

Many-body Dirac-Coulomb Hamiltonian are not understood mathematically.

3 - Spectral decomposition and functional calculus

Diagonalizable self-adjoint operators and Dirac's bra-ket notation

Let A be a self-adjoint operator that can be diagonalized in an orthonormal basis $(e_n)_{n\in\mathbb{N}}$ (this is not the case for many useful self-adjoint operators!).

Dirac's bra-ket notation: $$A = \sum_{n\in\mathbb{N}} \lambda_n |e_n\rangle\langle e_n|, \quad \lambda_n \in \mathbb{R}, \quad \langle e_m|e_n\rangle = \delta_{mn}.$$

Then,

- **the operator A is bounded if and only if $\|A\| = \sup_n |\lambda_n| < \infty$;**
- $D(A) = \left\{|u\rangle = \sum_{n\in\mathbb{N}} u_n|e_n\rangle \mid \sum_{n\in\mathbb{N}}(1+|\lambda_n|^2)|u_n|^2 < \infty\rangle\right\}$**;**
- $\sigma_{\rm p}(A) = \{\lambda_n\}_{n\in\mathbb{N}}$ **and** $\sigma_{\rm c}(A) = \big\{$**accumulation points of** $\{\lambda_n\}_{n\in\mathbb{N}}\big\}\setminus\sigma_{\rm p}(A)$**;**
- $\mathcal{H}_{\rm p} = \mathcal{H}$ **and** $\mathcal{H}_{\rm c} = \{0\}$ **(no scattering states);**
- **functional calculus for diagonalizable self-adjoint operators: for all $f:\mathbb{R}\to\mathbb{C}$, the operator $f(A)$ defined by**

$$D(f(A)) = \left\{|u\rangle = \sum_{n\in\mathbb{N}} u_n|e_n\rangle \mid \sum_{n\in\mathbb{N}}(1+|f(\lambda_n)|^2)|u_n|^2 < \infty\rangle\right\}, \quad f(A) = \sum_{n\in\mathbb{N}} f(\lambda_n)|e_n\rangle\langle e_n|$$

is independent of the choice of the spectral decomposition of A.

Theorem (functional calculus for bounded functions). **Let $\mathfrak{B}(\mathbb{R},\mathbb{C})$ be the $*$-algebra of bounded $\mathbb{C}$-valued Borel functions on $\mathbb{R}$ and let A be a self-adjoint operator on $\mathcal{H}$. Then there exists a unique map**

$$\Phi_A \,:\, \mathfrak{B}(\mathbb{R},\mathbb{C}) \ni f \mapsto f(A) \in \mathcal{B}(\mathcal{H})$$

satisfies the following properties:

1. **Φ_A is a homomorphism of $*$-algebras:**

$$(\alpha f + \beta g)(A) = \alpha f(A) + \beta g(A), \quad (fg)(A) = f(A)g(A), \quad \overline{f}(A) = f(A)^*;$$

2. $\|f(A)\| \leq \sup_{x\in\mathbb{R}} |f(x)|$**;**

3. **if $f_n(x) \to x$ pointwise and $|f_n(x)| \leq |x|$ for all n and all $x \in \mathbb{R}$, then**

$$\forall u \in D(A), \quad f_n(A)u \to Au \text{ in } \mathcal{H};$$

4. **if $f_n(x) \to f(x)$ pointwise and $\sup_n \sup_{x\in\mathbb{R}} |f_n(x)| < \infty$, then**

$$\forall u \in \mathcal{H}, \quad f_n(A)u \to f(A)u \text{ in } \mathcal{H};$$

In addition, if $u \in \mathcal{H}$ is such that $Au = \lambda u$, then $f(A)u = f(\lambda)u$.

Theorem (spectral projections and functional calculus).

Let A be a self-adjoint operator on $\mathcal{H}$.

- **For all $\lambda \in \mathbb{R}$, the bounded operator $P_\lambda^A := \mathbb{1}_{]-\infty,\lambda]}(A)$, where $\mathbb{1}_{]-\infty,\lambda]}(\cdot)$ is the characteristic function of $]-\infty,\lambda]$, is an orthogonal projection.**
- **For any $(u,v) \in \mathcal{H} \times \mathcal{H}$, $\lambda \mapsto \langle u|P_\lambda^A u\rangle$ is the distribution function of a finite positive Borel measure on $\mathbb{R}$, and $\lambda \mapsto \langle v|P_\lambda^A u\rangle$ is the distribution function of a finite complex Borel measure on $\mathbb{R}$.**
- **Spectral decomposition of A: for all $u \in D(A)$ and $v \in \mathcal{H}$, it holds**

$$\langle v|Au\rangle = \int_{\mathbb{R}} \lambda \, d\langle v|P_\lambda^A u\rangle, \quad \text{which we denote by} \quad A = \int_{\mathbb{R}} \lambda \, dP_\lambda^A.$$

- **Functional calculus: let f be a (not necessarily bounded) $\mathbb{C}$-valued Borel function on $\mathbb{R}$. The operator $f(A)$ can be defined by**

$$D(f(A)) := \left\{ u \in \mathcal{H} \mid \int_{\mathbb{R}} |f(\lambda)|^2 \, d\langle u|P_\lambda^A u\rangle < \infty \right\},$$

$$\forall (u,v) \in D(f(A)) \times \mathcal{H}, \ \langle v|f(A)u\rangle := \int_{\mathbb{R}} f(\lambda) \, d\langle v|P_\lambda^A u\rangle.$$