Inverse Problems in Image Science and Application to Tomography

Pierre Maréchal

August 30, 2004

Contents

Chapter 1

Introduction

1.1 Inverse problems

In this section, F and G are (separable) real or complex Hilbert spaces. Let $A\colon F \to G$ be a *continuous linear application* (in short, an *operator*). Consider the following problem:

Given $g \in G$, find $f \in F$ such that $g = Af$.

The problem is said to *well-posed* if

(1) $\forall g \in G,\ \exists!\, f \in F\colon g = Af$;

(2) the solution f depends continuously on g.

In other words, the problem is well-posed if A is invertible and its inverse $A^{-1}\colon G \to F$ is continuous. Existence and uniqueness of a solution for all $g \in G$ (Condition (1)) is equivalent to surjectivity and injectivity of A, respectively. Stability of the solution (Condition (2)) amounts to continuity of A^{-1}. Conditions (1) and (2) are referred to as the Hadamard Conditions. A problem that is not well-posed is said to be *ill-posed*.
Note that a 'mathematically' well-posed problem may be ill-posed in practice: the solution may (exist, be unique and) depend continuously on the data but still be very sensitive to small perturbations of it. An error δg produces the error $\delta f = A^{-1}\,\delta g$, which may have dramatic consequences on the interpretation of the solution. This can be understood by considering the inequality

$$\|\delta f\| \le \|A^{-1}\|\,\|\delta g\|,$$

in which $\|A^{-1}\|$ denotes the *spectral norm* of A^{-1}. We see that, if $\|A^{-1}\|$ is very large, errors may be strongly amplified by the action of A^{-1}.

In order to *regularize* an ill-posed problem, it will be necessary to restate the problem in such a way that the Hadamard conditions be satisfied. The preceeding remark also indicates that one should be able to estimate the effective *sensitivity* to perturbations of the data. Restating the problem alway implies reducing one's ambition: one should give up restoring all the information which an *ideal* solution would carry. The difficulty is then to find the right trade-off between the quantity of information to be retrieved and its accuracy. (This is why the concepts of information theory plaid such an important role in the history of inverse problems.)
In each area of the applied sciences where ill-posed problems occur, strategies for obtaining licit solutions have been developped. Many of them result from rather empirical approaches. Each of these methods can be described by means of an algorithm only, and the solution itself is not clearly defined: it is merely 'the point towards which the algorithm converges, or seems to converge'. In these lecture notes, we deal instead with methodologies that lead to solutions which are precisely defined as minimizers of functionals. They have the advantage of being more transparent: they allow for sensitivity analysis, and thus for *a priori* calibration of all the parameters that control the regularization.

1.2 Regularization heuristics

For simplicity, we assume here that

$$F = \mathbb{R}^n, \quad G = \mathbb{R}^m, \quad \text{and} \quad A \in \mathcal{M}_{m\times n}(\mathbb{R}).$$

We use the same notation for the linear applications from $\mathbb{R}^n$ to $\mathbb{R}^m$ and their matrix representations in the canonical bases of $\mathbb{R}^n$ and $\mathbb{R}^m$.
Consider the following problem:

Given $y \in \mathbb{R}^m$, find $x \in \mathbb{R}^n$ such that $y = Ax$.

Suppose first that A is injective but not surjective. The data y may not belong to the range of A. It seems natural to replace y by its projection onto the range of A. This amounts to replacing the original problem by that of finding the vector $\overline{x}$ which minimizes, over $\mathbb{R}^n$, the function

$$f(x) := \|y - Ax\|^2.$$

Equivalently, $\overline{x}$ is the unique solution of the *normal equation*:

$$A^\star y = A^\star Ax.$$

We have denoted by $A^\star$ the transpose of A. Note that, since A is assumed to be injective, $A^\star A$ is invertible. This reformulation of the problem clearly satisfies the Hadamard conditions. The vector $\overline{x}$ is called the *least square solution*.
Suppose now that A is surjective but not injective. Then all vectors in the affine subspace $A^{-1}(\{y\})$ are solutions, and we are facing the problem of selecting a meaningful one. We need to compensate for the lack of constraints on the unknown object x, and therefore to add constraints in a way or another. This amounts to introducing *a priori* information, which involves the choice of some *inference principle*. For example, the *minimum norm solution* is, needless to say, the unique vector which minimizes, over $A^{-1}(\{y\})$, the function $f(x) := \|x\|^2$. In the case where the unknown vector can be regarded as a probability distribution, information theoretic considerations may lead to the *maximum entropy solution*. The latter is the (unique) vector which maximizes, over $A^{-1}(\{y\})$, the entropy:

$$f(x) := -\sum_{k=1}^{n} x_j \ln x_j.$$

This has been extensively used in image science, even in contexts where the interpretation of x in terms of probability is not entirely natural. Many other functions can be designed, as attests the abundant (and often redundant) litterature on the subject.
If A is neither injective nor surjective, we may combine the preceeding strategies, and thus define the solution to be the vector $\overline{x}$ which optimizes some *entropy-like* function over the affine subspace $A^{-1}(\{\tilde{y}\})$, where $\tilde{y}$ is the projection of y onto the range of A. For example, the *minimum norm least square solution* is defined to be the vector which solves the optimization problem:

$$(\mathcal{P}) \quad \left| \begin{array}{ll} \text{minimize} & F(x) := \|x\|^2 \\ \text{subject to} & x \in A^{-1}(\{\tilde{y}\}). \end{array} \right.$$

Exercise 1.1 [Characterization of $A^{-1}(\{\tilde{y}\})$] Prove that the following statements are equivalent:

(a) x' solves $\tilde{y} = Ax$;

(b) x' minimizes $\|y - Ax\|^2$;

(c) x' solves the *normal equation* $A^\star y = A^\star Ax$.

Exercise 1.2 Prove the following relationships:

$$(\ker A)^{\perp} = \operatorname{ran} A^{\star} = \operatorname{ran} A^{\star}A \quad \text{and} \quad (\ker A^{\star})^{\perp} = \operatorname{ran} A = \operatorname{ran} AA^{\star}.$$

As we shall see later on, the results of Exercises 1.1 and 1.2 remain true in general Hilbert spaces.
The solution x^+ to Problem $(\mathcal{P})$ depends linearly on y. It is the image of y by the *pseudo-inverse* A^+ of A. As a linear mapping from $\mathbb{R}^m$ to $\mathbb{R}^n$, A^+ is, of course, continuous. However, if the norm of A^+ is very large, small perturbations of y may give rise to large perturbations of x^+. It can be shown that the choice of any other inference principle won't improve the situation. Sensitivity of a solution to perturbations of the data is inherent to the constraint equation (namely, the normal equations, which itself stems out from the least square principle). It appears that, in this case, the constraints need to be *relaxed*. In other words, some inertia has to be introduced in the reformulation of the problem. A standard strategy for doing so is to replace the original problem by that of finding the solution to an optimization problem of the form:

$$(\mathcal{P}) \quad \left| \begin{array}{ll} \text{minimize} & F(x) := \|y - Ax\|^2 + \alpha\rho(x) \\ \text{s. t.} & x \in C, \end{array} \right.$$

in which α is a *regularization parameter*, ρ is a *regularization function* (also called an entropy) and C is a constraint set which incarnates part of the *a priori* knowledge on x, such as non-negativity. The fit term (here, $\|y-Ax\|^2$) can be regarded as a tolerant expression of the constraint equation. Of course, one may replace $\|y - Ax\|^2$ by some other function attracting Ax to y.
In these notes, we intend to give an answer to the following question: How should we choose α, ρ (and C) so that the problem is adequately reformulated ? By 'adequately reformulated,' we mean that Problem $(\mathcal{P})$ should be such that

1. a unique and stable solution exists (Hadamard conditions);
2. the solution is 'physically relevant;'
3. the solution is computable.

We restrict our attention to problems of image reconstruction or restoration.

Chapter 2

Hilbert spaces

2.1 Inner product spaces

Definition 2.1 Let F be a vector space over the field of complex numbers (in short: a $\mathbb{C}$-vector space). An *inner product* on F is a mapping

$$\begin{array}{rccc} \langle\cdot,\cdot\rangle : & F\times F & \longrightarrow & \mathbb{C} \\ & (x,y) & \longmapsto & \langle x,y\rangle \end{array}$$

which satisfies the following axioms:

(1) for all fixed $y\in F$, $\langle\cdot,y\rangle$ is linear;

(2) for all $x,y\in F$, $\langle y,x\rangle=\overline{\langle x,y\rangle}$;

(3) for all $x\in F$, $\langle x,x\rangle\geq 0$ and $\langle x,x\rangle=0$ if and only if $x=0$.

The pair $(F,\langle\cdot,\cdot\rangle)$ is called an *inner product space*.

Example 2.1 The space l^2 of complex sequences $x=(x_k)_{k\in\mathbb{N}^*}$ such that $\sum|x_k|^2$ is convergent, with $\langle x,y\rangle=\sum x_k\overline{y_k}$.

Example 2.2 The space $L^2(\mathbb{R}^n)$ of complex-valued square-integrable functions on $\mathbb{R}^n$, with $\langle f,g\rangle=\int f\overline{g}$.

From now on, unless we explicitly indicate otherwise $(F,\langle\cdot,\cdot\rangle)$ will be a complex inner product space.

Theorem 2.1 [Schwarz' inequality] For all $x,y\in F$,

$$|\langle x,y\rangle|^2\leq\langle x,x\rangle\langle y,y\rangle.$$

Corollary 2.1 The mapping $x \mapsto \sqrt{\langle x, x \rangle}$ is a norm on F.

Definition 2.2 If F is complete relatively to the norm induced by $\langle \cdot, \cdot \rangle$, F is said to be a *Hilbert space*.

The spaces l^2 and $L^2(\mathbb{R}^n)$ are examples of Hilbert spaces.

Proposition 2.1 Let $(F, \langle \cdot, \cdot \rangle)$ be a complex inner product space. Then,

(1) $\forall x, y \in F$, $\|x+y\|^2 = \|x\|^2 + 2\mathcal{R}e(\langle x, y \rangle) + \|y\|^2$;

(2) $\forall x, y \in F$, $\|x+y\|^2 + \|x-y\|^2 = 2\big(\|x\|^2 + \|y\|^2\big)$;

(3) $\forall x, y \in F$, $4\langle x, y \rangle = \|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2$.

The first and second equalities are referred to as the *polarization identity* and the *parallelogram law*, respectively. As for the third equality, it shows that inner products are entirely determined by their associated norms.

Definition 2.3 A *linear manifold* is a non-empty subset M of F which is stable under linear combination. A *subspace* of F is a closed linear manifold.

It is readily seen that the intersection of any family of linear manifolds (resp. subspaces) is a linear manifold (resp. a subspace).

Definition 2.4 Given a subset S of F, the intersection of all subspaces that contain S is called the *span* of S. It is denoted by $\operatorname{span} S$. The set of all linear combinations of elements of S is called the *linear hull* of S. It is denoted by $\operatorname{vect} S$. In the particular case where S is a singleton $\{x\}$, we shall write: $\operatorname{vect}\{x\} = \langle x \rangle$.

Theorem 2.2 Let S be a non-empty subset of F. Then $\operatorname{span} S = \operatorname{cl} \operatorname{vect} S$.

Definition 2.5 A family $(x_\alpha)_{\alpha \in A} \subset F$ of vectors is said to be *linearly independent* if every finite subfamily of (x_α) is linearly independent.

2.2 Orthogonality

Throughout this section, F will denote an inner product space.

Definition 2.6 The vectors $x, y \in F$ are said to be *orthogonal* if $\langle x, y\rangle = 0$. Notation: $x \perp y$. The subsets $E_1, E_2 \subset F$ are said to be orthogonal if $x_1 \perp x_2$ for all $x_1 \in E_1$ and $x_2 \in E_2$. Finally, the *orthogonal complement* of a subset $S \subset F$ is the set defined by:

$$S^\perp := \{x \in F | \forall y \in S,\ x \perp y\}.$$

Definition 2.7 A family $(x_\alpha)_{\alpha\in A} \subset F$ is said to be *orthogonal* if for any two distinct $\alpha, \alpha' \in A$, $x_\alpha \perp x_{\alpha'}$, and *orthonormal* if in addition every vector in (x_α) has unit norm.

Theorem 2.3 [Bessel's inequality] Let $\{e_1, \ldots, e_n\}$ be a (finite) orthonormal family and x be any vector in F. Then,

$$\sum_{k=1}^{n} |\langle x, e_k\rangle|^2 \leq \|x\|^2$$

Theorem 2.4 [Pythagorean Theorem] Suppose $x, y \in F$ are orthogonal. Then $\|x+y\|^2 = \|x\|^2 + \|y\|^2$. Similarly, if $\{x_1, \ldots, x_n\}$ is orthogonal, then $\|x_1 + \cdots + x_n\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2$.

Theorem 2.5 Let $(e_\alpha)_{\alpha\in A} \subset F$ an orthonormal family. Then (e_α) is linearly independent.

Exercise 2.1 Let S be any subset of F. Prove that $S^\perp$ is a subspace of F.

Theorem 2.6 Let $E \subset F$ be a linear manifold. Then $E^{\perp\perp} := (E^\perp)^\perp = \operatorname{cl} E$. Consequently, if E is a subspace, then $E^{\perp\perp} = E$.

The sum of closed sets of a normed space is not, in general, a closed set, even if these sets are subspaces. However, for orthogonal subspaces of an inner product space, the following holds.

Theorem 2.7 Let E_1, E_2 be orthogonal subspaces of F. Then $E_1 + E_2$ is closed.

Theorem 2.8 Let E be a subspace of F. Then F is the direct sum of E and $E^\perp$, which we write: $F = E \oplus E^\perp$.

Definition 2.8 An *orthonormal basis* of F is an orthonormal family (x_α) such that $F = \operatorname{span} \cup_\alpha \langle x_\alpha\rangle$. If F is a Hilbert space, we say that (x_α) is a *Hilbert Basis* of F.

The next theorem provides a constructive method for obtaining an orthonormal sequence from any linearly independent sequence.

Theorem 2.9 [Gram-Schmidt orthonormalization] Let $\{e_1, e_2, \ldots\}$ be a linearly independent sequence in F. Then, there exists an orthonormal sequence $\{\varepsilon_1, \varepsilon_2, \ldots\}$ such that, for all $n \in \mathbb{N}$, $\{\varepsilon_1, \ldots, \varepsilon_n\}$ is a basis of $\mathrm{vect}\{e_1, \ldots, e_n\}$.

2.3 Infinite sums

Definition 2.9 [Summable families] Let F be a Banach space. A family $(x_\alpha)_{\alpha \in A} \subset F$ is said to be *summable* to x if for all $\varepsilon > 0$, there exists a finite subset I of A such that for every finite superset J of I,

$$\left\| x - \sum_{j \in J} x_j \right\| < \varepsilon.$$

We then write: $x = \sum_{\alpha \in A} x_\alpha$.

Theorem 2.10 The following statements are equilvalent:

(1) (x_α) is summable;

(2) $\forall \varepsilon > 0$, there exists a finite subset B_0 of A such that, for every finite subset C of A, $C \cap B_0 = \emptyset$ implies $\| \sum_{\alpha \in C} x_\alpha \| < \varepsilon$.

Theorem 2.11 The set of all non-zero elements of a summable family is at most countable.

Theorem 2.12 Let F be a Hilbert space. Suppose that (x_α) is an orthogonal family. The following statements are equivalent:

(1) (x_α) is summable;

(2) $(\|x_\alpha\|^2)$ is summable.

Furthermore, if $x = \sum_{\alpha \in A} x_\alpha$, then $\|x\|^2 = \sum_{\alpha \in A} \|x_\alpha\|^2$.

Example 2.3 A sequence $\{x_1, x_2, \ldots\}$ in F is summable if and only if there exists $x \in F$ such that

$$\left\| x - \sum_{k=1}^{n} x_k \right\|^2 \longrightarrow 0 \qquad \text{as} \qquad n \longrightarrow \infty.$$

Definition 2.10 [Infinite sums of subspaces] Given a family $(E_\alpha)_{\alpha\in A}$ of subspaces of F, the *sum* of (E_α) is defined to be the set

$$\sum_{\alpha\in A} E_\alpha := \left\{ \sum x_\alpha \,\middle|\, x_\alpha \in E_\alpha,\ (x_\alpha) \text{ is summable} \right\}.$$

The sum is said to be *direct* if for each $x \in \sum E_\alpha$, the representation $x = \sum x_\alpha$ (with $x_\alpha \in E_\alpha$) is unique.

Theorem 2.13 Let (E_α) be a family of subspaces of F. then

$$\operatorname{span} \cup_\alpha E_\alpha = \operatorname{cl} \sum_\alpha E_\alpha$$

Furthermore, if (E_α) is orthogonal (that is, if $E_\alpha \perp E_{\alpha'}$ whenever $\alpha \neq \alpha'$), the closure operation in the above equation can be ommitted and the sum is direct. In this case, (E_α) is called an *orthogonal decomposition* of $\operatorname{span} \cup E_\alpha$.

Theorem 2.14 Let $(e_\alpha) \subset F$ be an orthonormal family. The following for statements are equivalent:

(1) (e_α) is a basis of F;

(2) for all $x \in F$, $x = \sum \langle x, e_\alpha\rangle e_\alpha$;

(3) for all $x, y \in F$, $\langle x, y\rangle = \sum \langle x, e_\alpha\rangle \overline{\langle y, e_\alpha\rangle}$

(4) for all $x \in F$, $\|x\|^2 = \sum |\langle x, e_\alpha\rangle|^2$.

2.4 Separable inner product spaces

Definition 2.11 An inner product space is said to be *separable* if it contains a dense countable subset. A separable Hilbert space is a separable complete inner product space.

Theorem 2.15 Every orthonormal family of a separable inner product space is at most countable. In particular, every Hilbert basis of a separable inner product space is at most countable. Conversely, every separable inner product space possesses a countable orthonormal basis.

Theorem 2.16 Let F be a separable Hilbert space and $\{e_1, e_2, \ldots\}$ be a Hilbert basis of F. Then,

$$F = \left\{ \sum \alpha_k e_k \,\middle|\, \sum |\alpha_k|^2 < \infty \right\}.$$

2.5 Projections

Lemma 2.1 Let $(F, \langle \cdot, \cdot \rangle)$ be an inner product space, let $x, y, z \in F$, and let $m := (y+z)/2$. Then

$$\|x-y\|^2 + \|x-z\|^2 = 2\|x-m\|^2 + \frac{1}{2}\|y-z\|^2 .$$

Theorem 2.17 Let $(F, \langle \cdot, \cdot \rangle)$ be an inner product space, $K \subset F$ be a complete convex set and x be a vector in F. Then, there exists a unique $x_0 \in K$ such that

$$\|x - x_0\| = \inf\left\{\left\|x - x'\right\| \mid x' \in K\right\}.$$

This vector x_0 is called the *projection* of x onto K. Furthermore, x_0 is the unique vector of K satisfying

$$\forall y \in K, \quad \mathcal{Re}(\langle x - x_0, y - y_0 \rangle) \leq 0.$$

If F is a Hilbert space, a sufficient condition for K to be complete is that it be closed. In the particular case where K is a vector subspace, the next theorem holds:

Theorem 2.18 Let $(F, \langle \cdot, \cdot \rangle)$ be an inner product space, E be a complete linear manifold of F and x be a vector in F. The following are equivalent:

(1) x_0 is the projection of x onto E;

(2) $x_0 \in E$ and $x - x_0 \in E^{\perp}$.

Furthermore, the mapping P which maps $x \in F$ to its projection onto E enjoys the following properties:

(a) P is linear;

(b) for all $x \in E$, $\|P(x)\| \leq \|x\|$;

(c) $P \circ P = P$;

(d) $\ker P = E^{\perp}$ and $\operatorname{ran} P = E$.

Theorem 2.19 Let $\{e_1, \ldots, e_n\} \subset F$ be an orthonormal set and P be the orthogonal projection onto the subspace $\operatorname{vect}\{e_1, \ldots, e_n\}$. Then,

$$\forall x \in F, \quad P(x) = \sum_{k=1}^{n} \langle x, e_k \rangle \, e_k.$$

Theorem 2.20 [Riesz-Fischer] Let F be a Hilbert space and u be a continuous linear form, that is, a continuous linear mapping from F into $\mathbb{C}$. Then there exists a unique $a \in F$ such that, for all $x \in F$, $u(x) = \langle x, a \rangle$. Furthermore, the norm of u in the space $L(F, \mathbb{C})$ of all continuous linear forms on F is given by

$$\|u\|_{L(F,\mathbb{C})} := \sup \{ |u(x)| \mid \|x\| \le 1 \} = \|a\| .$$

Chapter 3

Operators

3.1 Hermitian operators

Throughout this section, F and G will be separable Hilbert spaces, with inner products denoted by $\langle\cdot,\cdot\rangle$. and A a continuous linear application (in short: an *operator*) from F to G. Recall that continuity of A is equivalent to boundedness:

$$\|A\| := \sup_{\|x\|\leq 1} \|Ax\| = \sup_{\|x\|=1} \|Ax\| < \infty.$$

Using the Riesz-Fischer Theorem, the reader may show that there is a unique continuous linear application $A^\star \colon G \to F$ such that

$$\forall x \in F, \quad \forall y \in G, \quad \langle Ax, y\rangle = \langle x, A^\star y\rangle \, .$$

This operator is called the adjoint of A.

Proposition 3.1 Let $A, B \colon F \to F$ be operators. Then the following holds:

(1) for all $\alpha, \beta \in \mathbb{C}$, $(\alpha A + \beta B)^\star = \overline{\alpha} A^\star + \overline{\beta} B^\star$;

(2) $(AB)^\star = B^\star A^\star$;

(3) $A^{\star\star} := (A^\star)^\star = A$;

(4) if A is invertible, then so is $A^\star$, and $(A^\star)^{-1} = (A^{-1})^\star$;

(5) on denoting by $\|\cdot\|$ the spectral norm of $L(F) := L(F, F)$, one has $\|A\|^2 = \|A^\star\|^2 = \|A^\star A\|$.

Proposition 3.2 Let $A\colon F \to G$ be linear and surjective. Then, the following statements are equivalent:

(1) A has a continuous inverse A^{-1};

(2) $\mu := \inf\{\|Ax\|^2 | \|x\| = 1\} > 0$.

The spectral norm of A^{-1} then satisfies: $\|A^{-1}\|^2 = \mu^{-1}$.

Definition 3.1 An operator $A\colon F \to F$ is said to be *Hermitian* (or *self-adjoint*) if $A^\star = A$ and *nonnegative Hermitian* if, in addition, $\langle Ax, x\rangle \geq 0$ for all $x \in F$.

Theorem 3.1 Let $A\colon F \to F$ be a Hermitian operator. Then, the spectral norm of A satisfies

$$\|A\| := \sup_{\|x\|=1} \|Ax\| = \sup_{\|x\|=1} |\langle Ax, x\rangle| \, .$$

3.2 Compact operators

We now turn to the crucial notion of compact operator.

Theorem 3.2 Let F and G be Banach spaces and $A\colon F \to G$ be a linear application. Then, the following statements are equivalent:

(1) the image by A of every bounded set is precompact;

(2) for all bounded sequence $(x_n) \subset F$, the sequence (Ax_n) has a convergent subsequence.

In this case, A is bounded, and A is called a *compact operator*.

It is easy to see that linear combinations and products of compact operators are compact.

Exercise 3.1 Let F and G be Banach spaces and $A\colon F \to G$ be a bounded linear application. Prove that, if $\dim \operatorname{ran} A < \infty$, then A is compact.

When $\operatorname{rk} A := \dim \operatorname{ran} A < \infty$, A is said to have *finite rank*.

Theorem 3.3 Let F and G be Banach spaces and (A_k) be a sequence of compact operators from F into G. Suppose there exists an operator A such that $\|A - A_k\|$ tends to 0 as $k \to \infty$. Then A is compact.

Theorem 3.4 Let F and G be Banach spaces and $A\colon F \to G$ be a bounded linear application. Then, the following statements are equivalent:

(1) A is compact;

(2) $A^\star$ is compact;

(3) there exists a sequence (A_k) of finite rank operators such that $\|A - A_k\| \to 0$ as k tends to infinity.

3.3 Spectral theorem

The results of this section are crucial to the theory of linear inverse problems in Hilbert spaces. Throughout, F will be a separable Hilbert space and $A\colon F \to F$ will be a Hermitian operator.

Definition 3.2 A vector subspace E of F is said to be *A-stable* if $AE \subset E$.

Recall that the vector $x \in F$ is said to be an *eigenvector* of A if $x \neq 0$ and $Ax = \lambda x$ for some $\lambda \in \mathbb{C}$. The complex number λ is then said to an *eigenvalue* of A, and the set $\Lambda(A)$ of all eigenvalues of A is called the *spectrum* of A. It is readily seen that $\Lambda(A)$ is the set of all complex numbers λ such that $A - \lambda I$ is not invertible. (We have denoted by I the identity of F.) The following facts are immediate.

Proposition 3.3 Let A be as above and $E \subset F$ be a linear manifold.

(a) If E is A-stable, then so is $\operatorname{cl} E$;

(b) if E is A-stable, then so is $E^\perp$;

(c) all eigenvalues are real: $\Lambda(A) \subset \mathbb{R}$;

(d) if x is an eigenvector, then $\langle x \rangle$ and $\langle x \rangle^\perp$ are A-stable;

(e) for all $\lambda \in \Lambda(A)$, the *eigenspace* $E_\lambda := \{x \in F | Ax = \lambda x\}$ is closed;

(f) if λ_1 and λ_2 are two distinct eigenvalues, then $E_{\lambda_1} \perp E_{\lambda_2}$;

(g) if E is a finite-dimensional A-stable subspace of F, then the restriction A_E of A to E is a Hermitian operator and there exists an orthonormal basis of E that consists of eigenvectors of A.

Lemma 3.1 Assume F is non trivial ($F \neq \{0\}$) and let $A \colon F \to F$ be a Hermitian compact operator. Then either $\|A\|$ or $-\|A\|$ belongs to the spectrum of A.

Theorem 3.5 Let $A \colon F \to F$ be a Hermitian compact operator. Then

$$F = \operatorname{span} \cup_{\lambda \in \Lambda(A)} E_\lambda = \sum_{\lambda \in \Lambda(A)} E_\lambda .$$

In other words, $\{E_\lambda | \lambda \in \Lambda(A)\}$ is an orthogonal decomposition of F.

Corollary 3.1 There exists an orthonormal basis that consists of eigenvectors of A.

Observe that, if λ is a nonzero eigenvalue, then $\dim E_\lambda < \infty$. For otherwise there would be an orthonormal sequence (e_k) whose image $(Ae_k) = \lambda(e_k)$ has no convergent subsequence. The same argument will show that, if μ is any positive number, only finitely many eigenvalues λ are such that $|\lambda| \geq \mu$. Therefore, the following holds:

Corollary 3.2 Suppose $\dim F = \infty$. Then 0 is an accumulation point of $\Lambda(A)$: $0 \in \operatorname{cl} \Lambda(A)$.

3.4 Hilbert-Schmidt Operators

Throughout this section, F and G will be separable, infinite dimensional, Hilbert space. Recall that separability implies existence of Hilbert bases.

Theorem 3.6 Let $(f_j)_{j \in \mathbb{N}^*}$ and $(g_k)_{k \in \mathbb{N}^*}$ be Hilbert bases of F and G, respectively, and let $A \in L(F, G)$. Then

$$\sum_{j \in \mathbb{N}^*} \|Af_j\|^2 = \sum_{k \in \mathbb{N}^*} \|Ag_k\|^2 .$$

Theorem 3.6 shows that $\sum_j \|Af_j\|^2$ is independent of the choice the Hilbert basis (f_j). The quantity $\sum_j \|Af_j\|^2$ is referred to as the *trace* of $A^\star A$ and is denoted by $\operatorname{tr} A^\star A$.
We say that $A \in L(F, G)$ is a *Hilbert-Schmidt operator* if $\operatorname{tr} A^\star A$ is finite. The Hilbert-Schmidt operators of F into G form a vector subspace of $L(F, G)$, denoted by $HS(F, G)$. Note that

$$\begin{array}{rccl} \langle \cdot, \cdot \rangle_{HS} : & HS(F,G)^2 & \longrightarrow & \mathbb{C} \\ & (A, B) & \longmapsto & \sum_j \langle Af_j, Bf_j \rangle \end{array}$$

is an inner product, which turns $HS(F,G)$ into a Hilbert space. Since the corresponding norm, namely

$$\begin{array}{rccl} \|\cdot\|_{HS}: & HS(F,G) & \longrightarrow & \mathbb{R}_+ \\ & A & \longmapsto & \|A\|_{HS} := \sqrt{\operatorname{tr} A^\star A}, \end{array}$$

is independent of the choice of the Hilbert basis, so is $\langle\cdot,\cdot\rangle_{HS}$ (see Proposition 2.1(3)). We now focuse on the case where $F = L^2_V(\mathbb{R}^n)$ and $G = L^2_W(\mathbb{R}^n)$, in which V and W are subsets of $\mathbb{R}^n$.

Proposition 3.4 Let $(f_j)_{j\in\mathbb{N}^*}$ and $(g_k)_{k\in\mathbb{N}^*}$ be Hilbert bases of $L^2_V(\mathbb{R}^n)$ and $L^2_W(\mathbb{R}^n)$, respectively. Then $(f_j\overline{g_k})_{j,k\in\mathbb{N}^*}$ is a Hilbert basis of $L^2_{V\times W}(\mathbb{R}^n)$.

Theorem 3.7 Let $\alpha \in L^2_{V\times W}(\mathbb{R}^n)$. Then, for all $f \in L^2_V(\mathbb{R}^n)$, the integral

$$F(y) := \int_V \alpha(x,y) f(x)\,\mathrm{d}x \tag{3.1}$$

is well-defined for almost all $y \in W$. The function F (defined almost everywhere in W) belongs to $L^2_W(\mathbb{R}^n)$. Moreover, the linear transformation

$$\begin{array}{rccl} A: & L^2_V(\mathbb{R}^n) & \longrightarrow & L^2_W(\mathbb{R}^n) \\ & f & \longmapsto & Af := F \end{array}$$

is bounded and satisfies $\|A\| \leq \|\alpha\|$.

The operator A defined in the previous theorem is called the *integral operator of kernel* α, and we write: $A = \operatorname{Int}\alpha$.

Theorem 3.8 Let $A \in L(L^2_V(\mathbb{R}^n), L^2_W(\mathbb{R}^n))$. The following statements are equivalent:

(1) A is a Hilbert-Schmidt operator;

(2) there exists $\alpha \in L^2_{V\times W}(\mathbb{R}^n)$ such that $A = \operatorname{Int}\alpha$.

Furthermore, we then have: $\|A\|_{HS} = \|\alpha\|$.

Theorem 3.9 Let $A \in L(L^2_V(\mathbb{R}^n), L^2_W(\mathbb{R}^n))$ be a Hilbert-Schmidt operator. Then A is compact.

Theorem 3.10 Let $\alpha \in L^2_{V\times W}(\mathbb{R}^n)$ and let $A \in L(L^2_V(\mathbb{R}^n), L^2_W(\mathbb{R}^n))$ be the integral operator of kernel α. Then $A^\star = \operatorname{Int}\alpha^\star$ where $\alpha^\star(y,x) = \overline{\alpha(x,y)}$.

Chapter 4

Convolution and approximation

4.1 Reminder of useful inequalities

Theorem 4.1 [Hölder's Inequality] Let p, q be conjugate exponents, with $p \in (1, \infty)$ (and thus $q \in (1, \infty)$). Let $(X, \mathcal{A}, \mu)$ be a measure space and let $f, g \colon X \to [0, \infty]$ be measurable. Then

$$\int_X f g \, \mathrm{d}\mu \leq \left(\int_X f^p \, \mathrm{d}\mu \right)^{1/p} \left(\int_X g^q \, \mathrm{d}\mu \right)^{1/q}.$$

Theorem 4.2 [Minkowski's Inequality] Let $p \in (1, \infty)$. Let $(X, \mathcal{A}, \mu)$ be a measure space and let $f, g \colon X \to [0, \infty]$ be measurable. Then

$$\left(\int_X (f + g)^p \, \mathrm{d}\mu \right)^{1/p} \leq \left(\int_X f^p \, \mathrm{d}\mu \right)^{1/p} + \left(\int_X g^p \, \mathrm{d}\mu \right)^{1/p}$$

Theorem 4.3 [Minkowski's Integral Inequality] Let $p \in [1, \infty)$. Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, and let $f : X \times Y \to \mathbb{R}^+$ be measurable. Then

$$\left(\int_X \left(\int_Y f \, \mathrm{d}\nu \right)^p \mathrm{d}\mu \right)^{1/p} \leq \int_Y \left(\int_X f^p \, \mathrm{d}\mu \right)^{1/p} \mathrm{d}\nu. \tag{4.1}$$

4.2 Convolution of L^p functions

Theorem 4.4 Let $f \in L^1(\mathbb{R}^n)$ and $g \in L^p(\mathbb{R}^n)$, where $p \in [1, \infty]$. Then, the function $\varphi_x : y \to f(x - y) g(y)$ is integrable for almost all $x \in \mathbb{R}^n$. The function $f * g$ defined for almost all x by

$$(f * g)(x) = \int f(x-y)g(y)\,\mathrm{d}y$$

belongs to $L^p(\mathbb{R}^n)$ and satisfies $\|f * g\|_p \leq \|f\|_1 \|g\|_p$.

Given any function f, on $\mathbb{R}^n$, we shall denote by f_x the *translate* of f by x: $f_x(y) := f(y-x)$.

Theorem 4.5 Let $f \in L^p(\mathbb{R}^n)$, where $p \in [1, \infty)$. Then $\|f_x - f\|_p \to 0$ as $\|x\| \to 0$.

Theorem 4.6 Let $\varphi \in L^1(\mathbb{R}^n)$ and $a := \int \varphi(x)\,\mathrm{d}x$. For all $\varepsilon > 0$, let φ_ε be defined by

$$\varphi_\varepsilon(x) = \frac{1}{\varepsilon^n}\varphi\left(\frac{x}{\varepsilon}\right).$$

For all $f \in L^p(\mathbb{R}^n)$, where $p \in [1, \infty)$, one has

$$\|f * \varphi_\varepsilon - af\|_p \longrightarrow 0 \quad \text{as} \quad \varepsilon \longrightarrow 0.$$

Theorem 4.7 Let $\psi \colon \mathbb{R}^n \to \mathbb{R}_+$ satisfying

(a) $\int \psi(x)\,\mathrm{d}x = 1$;

(b) for all $\eta > 0$, $\sup_{\|x\|>\eta} \psi_\alpha(x) \to 0$ as $\alpha \to 0$, in which $\psi_\alpha(x) := \alpha^{-n}\psi(x/\alpha)$.

Let $f \colon \mathbb{R}^n \to \mathbb{C}$ be integrable and continuous at x. Then $(f * \psi_\alpha)(x) \to f(x)$ as $\alpha \to 0$.

Remark 4.1 The function ψ given, for all $x \in \mathbb{R}^n$, by $\psi(x) = e^{-\pi\|x\|^2}$ satisfies the requirements of the last theorem, with $a = 1$.

The Schwartz space $\mathcal{S}(\mathbb{R}^n)$ is the set of all functions φ in $\mathcal{C}^\infty(\mathbb{R}^n)$ such that, for all $\alpha, \beta \in \mathbb{N}^n$,

$$\sup\left\{\left|x^\alpha D^\beta \varphi(x)\right| \,\middle|\, x \in \mathbb{R}^n\right\} < \infty$$

or, equivalently, such that $P(x)Q(D)\varphi(x)$ is bounded for all polynomials P and Q. Here, we have defined: $D := (\partial/\partial x_1, \ldots, \partial/\partial x_n)$. For all $p, q \in \mathbb{N}$,

$$\|f\|_{p,q} := \sup\left\{|(1+\|x\|)^p D^\alpha f(x)| \mid x \in \mathbb{R}^n,\ |\alpha| := \alpha_1 + \cdots + \alpha_n \leq q\right\}$$

is finite, and $\|\cdot\|_{p,q}$ is a norm on $\mathcal{S}(\mathbb{R}^n)$. A sequence $(f_k) \subset \mathcal{S}(\mathbb{R}^n)$ and a function $f \in \mathcal{S}(\mathbb{R}^n)$, we say that (f_k) *converges to f in $\mathcal{S}(\mathbb{R}^n)$* if

$$\forall p, q \in \mathbb{N}, \quad \|f_k - f\|_{p,q} \longrightarrow 0 \quad \text{as} \quad k \longrightarrow \infty.$$

The function $x \mapsto \exp(-\|x\|^2)$ is a well-known example of function in $\mathcal{S}(\mathbb{R}^n)$ which is not in $\mathcal{C}_0^\infty(\mathbb{R}^n)$, the space of all compactly supported infinitely differentiable functions on $\mathbb{R}^n$. For all $p \in [1, \infty]$, $\mathcal{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$.

Theorem 4.8 Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$ and $f \in L^p(\mathbb{R}^n)$, where $p \in [1, \infty)$. Then $f * \varphi \in \mathcal{C}^\infty(\mathbb{R}^n)$ and for all multi-index α,

$$D^\alpha(f * \varphi) = f * D^\alpha \varphi. \tag{4.2}$$

Theorem 4.9 Let $f, g \in \mathcal{C}_0(\mathbb{R}^n)$. Then $\operatorname{supp}((f * g)) \subset \operatorname{supp}(f) + \operatorname{supp}(g)$. In particular, $f * g$ has compact support.

Theorem 4.10 The space $\mathcal{C}_0^\infty(\mathbb{R}^n)$ of all compactly supported $\mathcal{C}^\infty$-functions on $\mathbb{R}^n$ is dense in $L^p(\mathbb{R}^n)$ for all $p \in [1, \infty)$.

Corollary 4.1 The space $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for all $p \in [1, \infty)$.

The latter density result can be strengthened. The following theorem will show that functions which belong to several L^p-spaces can be approximated by $\mathcal{C}_0^\infty$-functions with respect to all the L^p-norms in question, *simultaneously.*

Theorem 4.11 Let f be a measurable function. Then there exists a sequence $(f_k) \subset \mathcal{C}_0^\infty(\mathbb{R}^n)$ such that

$$\forall p \in [1, \infty) \colon f \in L^p(\mathbb{R}^n), \ \|f_n - f\|_p \longrightarrow 0.$$

We now give a generalization of Theorem 4.4.

Theorem 4.12 Let $f \in L^p(\mathbb{R}^n)$ and $g \in L^q(\mathbb{R}^n)$, where $p, q \in [1, \infty]$ are such that $r^{-1} := p^{-1} + q^{-1} - 1 \in [0, 1]$. Then, the function $\varphi_x \colon y \to f(x-y)g(y)$ is integrable for almost all $x \in \mathbb{R}^n$. Furthermore, the function h defined for almost all x by

$$h(x) := (f * g)(x) = \int f(x-y)g(y)\,\mathrm{d}y$$

belongs to $L^r(\mathbb{R}^n)$ and satisfies $\|h\|_r \leq \|f\|_p \|g\|_q$.

Chapter 5

Fourier transforms

5.1 Fourier transforms of integrable functions

Let $f \in L^1(\mathbb{R}^n)$. The function $x \mapsto e^{-2i\pi\langle x,\xi\rangle} f(x)$ is, needless to say, integrable for all $\xi \in \mathbb{R}^n$. We then define the functions $\hat{f}$ and $\check{f}$ by

$$\hat{f}(\xi) := \int e^{-2i\pi\langle x,\xi\rangle} f(x)\,\mathrm{d}x \quad \text{and} \quad \check{f}(\xi) := \int e^{2i\pi\langle x,\xi\rangle} f(x)\,\mathrm{d}x, \quad x \in \mathbb{R}^n.$$

The functions $\hat{f}$ and $\check{f}$ will also be denoted by Uf and $\overline{U}f$, respectively. It is clear that $\check{f}(\xi) = \hat{f}(-\xi)$. For that reason, every statement concerning $\hat{f}$ gives rise to a similar statement about $\check{f}$, which will be assumed tacitly in the sequel.

Clearly, the transformation U acts linearly on $L^1(\mathbb{R}^n)$:

$$\forall \alpha, \beta \in \mathbb{C}, \quad \forall f, g \in L^1, \quad U(\alpha f + \beta g) = \alpha U f + \beta U g.$$

Theorem 5.1 Let $f, g \in L^1(\mathbb{R}^n)$. Then $f * g \in L^1(\mathbb{R}^n)$ (by Theorem 4.4), and

$$U(f * g) = Uf \cdot Ug.$$

Theorem 5.2 Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Then

(1) for all multi-index α, $UD^\alpha\varphi(\xi) = (2i\pi\xi)^\alpha \cdot U\varphi(\xi)$;

(2) for all multi-index β, $D^\beta U\varphi(\xi) = U\big((-2i\pi x)^\beta\varphi\big)(\xi)$, in which $(-2i\pi x)$ stands for the function $x \mapsto (-2i\pi x)$;

(3) $U\varphi \in \mathcal{S}(\mathbb{R}^n)$.

A function h on $\mathbb{R}^n$ is said to *vanish at infinity* if $h(x) \to 0$ as $\|x\| \to \infty$.

Exercise 5.1 Let (h_n) be a sequence of functions on $\mathbb{R}^n$ which converges uniformly to a function h. Show that if h_n vanishes at infinity for all n, then so does h.

Theorem 5.3 [Riemann-Lebesgue] Let $f \in L^1(\mathbb{R}^n)$. Then

(1) $\hat{f}$ is a bounded function, and $\|\hat{f}\|_\infty \leq \|f\|_1$;

(2) $\hat{f}$ is continuous and vanishes at infinity.

We shall denote by $\Gamma(\mathbb{R}^n)$ the set of all continuous functions which vanish at infinity. Members of $\Gamma(\mathbb{R}^n)$ are, of course, measurable functions. As such, each member of $\Gamma(\mathbb{R}^n)$ identified to its *almost everywhere*-class. However, given $\varphi \in \Gamma(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$, the expression $\varphi(x)$, which is meaningless if φ is considered as a class, will always be understood as the value at x of the original (continuous) member of the class. Conversely, if φ is *a priori* a member of some L^p-space, a statement such as $\varphi \in \Gamma(\mathbb{R}^n)$ will be understood as: 'φ is almost everywhere equal to a continuous function on $\mathbb{R}^n$ which vanishes at infinity.'

Exercise 5.2 For all function f on $\mathbb{R}^n$, let $T_y f$, $M_y f$ (where $y \in \mathbb{R}^n$) and $S_a f$ (where $a \in \mathbb{R}^*$) be defined by

$$(T_y f)(x) = f(x-y), \quad (M_y f)(x) = e^{-2i\pi\langle x,y\rangle} f(x), \quad (S_a f)(x) = f(ax).$$

Let $f \in L^1(\mathbb{R}^n)$. Prove that, for all $\xi \in \mathbb{R}^n$,

(1) $UT_y f(\xi) = M_y U f(\xi)$;

(2) $UM_y f(\xi) = T_{-y} U f(\xi)$;

(3) $US_a f(\xi) = |a|^{-n} S_{1/a} U f(\xi)$.

Exercise 5.3 Let $\varphi \colon \mathbb{R} \to \mathbb{R},\ x \mapsto e^{-\pi x^2}$.

(1) Show that $\int \varphi(x)\,\mathrm{d}x = 1$;

(2) Prove that $\hat{\varphi}$ satisfies the differential equation $y'(\xi) = -2\pi\xi y(\xi)$;

(3) Deduce that $\hat{\varphi} = \varphi$ and that, if $\psi \colon \mathbb{R}^n \to \mathbb{R},\ x \mapsto e^{-\pi\|x\|^2}$, then $\hat{\psi} = \psi$.

Theorem 5.4 Let $f, g \in L^1(\mathbb{R}^n)$. Then $\hat{f}g$ and $f\hat{g}$ are integrable, and

$$\int \hat{f}(x)g(x)\,\mathrm{d}x = \int f(x)\hat{g}(x)\,\mathrm{d}x.$$

Theorem 5.5 Let $f \in \mathcal{S}(\mathbb{R}^n)$. Then, for all $x \in \mathbb{R}^n$,

$$f(x) = \int e^{2i\pi\langle x,\xi\rangle}\hat{f}(\xi)\,\mathrm{d}\xi.$$

Otherwise expressed, $U\colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ is bijective, and $U^{-1} = \overline{U}$.

Remark 5.1 The mapping

$$\begin{array}{rccc} \langle\cdot,\cdot\rangle : & \mathcal{S}(\mathbb{R}^n)\times\mathcal{S}(\mathbb{R}^n) & \longrightarrow & \mathbb{C} \\ & (f,g) & \longmapsto & \displaystyle\int f\overline{g} \end{array}$$

turns $\mathcal{S}(\mathbb{R}^n)$ into an inner product space. Relatively to this inner product, U and U^{-1} are adjoint to each other, since

$$\begin{aligned} \langle Uf, g\rangle &= \int \left(\int e^{-2i\pi\langle x,\xi\rangle} f(x)\right)\overline{g}(\xi)\,\mathrm{d}\xi \\ &= \int f(x)\left(\int e^{-2i\pi\langle x,\xi\rangle}\overline{g}(\xi)\,\mathrm{d}\xi\right)\mathrm{d}x \\ &= \int f(x)\overline{\int e^{2i\pi\langle x,\xi\rangle} g(\xi)\,\mathrm{d}\xi}\,\mathrm{d}x \\ &= \left\langle f, U^{-1}g\right\rangle, \end{aligned} \tag{5.1}$$

in which the second equality results from Fubini's Theorem. ∎

Recall that, if $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is such that $\int f(x)\varphi(x)\,\mathrm{d}x = 0$ for all $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, then $f = 0$ almost everywhere. This will allow us to prove the following Theorem.

Theorem 5.6 Let $f \in L^1(\mathbb{R}^n)$ be such that $\hat{f} \in L^1(\mathbb{R}^n)$. Then $f = \overline{U}\hat{f}$ almost everywhere.

The following approximation result will prove to be useful later on.

Theorem 5.7 Let $f \in L^1(\mathbb{R}^n)$ be such that $\hat{f} \in L^1(\mathbb{R}^n)$. Let $\psi \in L^1(\mathbb{R}^n)$ be such that $\int \psi(x)\,\mathrm{d}x = 1$, and let ψ_α be defined by

$$\psi_\alpha(x) = \frac{1}{\alpha^n}\psi\left(\frac{x}{\alpha}\right), \quad x \in \mathbb{R}^n.$$

Then $f * \psi_\alpha \to f$ uniformly on $\mathbb{R}^n$ as $\alpha \to 0$.

Our aim is now to show that the convolution of two functions of $\mathcal{S}(\mathbb{R}^n)$ belongs to $\mathcal{S}(\mathbb{R}^n)$ (see Theorem 5.8 below). In order to achieve this goal, we shall establish a few technical results.

Lemma 5.1 [Multinomial Formula] Let $n \in \mathbb{N}^*$, let $a := (a_1, \ldots, a_n) \in \mathbb{C}^n$, and let $N \in \mathbb{N}$. Then

$$(a_1 + \cdots + a_n)^N = \sum_{|\alpha|=N} \frac{N!}{\alpha!} a^\alpha,$$

in which, as usual, $|\alpha| := \alpha_1 + \cdots + \alpha_n$ and $\alpha! := \alpha_1! \cdots \alpha_n!$.

Lemma 5.2 For all $x \in \mathbb{R}^n$ and all $\alpha \in \mathbb{N}^n$, $|x^\alpha| \leq \|x\|^{|\alpha|}$.

Lemma 5.3 Let $\gamma\colon [0, \infty]$ be any function. The following statements are equivalent:

(i) $\sup_{x \in \mathbb{R}^n} \big\{ |x^\alpha| \, \gamma(x) \big\} < \infty$ for all $\alpha \in \mathbb{N}^n$;

(ii) $\sup_{x \in \mathbb{R}^n} \big\{ \|x\|^N \gamma(x) \big\} < \infty$ for all $N \in \mathbb{N}$.

Theorem 5.8 Let $f, g \in \mathcal{S}(\mathbb{R}^n)$. Then $f * g \in \mathcal{S}(\mathbb{R}^n)$.

Remark 5.2 Let $\cdot$ denote the product $\mathbb{R} \times \mathcal{S}(\mathbb{R}) \to \mathcal{S}(\mathbb{R}^n)$, $(\lambda, f) \mapsto \lambda f$. Then $(\mathcal{S}(\mathbb{R}^n), +, \cdot)$ is obviously a vector space. Theorem 5.8 shows that the convolution product $*$ is internal for $\mathcal{S}(\mathbb{R}^n)$. Furthermore, the following properties hold:

(1) $\forall f, g_1, g_2 \in \mathcal{S}(\mathbb{R}^n)$, $f * (g_1 + g_2) = f * g_1 + f * g_2$ and
$\forall f_1, f_2, g \in \mathcal{S}(\mathbb{R}^n)$, $(f_1 + f_2) * g = f_1 * g + f_2 * g$;

(2) $\forall \lambda \in \mathbb{C}$, $\forall f, g \in \mathcal{S}(\mathbb{R}^n)$, $\lambda(f * g) = (\lambda f) * g = f * (\lambda g)$.

All these properties turn $(\mathcal{S}(\mathbb{R}^n), +, \cdot, *)$ into an algebra, which is commutative and associative, since

(3) $\forall f, g \in \mathcal{S}(\mathbb{R}^n)$, $f * g = g * f$;

(4) $\forall f, g, h \in \mathcal{S}(\mathbb{R}^n)$, $f * (g * h) = (f * g) * h$.

Note that Property (4) is immediate from Theorem 5.1 and the associativity of the standard product of functions. Finally, since the standard product of functions, denoted here by $\times$, is *dual* of $*$ by the Fourier transformation, and since the Fourier transformation is a bijection of $\mathcal{S}(\mathbb{R}^n)$ onto itself (Theorem 5.5), $(\mathcal{S}(\mathbb{R}^n), +, \cdot, \times)$ is also a commutative and associative algebra.

5.2 The Fourier operator on $L^2(\mathbb{R}^n)$

We shall now define the Fourier operator on $L^2(\mathbb{R}^n)$, as an extention of the Fourier transformation on the Schwartz space.

Proposition 5.1 Let F, G be Banach spaces whose norms are both denoted by $\|\cdot\|$, and let E be a dense subspace of F. Let $A\colon E \to G$ be a linear mapping such that:

$$\exists k, K > 0\colon \forall x \in E,\ k\,\|x\| \leq \|Ax\| \leq K\,\|x\|\,. \tag{5.2}$$

Then, there exists a unique continuous linear mapping $\mathbb{A}\colon F \to G$ also denoted by $\operatorname{cl} A$, whose restriction to E coincides with A. Furthermore, the following holds:

(1) $\forall x \in F,\ k\|x\| \leq \|\mathbb{A}x\| \leq K\|x\|$;

(2) the range of $\mathbb{A}$ is the closure of the range of A;

(3) $\mathbb{A}\colon F \to \operatorname{ran}\mathbb{A}$ is bijective and bicontinuous, and $(\operatorname{cl} A)^{-1} = \operatorname{cl}(A^{-1})$.

Theorem 5.9 The Fourier transformation $U\colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ can be extended to a unique continuous linear mapping $\mathbb{U}\colon L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ such that, for all $f \in L^2(\mathbb{R}^n)$, $\|\mathbb{U}f\| = \|f\|$ (i.e. $\mathbb{U}$ is an isometry).

The fact that $\mathbb{U}^{-1} = \operatorname{cl}\overline{U} = S_{-1}U$ is straightforward. Note that the Fourier operator $\mathbb{U}$ turns out to be a Hilbert space isomorphism, since, by Proposition 2.1,

$$\forall f, g \in L^2(\mathbb{R}^n),\ \langle f, g\rangle = \langle \mathbb{U}f, \mathbb{U}g\rangle\,.$$

It should be stressed that, for functions in $L^2(\mathbb{R}^n)$ which are not integrable, the Fourier integral

$$\int e^{-2i\pi\langle x,\xi\rangle} f(x)\,\mathrm{d}x$$

may be undefined, as a Lebesgue integral. What is then the explicit link between the Fourier operator $\mathbb{U}$ and the previously defined integral transformation ? In the particular case where $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, do we have $\mathbb{U}f = \hat{f}$? The following theorem adresses these questions.

Theorem 5.10 (1) Suppose that $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Then $\mathbb{U}f = \hat{f}$, in which $\hat{f}$ is identified with its class in $L^2(\mathbb{R}^n)$.

(2) Suppose now that f is any member of $L^2(\mathbb{R}^n)$, and let Φ_R and F_R be defined, for all $R > 0$, by

$$\Phi_R(\xi) := \int_{\|x\| \leq R} e^{-2i\pi\langle x,\xi\rangle} f(x)\,\mathrm{d}x$$

$$\text{and} \quad F_R(x) := \int_{\|\xi\| \leq R} e^{2i\pi\langle x,\xi\rangle} (\mathbb{U}f)(\xi)\,\mathrm{d}\xi.$$

Then $\|\Phi_R - \mathbb{U}f\|_2 \to 0$ and $\|F_R - f\|_2 \to 0$ as $R \to \infty$.

The use distinct notation for U and $\mathbb{U}$ is superfluous in general, and both linear transformations will be denoted by U in the sequel.

5.3 Compactly supported functions

Functions with compact support have the important property that their Fourier transforms are analytic (see Theorem 5.20 below). We start with a survey of n-dimensional complex analysis.

Let $\mathbb{K}$ be $\mathbb{R}$ or $\mathbb{C}$, and let F be a Banach space on $\mathbb{K}$. Let Ω be an open subset of $\mathbb{K}^n$, where $n \in \mathbb{N}^*$, and let $x_0 \in \Omega$. A function $f\colon \Omega \to F$ is said to be $\mathbb{K}$-*differentiable at* x_0 if there exists a linear mapping $L\colon \mathbb{K}^n \to F$ such that

$$\|f(x) - f(x_0) - L(x - x_0)\|_F = o(\|x - x_0\|). \tag{5.3}$$

In the particular case where $\mathbb{K} = \mathbb{C}$, $n = 1$ and $F = \mathbb{C}$, the variable of f is denoted by z, and the complex number

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0},$$

whose existence is provided by (5.3), is called the $\mathbb{C}$-derivative of f at z_0. We say that f is *holomorphic at* z_0 if f is $\mathbb{C}$-differentiable at every point of some neighborhood of z_0, and that f is holomorphic on Ω if f is $\mathbb{C}$-differentiable at every point of Ω. The set of all functions which are holomorphic on Ω is denoted by $H(\Omega)$. We say that f is entire if $f \in H(\mathbb{C})$.

On writing $z = x + iy$ with $(x, y) \in \mathbb{R}^2$, every function $f(z)$ can be regarded as a function on $\mathbb{R}^2$. Its partial derivatives with respect to x and y are denoted, when they exist, by f_x and f_y, or by $\partial f/\partial x$ and $\partial f/\partial y$, respectively:

$$f_x(z_0) = f_x(x_0 + iy_0) = \lim_{x \to x_0} \frac{f(x + iy_0) - f(x_0 + iy_0)}{x - x_0}$$

$$\text{and} \quad f_y(z_0) = f_y(x_0 + iy_0) = \lim_{y \to y_0} \frac{f(x_0 + iy) - f(x_0 + iy_0)}{y - y_0}.$$

Theorem 5.11 Let Ω be an open subset of $\mathbb{C}$, and let $f : \Omega \to \mathbb{C}$.

(1) If f is $\mathbb{C}$-differentiable at $z_0 \in \Omega$, then $f_x(z_0)$ and $f_y(z_0)$ satisfy the Cauchy-Riemann condition: $f_y(z_0) = i f_x(z_0)$.

(2) Conversely, suppose that f_x and f_y exist in some open neighborhood of z_0, that f_x and f_y are continuous in it, and that $f_y = i f_x$. Then f is $\mathbb{C}$-differentiable at z_0.

Let again $\mathbb{K}$ be $\mathbb{R}$ or $\mathbb{C}$, and let F be a Banach space on $\mathbb{K}$. An *n-dimensional power series* (or merely a *power series*) is a family of monomials $(c_k x^k)_{k \in \mathbb{N}^n}$ in which $c_k \in F$ for all k and $x = (x_1, \ldots, x_n) \in \mathbb{K}^n$. The series $(c_k x^k)_{k \in \mathbb{N}^n}$ is also denoted by $\sum c_k x^k$. The *set of summability* of $(c_k x^k)_{k \in \mathbb{N}^n}$ is, by definition, the set of all $x \in \mathbb{K}^n$ such that the family is summable. For such values of x, the sum is denoted by $\sum_{k \in \mathbb{N}^n} c_k x^k$. Given $x \in \mathbb{K}^n$ and $r \in (\mathbb{R}_+^*)^n$, the set

$$D(x, r) := \{u \in \mathbb{K}^n \mid \forall j \in \{1, \ldots, n\}, |u_j - x_j| < r_j\}$$

is referred to as the *open polydisc of radius r centered at x*.

Theorem 5.12 Let $\sum c_k x^k$ be a power series, and let Ω be the interior of its set of summability, which is assumed to be nonempty. For all $x \in \Omega$, let

$$f(x) := \sum_{k \in \mathbb{N}^n} c_k x^k.$$

Then, for all $\alpha \in \mathbb{N}^n$, the power series

$$\left(D^\alpha (c_k x^k)\right)_{k \geq \alpha} = \left(\frac{k!}{(k - \alpha)!} c_k x^{k-\alpha}\right)_{k \geq \alpha}$$

is summable at every $x \in \Omega$. Moreover, f is infinitely $\mathbb{K}$-differentiable on Ω, and for all $\alpha \in \mathbb{N}^n$ and all $x \in \Omega$,

$$\frac{1}{\alpha!} D^\alpha f(x) = \sum_{k \geq \alpha} C_k^\alpha c_k x^{k-\alpha},$$

in which $C_k^\alpha = C_{k_1}^{\alpha_1} \ldots C_{k_n}^{\alpha_n}$. In particular, $(1/\alpha!) D^\alpha f(0) = c_\alpha$ for all $\alpha \in \mathbb{N}^n$, so that f determines uniquely the family $(c_k)_{k \in \mathbb{N}^n}$.

Theorem 5.13 Let $\sum c_k x^k$, Ω and f be as in the previous theorem. Then, for all $\alpha \in \mathbb{N}^n$, the power series

$$\left(\frac{k!}{(k+\alpha)!} c_k x^{k+\alpha}\right)_{k \in \mathbb{N}^n}$$

is summable at every $x \in \Omega$. Its sum $g(x)$ satisfies $D^\alpha g = f$. Furthermore, if $r \in (\mathbb{R}_+^*)^n$ is such that $D(0,r) \subset \Omega$ and if $x_0 \in D(0,r)$, then the power series

$$\left(\frac{1}{k!} D^k f(x_0)(x - x_0)^k\right)_{k \in \mathbb{N}^n}$$

is summable to $f(x)$ for all $x \in \mathbb{K}^n$ such that $|x_j - x_{0j}| < r_j - |x_{0j}|$ for all $j \in \{1, \ldots, n\}$.

Let now $f \colon \Omega \to F$, and let $x_0 \in \Omega$. We say that $\sum c_k (x - x_0)^k$ is a *power series development of* f *about* x_0 if there exists a neighborhood V of x_0 in Ω such that $\sum c_k (x - x_0)^k$ is summable to f for all $x \in V$. We say that f is *analytic* if, for all $x_0 \in \Omega$, f has a power series development about x_0. The set of all function $f \colon \Omega \to F$ which are analytical is denoted by $A(\Omega, F)$.

Theorem 5.14 The set $A(\Omega, F)$ is a $\mathbb{K}$-vector space.

Theorem 5.15 Let $f \in A(\Omega, F)$. Then f is infinitely $\mathbb{K}$-differentiable, and $D^\alpha f \in A(\Omega, F)$ for all $\alpha \in \mathbb{N}^n$.

Theorem 5.16 The sum of a power series is analytical on the interior of its set of convergence.

Theorem 5.17 Let $f \in A(\Omega, F)$ and let $x_0 \in \Omega$. Then f has a unique power series development about x_0, namely,

$$\sum \frac{1}{k!} D^k f(x_0)(x - x_0)^k.$$

The following two theorems are of crucial importance.

Theorem 5.18 Let $f \in A(\Omega, F)$, where Ω is connected, and let y be in the range of F. If f is not identically equal to y on Ω, then $f^{-1}(y)$ has empty interior.

Corollary 5.1 Let $f, g \in A(\Omega, F)$, where Ω is connected. If f and g coincide on an open subset Ω_1 of Ω, then they coincide on Ω.

Theorem 5.19 [Osgood] Let Ω be an open subset of $\mathbb{C}^n$ and let F be a Banach space on $\mathbb{C}$. Let $f\colon \Omega \to F$. The following statements are equivalent:

(1) $f \in A(\Omega, F)$;

(2) f is $\mathbb{C}$-differentiable;

(3) f is continuous on Ω and, at every $z = (z_1, \ldots, z_n)$ in Ω, the partial derivatives $\partial f/\partial z_1, \ldots, \partial f/\partial z_n$ exist.

Remark 5.3 According to a Theorem due to Hartog, the continuity assumption in Condition (3) of the above theorem can be omitted. ∎

Let now $f \in L^1_V(\mathbb{R}^n)$, where V is compact. For all $\zeta = \xi + i\eta$ in $\mathbb{C}^n$, the function $x \mapsto e^{-2i\pi\langle x,\xi\rangle} e^{2\pi\langle x,\eta\rangle} f(x)$ is integrable. The function

$$\check{f}(\zeta) = \check{f}(\xi + i\eta) := \int e^{-2i\pi\langle x,\xi\rangle} e^{2\pi\langle x,\eta\rangle} f(x)\,\mathrm{d}x$$

coincides with $\hat{f}$ for real values of the argument.

Theorem 5.20 Let V be a compact subset of $\mathbb{R}^n$ and let $f \in L^1_V(\mathbb{R}^n)$. Then $f \in A(\mathbb{C}^n, \mathbb{C})$.

Corollary 5.2 Let V be a compact subset of $\mathbb{R}^n$ and let $f \in L^1_V(\mathbb{R}^n)$. Then $f \in A(\mathbb{R}^n, \mathbb{C})$. Consequently, if Ω is an open subset of $\mathbb{R}^n$, the mapping

$$\begin{array}{rccl} A\colon & L^1_V(\mathbb{R}^n) & \longrightarrow & L^2_\Omega(\mathbb{R}^n) \\ & f & \longmapsto & Af := \mathbb{1}_\Omega \hat{f} \end{array}$$

is injective.

Chapter 6

Ill-posedness and regularization

We now return to the study inverse problems. Following the reformulation heuristics of Section 1.2, we first consider the least square solutions of the problem.

6.1 Least squares

Proposition 6.1 Let F and G be Hilbert spaces and let $A\colon F \to G$ be an operator. Then

(1) $\ker A$ and $\operatorname{cl}\operatorname{ran} A^\star$ form an orthogonal decomposition of F.

(2) $\ker A^\star$ and $\operatorname{cl}\operatorname{ran} A$ form an orthogonal decomposition of G.

Exercise 6.1 Prove that $\ker A^\star A = \ker A$ and $\ker AA^\star = \ker A^\star$.

Theorem 6.1 Let F and G be Hilbert spaces and let $A\colon F \to G$ be an operator. Let P denote the orthogonal projection onto $\operatorname{cl}\operatorname{ran} A$. Let $y \in G$ and let $\tilde{y} := Py$. Then the following statements are equivalent:

(1) $\tilde{y} = Ax_0$;

(2) x_0 minimizes $\|y - Ax\|^2$;

(3) x_0 solves the *normal equation*: $A^\star y = A^\star A x_0$.

Note that if $y \in \operatorname{ran} A + (\operatorname{ran} A)^{\perp}$, then $Py \in \operatorname{ran} A$. In this case, $A^{-1}(Py)$ is nonempty. According to the previous theorem, it is given by

$$A^{-1}(Py) = \{x \in F \mid A^{\star}y = A^{\star}Ax\}.$$

Therefore, $A^{-1}(Py)$ is an affine manifold parallel to $\ker A^{\star}A = \ker A$.

Theorem 6.2 Let $y \in \operatorname{ran} A + (\operatorname{ran} A)^{\perp}$. Then there exists a unique $x^{+} \in \operatorname{cl} \operatorname{ran} A^{\star}$ minimizing $\|\cdot\|$ over $A^{-1}(Py)$. Moreover, $A^{-1}(Py) = \{x^{+}\} + \ker A$, and the mapping

$$\begin{array}{rccl} A^{+}\colon & \mathcal{D}(A^{+}) & \longrightarrow & F \\ & y & \longmapsto & A^{+}y := x^{+}, \end{array}$$

in which $\mathcal{D}(A^{+}) := \operatorname{ran} A + (\operatorname{ran} A)^{\perp}$, is linear. It is referred to as the *generalized inverse* (or *pseudo-inverse*) of A.

Remark 6.1 If A is injective, then so is $A^{\star}A$ (see Exercise 6.1), and $A^{\star}A\colon F \to \operatorname{ran} A^{\star}A$ is invertible. In this case, for all $y \in \operatorname{ran} A + (\operatorname{ran} A^{\perp}$, $(A^{\star}A)^{-1}A^{\star}y$ is the unique solution to the normal equation, so that

$$A^{-1}(Py) = \{(A^{\star}A)^{-1}A^{\star}y\}.$$

In the case where A is invertible, then $A^{+} = A^{-1}$. ∎

The preceeding discussion shows that the reformulation of the inverse problem under consideration which consists in searching for 'the' minimum norm least square solution may not be successful. As a matter of fact, the latter may not be defined for every $y \in F$: if the range of A happens to be not closed, $A^{-1}(\{\tilde{y}\})$ may be empty. Furthermore, the linear mapping A^{+} may not be continuous, which will result in the violation of the stability condition. In the next section, we study a class of inverse problems that are subject to both pathologies. The problem of *Fourier synthesis*, which is central to image science, pertains to this class.

6.2 A class of ill-posed problems

To begin with, let us recall the so-called Open Mapping Theorem. For all $r > 0$, let B_r denote the open ball of radius r centered at the origin.

Theorem 6.3 Let F, G be Banach spaces and let $A\colon F \to G$ be a surjective operator. Then

$$\exists c > 0\colon AB_1 \supset B_c. \tag{6.1}$$

In particular, if A is bijective, then A^{-1} is continuous. As a matter of fact, Condition (6.1) then says that

$$\forall y \in G, \quad \|y\| < 1 \Longrightarrow \left\|A^{-1} y\right\| < c^{-1}.$$

Theorem 6.4 Let F and G be separable Hilbert spaces, where $\dim F = +\infty$ and let $A\colon F \to G$ be an injective Hilbert-Schmidt operator. Then

(1) $A^{-1}\colon \operatorname{ran} A \to F$ is not continuous;

(2) $\Lambda(A^\star A)$ is a subset of $\mathbb{R}_+^*$ and has 0 as accumulation point;

(3) $\operatorname{ran} A$ is not closed, so that $\mathcal{D}(A^+) = \operatorname{ran} A + (\operatorname{ran} A)^\perp \subsetneq G$.

More precisely, the spectral theorem says that the Hilbert basis exhibited in the above proof can chosen in such a way that

$$\forall k \in \mathbb{N}^*, \quad A^\star A f_k = \lambda_k f_k,$$

in which the sequence (λ_k) is positive, decreasing and converging to 0. We may then consider the following system of equations:

$$\begin{cases} g_k & := & \dfrac{1}{\sqrt{\lambda_k}} A f_k, \\ f_k & = & \dfrac{1}{\sqrt{\lambda_k}} A^\star g_k, \end{cases}$$

in which the second equation is immediate from the first one. The above system is referred to as a *Singular Value Decompostion* of A (in short, SVD), and the numbers $\sqrt{\lambda_k}$ are called the *singular values* of A.

Proposition 6.2 The sequence (g_k) forms a Hilbert basis of $\operatorname{cl} \operatorname{ran} A$.

Remark 6.2 In Theorem 6.4, we have seen that the range of A fails to be closed. We can easily exhibit vectors g in $\operatorname{cl} \operatorname{ran} A \setminus \operatorname{ran} A$. Since A is a Hilbert-Schmidt opeartor,

$$\sum_k \|A f_k\|^2 = \sum_k \langle f_k, A^\star A f_k \rangle = \sum_k \lambda_k < \infty.$$

If (β_k) is any sequence such that $|\beta_k|^2 = \lambda_k$, then clearly $g := \sum_k \beta_k g_k$ belongs to $\operatorname{cl} \operatorname{vect}(g_k) = \operatorname{cl} \operatorname{ran} A$. But $\sum_k |\beta_k|^2 / \lambda_k = \infty$, which shows that $g \notin \operatorname{ran} A$. ∎

Theorem 6.5 Let F, G and A be as in Theorem 6.4, and consider the above Singular Value Decomposition of A. Then

(1) $\forall f \in F$, $Af = \sum_k \lambda_k^{1/2} \langle f, f_k \rangle g_k$;

(2) $\forall g \in G$, $A^\star g = \sum_k \lambda_k^{1/2} \langle g, g_k \rangle f_k$;

(3) $\forall g \in \mathcal{D}(A^+)$, $A^+ g = \sum_k \lambda_k^{-1/2} \langle g, g_k \rangle f_k$.

Corollary 6.1 Under the previous assumptions, A^+ is unbounded.

6.3 Regularization of A^+

Let again F and G be separable Hilbert spaces with $\dim F = +\infty$ and $A\colon F \to G$ be an injective operator, such that $\operatorname{tr} A^\star A < \infty$. In the preceeding section, it appeared that the corresponding inverse problem together with its 'minimum-norm-least-square' reformulation were ill-posed. In order to cope with these difficulties, we introduce the notion of regularization of A^+.

Definition 6.1 A family $(T_\alpha)_{\alpha > 0}$ of continuous linear applications from G to F is called a *regularization of* A^+ if

$$\forall g \in \mathcal{D}(A^+), \qquad \|T_\alpha g - A^+ g\| \longrightarrow 0 \quad \text{as} \quad \alpha \longrightarrow 0.$$

Exercise 6.2 Prove that $\|T_\alpha\| \longrightarrow \infty$ as $\alpha \longrightarrow 0$.

Proposition 6.3 Suppose that $(g_\varepsilon)_{\varepsilon > 0}$ is such that $\|g_\varepsilon - g\| \leq \varepsilon$. Then there exists a function $\alpha(\varepsilon)$ such that

$$\|T_{\alpha(\varepsilon)} g_\varepsilon - A^+ g\| \longrightarrow 0 \quad \text{as} \quad \varepsilon \longrightarrow 0.$$

Exercise 6.3 Define (T_α) by

$$T_\alpha g := \sum_{k \leq \frac{1}{\alpha}} \frac{1}{\sqrt{\lambda_k}} \langle g, g_k \rangle f_k.$$

(1) Prove that, for all $\alpha > 0$, T_α is continuous.

(2) Compute the adjoint $T_\alpha^\star$ of T_α, the spectrum of $T_\alpha^\star T_\alpha$. Deduce the value of $\|T_\alpha\|$.

(3) Show that, for all $g \in \mathcal{D}(A^+)$, $T_\alpha g$ tends to A^+g as α tends to 0. Conclude.

Another example of regularization of A^+ is provided by Tikhonov's regularization principle.
Recall that A is assumed to be injective. therefore,if $g \in \mathcal{D}(A^+) = \operatorname{ran} A + (\operatorname{ran} A)^\perp$, then A^+g is the unique solution to the normal equation: $A^\star g = A^\star A f$. Since $\ker A = \ker A^\star A = \{0\}$, the linear application

$$A^\star A \colon F \longrightarrow \operatorname{ran} A^\star A$$

is invertible. Consequently, $A^\star g \in \operatorname{ran} A^\star A$ and

$$A^+ g = (A^\star A)^{-1} A^\star g.$$

Clearly, $(A^\star A)^{-1}$ is not continuous, for otherwise A^+ would be continuous. The spectrum of $A^\star A$ (is a subset of $\mathbb{R}_+^*$ and) has 0 as accumulation point. A rather natural idea is then to 'shift' the spectrum to the right, so that it accumulates to a positive number. This amounts to define

$$T_\alpha := (A^\star A + \alpha I)^{-1} A^\star.$$

Obviously, for all $\alpha > 0$, T_α is bounded.

Theorem 6.6 The family T_α defined above is a regularization of A^+. It is called the Tikhonov regularization.

Exercise 6.4 Show that $T_\alpha g$ is the unique minimizer of the functional

$$F(f) := \frac{1}{2}\|g - Af\|^2 + \frac{\alpha}{2}\|f\|^2.$$

Chapter 7

Fourier Synthesis

7.1 Fourier extrapolation

Let V and W be bounded subsets of $\mathbb{R}^n$, where W is assumed to contain some open subset, and let

$$\begin{array}{rccl} A\colon & L^2_V(\mathbb{R}^n) & \longrightarrow & L^2_W(\mathbb{R}^n) \\ & f & \longmapsto & Af := \mathbb{1}_W \hat{f}. \end{array}$$

More explicitly, the function Af is given by

$$\begin{aligned} (Af)(y) &= \int_V e^{-2i\pi\langle x,y\rangle} f(x)\,\mathrm{d}x \cdot \mathbb{1}_W(y) \\ &= \int e^{-2i\pi\langle x,y\rangle} \mathbb{1}_V(x)\mathbb{1}_W(y) f(x)\,\mathrm{d}x. \end{aligned}$$

Since V and W are bounded, the kernel $\alpha(x,y) := e^{-2i\pi\langle x,y\rangle}\mathbb{1}_V(x)\mathbb{1}_W(y)$ is in $L^2(\mathbb{R}^{2n})$. Therefore, A is a Hilbert-Schmidt operator. By Theorem 3.10, the adjoint of A is given by

$$\begin{aligned} A^\star g(x) &= \int \overline{\alpha(x,y)} g(y)\,\mathrm{d}y \\ &= \int e^{2i\pi\langle x,y\rangle} \mathbb{1}_V(x)\mathbb{1}_W(y) g(y)\,\mathrm{d}y \\ &= \mathbb{1}_V(x) \cdot \int_W e^{2i\pi\langle x,y\rangle} g(y)\,\mathrm{d}y. \end{aligned}$$

Since g vanishes (almost everywhere) outside W, we see that $A^\star = \mathbb{1}_V \overline{U} = \mathbb{1}_V U^{-1}$. Thus $A^\star A = \mathbb{1}_V U^{-1} \mathbb{1}_W U$. Furthermore, since $L^2_V(\mathbb{R}^n) \subset L^1_V(\mathbb{R}^n)$, Theorem 5.2 implies that A is injective.

Consequently, the problem of *Fourier extrapolation* is ill-posed:

(1) $A^{-1}\colon \operatorname{ran} A \to L^2_V(\mathbb{R}^n)$ exists but it is not continuous;

(2) the range of A is not closed, so that the domain of A^+ is a proper subset of $L^2_W(\mathbb{R}^n)$;

(3) $A^+\colon \operatorname{ran} A + (\operatorname{ran} A)^\perp \to L^2_V(\mathbb{R}^n)$ is unbounded.

Proposition 7.1 The largest eigenvalue λ_1 of $A^\star A$ satisfies $\lambda_1 < 1$, so that $\Lambda(A^\star A)$ is actually contained in $(0,1)$ and $\|A^\star A\| < 1$.

We may, of course, adopt one of the strategies of Section 6.3 to regularize this problem. The study of the problem of Fourier interpolation will suggest another approach.

7.2 Fourier interpolation

Let again V be a bounded subset of $\mathbb{R}^n$, and let W be a subset of $\mathbb{R}^n$ such that $W^c := \mathbb{R}^n \setminus W$ is bounded and contains an open subset. Let

$$\begin{array}{rccl} B\colon & L^2_V(\mathbb{R}^n) & \longrightarrow & L^2_W(\mathbb{R}^n) \\ & f & \longmapsto & Bf := 1\!\!1_W \hat{f}. \end{array}$$

Then $B^\star = 1\!\!1_V U^{-1}$ and

$$B^\star B = 1\!\!1_V U^{-1} 1\!\!1_W U = 1\!\!1_V U^{-1}(I - 1\!\!1_{W^c})U = I - 1\!\!1_V U^{-1} 1\!\!1_{W^c} U,$$

where I denotes the identity of $L^2_V(\mathbb{R}^n)$. On denoting by A the operator $1\!\!1_{W^c} U$, we see that $B^\star B = I - A^\star A$, where $\|A^\star A\| \le 1$ by Proposition 7.1 (since W^c satisfies the assumptions of the set W involved in the problem of Fourier extrapolation).

Theorem 7.1 Let F be a Banach space and let $T \in L(F)$ be such that $\|T\| < 1$. Then $\operatorname{ran}(I - T) = F$, the inverse of $I - T$ exists and belong to $L(F)$, and

$$(I - T)^{-1} = \sum_{k=0}^{\infty} T^k$$

(in which the series converges in the topology of the norm of $L(F)$). Furthermore,

$$\left\|(I - T)^{-1}\right\| \le \frac{1}{1 - \|T\|}.$$

The last theorem shows that $B^\star B$ has a continuous inverse, and that

$$\left\|(B^\star B)^{-1}\right\| \le \frac{1}{1-\|A^\star A\|} = \frac{1}{1-\lambda_1}.$$

Remark 7.1 The operator $B\colon L^2_V(\mathbb{R}^n) \to \operatorname{ran} B \subset L^2_W(\mathbb{R}^n)$ is injective and surjective, so that $B^{-1}\colon \operatorname{ran} B \to L^2_V(\mathbb{R}^n)$ exists. However, $B\colon L^2_V(\mathbb{R}^n) \to L^2_W(\mathbb{R}^n)$ is not surjective and thus not invertible. As a matter of fact, for all $f \in L^2_V(\mathbb{R}^n)$, $\hat{f}$ is analytic on any open subset Ω of W, and we can certainly find $g \in L^2_W(\mathbb{R}^n)$ such that the restriction of g to Ω is not analytic. ∎

Proposition 7.2 With the previous notation and assumptions,

(1) $B^{-1}\colon \operatorname{ran} B \to L^2_V(\mathbb{R}^n)$ is continuous;

(2) $\operatorname{ran} B$ is a closed subset of $L^2_W(\mathbb{R}^n)$;

(3) $\Lambda(B^\star B)$ can be arranged to form an increasing sequence $\mu_1 \le \mu_2 \le \dots$ such that $\mu_1 > 0$ and $\mu_k \to 1$ as $k \to \infty$.

As a conclusion, the inverse problem of Fourier interpomation fails to be well-posed only by the fact that B is not surjective. Its reformulation in terms of least squares is well-posed. In other words, the equation

$$B^\star B f = B^\star g$$

has a unique solution for all $g \in L^2_W(\mathbb{R}^n)$, which depends continuously on g.

7.3 Fourier regularization

In order to regularize the problem of Fourier extrapolation, one may invoke Tikhonov's regularization principle. The last paragraph suggests another approach.

Let $W_\beta := \mathbb{R}^n \setminus \beta^{-1}B$, where B denotes the unit ball centered at the origin. Since W is bounded, $W \cap W_\beta = \emptyset$ for β small enough. In any case, $(W \cup W_\beta)^c$ is a bounded subset of $\mathbb{R}^n$. We then define

$$\begin{array}{rccl} B_\beta\colon & L^2_V(\mathbb{R}^n) & \longrightarrow & L^2_{W\cup W_\beta}(\mathbb{R}^n) \\ & f & \longmapsto & B_\beta f := \mathbb{1}_{W\cup W_\beta}\hat{f}, \end{array}$$

and $T_\beta = (B_\beta^\star B_\beta)^{-1} B_\beta^\star$. For all $g \in L^2_V(\mathbb{R}^n)$, $T_\beta g$ minimize the functional

$$\begin{aligned}
\mathcal{F}_\beta(f) &= \frac{1}{2}\|g - B_betaf\|^2 \\
&= \frac{1}{2}\int_{W\cup W_\beta} |g(y) - (B_\beta f)(y)|^2 \,\mathrm{d}y \\
&= \frac{1}{2}\int_{W} |g(y) - (B_\beta f)(y)|^2 \,\mathrm{d}y + \frac{1}{2}\int_{W_\beta} |(B_\beta f)(y)|^2 \,\mathrm{d}y \\
&= \frac{1}{2}\|g - Af\|^2 + \frac{1}{2}\|C_\beta f\|^2,
\end{aligned}$$

in which

$$\begin{aligned}
C_\beta\colon \quad L^2_V(\mathbb{R}^n) &\longrightarrow L^2_{W_\beta}(\mathbb{R}^n) \\
f &\longmapsto \mathbb{1}_{W_\beta}\hat{f}.
\end{aligned}$$

Furthermore, using the obvious identification of $L_{W\cup W_\beta}(\mathbb{R}^n)$ with $L^2_W(\mathbb{R}^n)\times L_{W_\beta}(\mathbb{R}^n)$, we can write

$$\begin{aligned}
B_\beta\colon \quad L^2_V(\mathbb{R}^n &\longrightarrow L^2_W(\mathbb{R}^n)\times L_{W_\beta}(\mathbb{R}^n) \\
f &\longmapsto (Af, C_\beta f).
\end{aligned}$$

We then have:

$$\begin{aligned}
\langle B_\beta f, (g,h)\rangle &= \langle (Af, C_\beta f), (g,h)\rangle \\
&= \langle Af, g\rangle + \langle C_\beta f, h\rangle \\
&= \langle f, A^\star g\rangle + \langle f, C_\beta^\star h\rangle \\
&= \left\langle f, A^\star g + C_\beta^\star h\right\rangle.
\end{aligned}$$

Thus $B_\beta^\star(g,f) = A^\star g + C_\beta^\star h$ for all $(g,h)\in L^2_W(\mathbb{R}^n)\times L_{W_\beta}(\mathbb{R}^n)$, so that

$$B_\beta^\star B_\beta f = B_\beta^\star(Af, C_\beta g) = A^\star A f + C_\beta^\star C_\beta f = (A^\star A + C_\beta^\star C_\beta) f.$$

This shows that $T_\beta = (A^\star A + C_\beta^\star C_\beta)^{-1}A^\star$. This makes the comparison with Tikhonov's regularization easier. Recall that the corresponding family of operators is given by

$$T_\alpha = (A^\star A + \alpha I)^{-1}A^\star.$$

Notice that Tikhonov's regularization amounts to the minimization of

$$\mathcal{F}_\alpha(f) := \frac{1}{2}\|g - Af\|^2 + \frac{\alpha}{2}\|f\|^2 = \frac{1}{2}\|g - Af\|^2 + \frac{\alpha}{2}\|Uf\|^2.$$

We see that Tikhonov's regularization term acts everywhere in the Fourier domain. The restriction of $\hat{f}$ to W is requested to fit to g by the first component of $\mathcal{F}_\alpha$ and to 0 by the second component. Thus, in this particular case, Tikhonov's regularization introduces constraints which are in conflict with the *experimental constraints.*

7.4 Deconvolution

In essence, the problem of deconvolution is the same as that of Fourier synthesis.

Exercise 7.1 Let V be a bounded subset of $\mathbb{R}^n$, and let $k \in L^2(\mathbb{R}^n) \cap L^1(\mathbb{R}^n)$. Let U denote the Fourier operator in $L^2(\mathbb{R}^n)$. We assume that $\operatorname{supp}(Uk)$ contains an open set.

(1) Show that

$$\begin{array}{rccc} A\colon & L^2_V(\mathbb{R}^n) & \longrightarrow & L^2(\mathbb{R}^n) \\ & f & \longmapsto & Af := k * f \end{array}$$

is a well-defined operator, and that $\|A\| \leq \|k\|_1$.

(2) Show that $A = \overline{U} A_0$, where $A_0\colon L^2_V(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is an operator to be specified.

(3) Show that A_0 and A are Hilbert-Schmidt operators.

(4) Show that A is injective. Conclude.

The reader is invited to write explictly the Fourier regularization of the problem.

7.5 Sampling theorems

Recall that the sinc function is defined, in dimension 1, by

$$\operatorname{sinc} x := \begin{cases} \sin x / x & \text{if } x \neq 0, \\ 1 & \text{if } x \neq 0. \end{cases}$$

If $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, we define: $\operatorname{sinc} x := \operatorname{sinc} x_1 \ldots \operatorname{sinc} x_n$. For all $r > 0$, we define:

$$B_r := \{\xi \in \mathbb{R}^n \mid \|\xi\| \leq r/2\} \quad \text{and} \quad C_r := [-r/2, r/2]^n.$$

Let $a > 0$. For all $k = (k_1, \dots, k_2) \in \mathbb{Z}^n$, we define the function φ_k by:

$$\varphi_k(\xi) = \begin{cases} a^{-n/2} e^{-2i\pi\langle \xi, k/a\rangle} & \text{if } \xi \in C_a, \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to check that $(\varphi_k)_{k\in\mathbb{Z}^n}$ is an orthonormal set in the Hilbert space $L^2_{C_a}(\mathbb{R}^n)$. We shall admit the following result, which is a consequence of the Stone-Weierstrass Theorem.

Theorem 7.2 The set $(\varphi_k)_{k\in\mathbb{Z}^n}$ is a Hilbert basis of $L^2_{C_a}(\mathbb{R}^n)$.

Definition 7.1 A function $f \in L^2(\mathbb{R}^n)$ is said to be *band-limited* if there exists $b > 0$ such that its Fourier transform $\hat{f}$ vanishes (almost everywhere) outside B_b. The smallest such b is then called the *band-width* of f.

Notice that a band-limited function must be analytical, since it is the (inverse) Fourier transform of a compactly supported function.

Example 7.1 Let $n = 1$. The function $x \to \operatorname{sinc} \pi x$ is band-limited, with band-width equal to 1. As a matter of fact,

$$\operatorname{sinc} \pi x = \int_{-1/2}^{1/2} e^{2i\pi\xi x}\, \mathrm{d}\xi = U^{-1}\mathbb{1}_{[-\frac{1}{2},\frac{1}{2}]}(x). \blacksquare$$

Lemma 7.1 For all $a > 0$, $U^{-1}\mathbb{1}_{C_a}(x) = a^n \operatorname{sinc} \pi a x$.

Theorem 7.3 [Shannon's Sampling Theorem] Let $f \in L^2(\mathbb{R}^n)$ be band-limited, with band-width $b \in (0, a]$, where a is some positive number. Then f is entirely determined by the samples $f(k/a)$, $k \in \mathbb{Z}^n$. More precisely, f is given by the following interpolation formula:

$$f(x) = \sum_{k\in\mathbb{Z}^n} f\left(\frac{k}{a}\right) \operatorname{sinc} \pi a \left(x - \frac{k}{a}\right). \tag{7.1}$$

Furthermore, $\hat{f}$ is given on C_a by

$$\hat{f}(\xi) = \frac{1}{a^n} \sum_{k\in\mathbb{Z}^n} f\left(\frac{k}{a}\right) \exp\left[-2i\pi \left\langle \frac{k}{a}, \xi \right\rangle\right], \tag{7.2}$$

and if g is another band-limited function with band-width b, then

$$\langle f, g\rangle := \int f(x)\overline{g}(x)\,\mathrm{d}x = \frac{1}{a^n} \sum_{k\in\mathbb{Z}^n} f\left(\frac{k}{a}\right) \overline{g}\left(\frac{k}{a}\right). \tag{7.3}$$

In the case where b is strictly less that a, the rate of convergence of the series in Equation (7.1) may be improved, as shown in the following theorem.

Theorem 7.4 Let $f \in L^2(\mathbb{R}^n)$ be band-limited, with band-width $b \in (0, a)$, where a is some positive number. Let γ be a C^∞-function such that

(i) $\gamma(\xi) = 0$ if $\|\xi\| \geq (a-b)/2$;

(ii) $\int \gamma(\xi)\,\mathrm{d}\xi = 1$.

Then f is given by the following interpolation formula:

$$f(x) = \sum_{k \in \mathbb{Z}^n} f\left(\frac{k}{a}\right) \check{\gamma}\left(x - \frac{k}{a}\right) \operatorname{sinc} \pi a \left(x - \frac{k}{a}\right). \tag{7.4}$$

Chapter 8

The Radon transformation

8.1 Definition and basic properties

We first define the Radon transform of functions in the Schwartz space. Extension to Sobolev spaces will be considered later on.
From now on, we assume that $n \geq 2$. Let S^{n-1} denote the unit sphere of dimension $n-1$:

$$S^{n-1} := \{\theta \in \mathbb{R}^n \mid \|\theta\| = 1\}.$$

For all $(\theta, r) \in Z := S^{n-1} \times \mathbb{R}$, let $H(\theta, r)$ denote the hyperplane orthogonal to θ through $r\theta$:

$$H(\theta, r) := \{x \in \mathbb{R}^n \mid \langle \theta, x \rangle = r\}.$$

Every $x \in \mathbb{R}^n$ can be written as $x = \langle \theta, x \rangle \theta + v$, where v belongs to the orthogonal complement of $\langle \theta \rangle := \{u\theta | u \in \mathbb{R}\}$, that is, to $H(\theta, 0)$.
The *Radon transform* of a function $f \in \mathcal{S}(\mathbb{R}^n)$ is the function Rf defined on $Z := S^{n-1} \times \mathbb{R}$ by

$$(Rf)(\theta, r) := \int_{H(\theta,0)} f(r\theta + v)\, \mathrm{d}v. \tag{8.1}$$

We shall call θ and r the *angular* and *radial* variables, respectively. Note that the integral in Equation (8.1) is convergent since f is assumed to lie in the Schwartz space. Clearly, the Radon transformation acts linearly on $\mathcal{S}(\mathbb{R}^n)$.
When dealing with integrals, we shall often use the following change of variable:

$$\begin{cases} x = r\theta + v \\ r = \langle \theta, x \rangle \\ \mathrm{d}x = \mathrm{d}r\, \mathrm{d}v. \end{cases} \tag{8.2}$$

For example, the integral of a function $g \in \mathcal{S}(\mathbb{R}^n)$ may be written as:

$$\int_{\mathbb{R}^n} g(x)\,\mathrm{d}x = \int_{\mathbb{R}} \left(\int_{H(\theta,0)} g(r\theta + v)\,\mathrm{d}v \right) \mathrm{d}r.$$

For all $\theta \in S^{n-1}$, the function $(Rf)(\theta, \cdot)$ is integrable. In fact,

$$\int_{\mathbb{R}} (Rf)(\theta, r)\,\mathrm{d}r = \int_{\mathbb{R}} \left(\int_{H(\theta,r)} f(r\theta + v)\,\mathrm{d}v \right) \mathrm{d}r = \int_{\mathbb{R}^n} f(x)\,\mathrm{d}x$$

by Fubini's Theorem. We may then consider the Fourier transform of $(Rf)(\theta, \cdot)$. We shall denote by U_r the Fourier transformation relative to the radial variable, and by ρ the *dual* variable of r:

$$(U_r Rf)(\theta, \rho) := \int_{\mathbb{R}} e^{-2i\pi r\rho} (Rf)(\theta, r)\,\mathrm{d}r.$$

The following theorem establishes the connection between The Radon and Fourier transforms.

Theorem 8.1 [Fourier Slice Theorem] Let $f \in \mathcal{S}(\mathbb{R}^n)$. Then,

$$\forall (\theta, \rho) \in Z, \quad (U_r Rf)(\theta, \rho) = \hat{f}(\rho\theta).$$

Corollary 8.1 Let $f \in \mathcal{S}(\mathbb{R}^n)$. Then $(Rf)(\theta, \cdot) \in \mathcal{S}(\mathbb{R})$.

Corollary 8.2 Let $f, g \in \mathcal{S}(\mathbb{R}^n)$. Then

$$R(f * g) = Rf \circledast Rg,$$

in which $\circledast$ denotes convolution with respect to the radial variable.

Remark 8.1 In tomography, an approximation $(Rf)(\theta, s)$ is available for $\theta \in \{\theta_i\}_{i \in I}$ and $s \in \{s_j\}_{j \in J}$, where I and J are finite. Assume continuous knowledge of $(Rf)(\theta_i, s)$ for each i. By the Fourier slice theorem, this knowledge is equivalent to the knowledge of $\hat{f}$ on the lines $\mathbb{R}\{\theta_i\}$, $i \in I$. Therefore, the reconstruction of f can once again be regarded as a problem of Fourier Synthesis. In practice, due to the resolution limit inherent to all experimental device, the knowledge of $\hat{f}$ is also limited in diameter. In any case, one is facing a problem of Fourier extrapolation. Therefore, one should give up reconstructing f, and aim at reconstructing a smoothed version of f: $f * \kappa$, where κ is some convolution kernel (or *point spread function*).

Proposition 8.1 Let $f \in \mathcal{S}(\mathbb{R}^n)$. Then,

(1) Rf is an even function;

(2) for all $N \in \mathbb{N}$, $\int_{\mathbb{R}} s^N (Rf)(\theta, r)\,\mathrm{d}r = P_N(\theta)$, where P_N is a homogeneous polynomial of degree N.

Exercise 8.1 Let $f \in \mathcal{S}(\mathbb{R}^n)$ and let f_{x_0} be defined by $f_{x_0}(x) := f(x - x_0)$. Show that $(Rf_{x_0})(\theta, r) = (Rf)(\theta, r - \langle \theta, x_0 \rangle)$.

Proposition 8.2 Let $f \in \mathcal{S}(\mathbb{R}^n)$, let $\varphi(x) := e^{-\pi x^2}$ and let $\varphi_\alpha(x) := \alpha^{-1}\varphi(\alpha^{-1}x)$ for all $\alpha > 0$. Let

$$(R_\alpha f)(\theta, r) := \int_{\mathbb{R}^n} f(x)\varphi_\alpha(r - \langle \theta, x \rangle)\,\mathrm{d}x.$$

Then $(R_\alpha f)(\theta, r) \to (Rf)(\theta, r)$ for all $(\theta, r) \in Z$, the convergence being uniform in r.

In Proposition 8.2, φ_α appears as an approximation of the Dirac *delta function*. This is why one sometimes write:

$$(Rf)(\theta, s) = \int f(x)\delta(s - \langle \theta, x \rangle)\,\mathrm{d}x.$$

Proposition 8.2 also suggests to extend Rf from Z to $\mathbb{R}_+^* Z = \mathbb{R}_+^* S^{n-1} \times \mathbb{R}$ as follows: for all $t \in \mathbb{R}_+^*$, we may write

$$\begin{aligned} (Rf)(t\theta, tr) &= \lim_{\alpha \to 0} \int_{\mathbb{R}^n} f(x)\varphi_\alpha(tr - \langle t\theta, x \rangle)\,\mathrm{d}x \\ &= \lim_{\alpha \to 0} \int_{\mathbb{R}^n} f(x) t^{-1} \varphi_{t^{-1}\alpha}(s - \langle \theta, x \rangle)\,\mathrm{d}x. \\ &= t^{-1}(Rf)(\theta, r). \end{aligned} \tag{8.3}$$

Notice that this extension, which is assumed throughout, is positively homogeneous of degree -1.

Exercise 8.2 Let M be an invertible $(n \times n)$-matrix, and let $f \in \mathcal{S}(\mathbb{R}^n)$. Let f_M denote the function given by $f_M(x) = f(M^{-1}x)$. Prove that

$$(Rf_M)(\theta, r) = |\det M|\,(Rf)(M^t\theta, r).$$

8.2 Radon transformation and differentiation

Theorem 8.2 Let $f \in \mathcal{S}(\mathbb{R}^n)$. Then, for all $\alpha \in \mathbb{N}^n$ and all $(\theta, r) \in Z$,

$$(RD^\alpha f)(\theta, r) = \theta^\alpha \left(D_r^{|\alpha|} Rf \right)(\theta, r),$$

where D_r^N denotes the differential operator of order N with respect to the second argument.

Our aim is now to obtain formulae for the derivatives of Rf with respect to $\theta \in \mathbb{R}_+^* S^{n-1}$.

Lemma 8.1 Let $f \in \mathcal{S}(\mathbb{R}^n)$ and let $\theta \in S^{n-1}$. Then $(R_\alpha f)(\theta, \cdot) \in \mathcal{S}(\mathbb{R})$ and, for all $N \in \mathbb{N}$, $(D_r^N R_\alpha f)(\theta, \cdot)$ converges to $(D_r^N Rf)(\theta, \cdot)$ uniformly as α tends to 0, where D_r^N denotes the differential operator $\partial^N / \partial r^N$.

Remark 8.2 By homogeneity, Lemma 8.1 remains true for all $\theta \in \mathbb{R}_+^* S^{n-1}$.

Lemma 8.2 Let $f \in \mathcal{S}(\mathbb{R}^n)$. Then $(R_\alpha f)(\cdot, s) \in \mathcal{C}^\infty(\mathbb{R}_+^* S^{n-1})$, and we have, for all $\beta \in \mathbb{N}^n$ and all $(\theta, r) \in \mathbb{R}_+^* Z$,

$$(D_\theta^\beta R_\alpha f)(\theta, r) = (-1)^{|\beta|} D_r^{|\beta|} R_\alpha (x^\beta f)(\theta, r),$$

where D_θ^β denotes the differential operator $\partial^{|\beta|} / \partial \theta_1^{\beta_1} \ldots \partial \theta_n^{\beta_n}$, and where $x^\beta f$ denotes the function $x \mapsto x^\beta f(x)$.

Theorem 8.3 Let $f \in \mathcal{S}(\mathbb{R}^n)$. Then $(Rf)(\cdot, s) \in \mathcal{C}^\infty(\mathbb{R}^n \setminus \{0\})$, and for all $\beta \in \mathbb{N}^n$ and all $(\theta, r) \in \mathbb{R}_+^* Z$,

$$(D_\theta^\beta Rf)(\theta, r) = (-1)^{|\beta|} D_r^{|\beta|} R(x^\beta f)(\theta, r).$$

From the preceeding developments, the Radon transform of a function $f \in \mathcal{S}(\mathbb{R}^n)$ satisfies:

(1) $\forall \theta \in \mathbb{R}_+^* S^{n-1}$, $(Rf)(\theta, \cdot) \in \mathcal{S}(\mathbb{R})$;

(2) $\forall r \in \mathbb{R}$, $(Rf)(\cdot, r) \in \mathcal{C}^\infty(\mathbb{R}_+^* S^{n-1})$.

We are led to define the space $\mathcal{S}(Z)$ as the sapce of all functions $g \colon Z \to \mathbb{C}$ which are restrictions od functions $g_e \colon \Omega \times \mathbb{R} \to \mathbb{C}$, where Ω is an open set containing S^{n-1}, such that

(1) $\forall \theta \in \Omega$, $g(\theta, \cdot) \in \mathcal{S}(\mathbb{R})$;

(2) $\forall r \in \mathbb{R}$, $g(\cdot, r) \in \mathcal{C}^\infty(\mathbb{R}_+^* S^{n-1})$.

Clearly, $R\mathcal{S}(\mathbb{R}^n) \subset \mathcal{S}(Z)$. In the next section, we shall identify the subspace $R\mathcal{S}(\mathbb{R}^n)$.

8.3 The range of R

Definition 8.1 A function $g\colon Z \to \mathbb{R}$ is said to satisfy the *Helgason-Ludwig conditions* if

(1) g is even;

(2) for all $N \in \mathbb{N}$, $\int_{\mathbb{R}} r^N g(\theta, r)\,\mathrm{d}r = P_N(\theta)$, where P_N is a homogeneous polynomial of degree N.

The subspace of $\mathcal{S}(Z)$ of all functions which satisfy the Helgason-Ludwig conditions will be denoted by $\mathcal{S}_{HL}(Z)$.

From the previous section and from Proposition 8.1, we know that the Radon transformation maps $\mathcal{S}(\mathbb{R}^n)$ into $\mathcal{S}_{HL}(Z)$. We shall see in Theorem 8.4 below that R maps $\mathcal{S}(\mathbb{R}^n)$ onto $\mathcal{S}_{HL}(Z)$.

Proposition 8.3 Let Ω be an open subset of $\mathbb{R}^n$, with $n \geq 2$, and let $f\colon \Omega \to \mathbb{R}^m$ be a continuous function. Suppose that $\mathrm{d}^2 f$ exists on

$$\dot{B}(x_0, \varepsilon) := B(x_0, \varepsilon) = \{x \in \mathbb{R}^n \mid 0 < \|x - x_0\| < \varepsilon\},$$

and that $\|\mathrm{d}^2 f(z)\| \leq M$ for all $z \in \dot{B}(x_0, \varepsilon)$, where M is some positive constant. Then f is continuously differentiable on $B(x_0, \varepsilon)$.

Corollary 8.3 Let Ω be an open subset of $\mathbb{R}^n$, and let $x \in \Omega$. Suppose that $f\colon \Omega \to \mathbb{R}$ is continuous, that f is $\mathcal{C}^\infty$ in $\Omega \setminus \{x\}$, and that $D^\alpha f$ is bounded on $\Omega \setminus \{x\}$ for all $\alpha \in \mathbb{N}^n$. Then $f \in \mathcal{C}^\infty(\Omega)$.

Theorem 8.4 Let $g \in \mathcal{S}_{HL}(Z)$. Then, there exists a unique $f \in \mathcal{S}(\mathbb{R}^n)$ such that $g = Rf$. Furthermore, if $g \in \mathcal{S}_{HL}(Z)$ is such that

$$\forall r \in \mathbb{R} \setminus [-a, a], \quad g(\theta, r) = 0$$

(where a is some positive number), then $f(x) = 0$ for all $x \in \mathbb{R}^n$ such that $\|x\| \geq a$.

8.4 Backprojection

Let g be any function on Z. On denoting by $\mathrm{d}\theta$ the natural measure on S^{n-1}, let $R^\diamond g$ be defined on $\mathbb{R}^n$ by

$$(R^\diamond g)(x) := \int_{S^{n-1}} g(\theta, \langle \theta, x\rangle)\,\mathrm{d}\theta,$$

provided the integral is exists. Notice that, if $g = Rf$, $(R^\diamond g)(x)$ appears as the sum over all directions θ of the integrals of f over the hyperplane orthogonal to θ which contains x. This is why $R^\diamond$ is sometimes referred to as the *backprojection operator*. It is clear that $R^\diamond$ acts linearly on $\mathcal{S}(Z)$. The notation $R^\diamond$ is justified by the fact that R and $R^\diamond$ are *dual* to each other, in the following sense: for all $f \in \mathcal{S}(\mathbb{R}^n)$ and $g \in \mathcal{S}(Z)$,

$$\begin{aligned}
\int_Z (Rf)(\theta, r) g(\theta, r)\, \mathrm{d}r\, \mathrm{d}\theta &= \int_{S^{n-1}} \int_{\mathbb{R}} \int_{H(\theta,0)} f(r\theta + v) g(\theta, r)\, \mathrm{d}v\, \mathrm{d}r\, \mathrm{d}\theta \\
&= \int_{S^{n-1}} \int_{\mathbb{R}^n} f(x) g(\theta, \langle \theta, x\rangle)\, \mathrm{d}x \\
&= \int_{\mathbb{R}^n} f(x) \int_{S^{n-1}} g(\theta, \langle \theta, x\rangle)\, \mathrm{d}\theta\, \mathrm{d}x \\
&= \int_{\mathbb{R}^n} f(x) (R^\diamond g)(x)\, \mathrm{d}x.
\end{aligned}$$

Theorem 8.5 Let $f \in \mathcal{S}(\mathbb{R}^n)$ and let $g \in \mathcal{S}(Z)$. Then

$$(R^\diamond g) * f = R^\diamond (g \circledast (Rf)).$$

8.5 Inversion Formulae

Definition 8.2 Let $f \in \mathcal{S}(\mathbb{R}^n)$ and let $a < n$. The *Riesz potential* is the (well-defined) linear application given by

$$I^a f := U^{-1} \frac{1}{\|\xi\|^a} U f.$$

For functions $g \in \mathcal{S}(Z)$, the Riesz potential is defined by

$$I^a g = U_s^{-1} \frac{1}{\sigma^a} U_s g.$$

Theorem 8.6 Let $f \in \mathcal{S}(\mathbb{R}^n)$ and $a < n$. Then

$$f = \frac{1}{2} I^{-a} R^\star I^{a-(n-1)} R f.$$

In particular, if $a = 0$, we obtain $f = \frac{1}{2} R^\star I^{-(n-1)} R f$.

Remark 8.3 In the case where $n = 2$, we obtain

$$R^{-1} = R^\star K, \quad \text{where} \quad K := \frac{1}{2} I^{-1} = \frac{1}{4\pi} H D,$$

in which H denotes the Hilbert transform.

8.6 Extension to Sobolev spaces

Let $\alpha \geq 0$. The *Sobolev space* $H^\alpha(\mathbb{R}^n)$ is, by definition, the set

$$H^\alpha(\mathbb{R}^n) := \left\{ f \in L^2(\mathbb{R}^n) \,\middle|\, (1+\|\xi\|^2)^{\alpha/2}(Uf)(\xi) \in L^2(\mathbb{R}^n) \right\}.$$

Clearly, $H^0(\mathbb{R}^n) = L^2(\mathbb{R}^n)$. Increasing α amounts to imposing a stronger decay of the Fourier transform of f. The mapping

$$\begin{array}{rccl} \langle \cdot, \cdot \rangle_\alpha : & H^\alpha(\mathbb{R}^n)^2 & \longrightarrow & \mathbb{C} \\ & (f_1, f_2) & \longmapsto & \displaystyle\int_{\mathbb{R}^n} (Uf_1)(\xi)\overline{(Uf_2)(\xi)}(1+\|\xi\|^2)^\alpha \,\mathrm{d}\xi \end{array}$$

turns $H^\alpha(\mathbb{R}^n)$ into a Hilbert space. The corresponding norm is denoted by $\|\cdot\|_\alpha$. It is given by

$$\|f\|_\alpha = \left(\int_{\mathbb{R}^n} |(Uf)(\xi)|^2 \,(1+\|\xi\|^2)^\alpha \,\mathrm{d}\xi \right)^{1/2}.$$

Let B denote the closed unit ball. We shall denote by $H_0^\alpha(B)$ the closure of $\mathcal{C}_0^\infty(B) := \{f \in \mathcal{C}^\infty(\mathbb{R}^n) | \operatorname{supp}(f) \subset B\}$ in $H^\alpha(\mathbb{R}^n)$.
By analogy with the previous definitions, let

$$H^\alpha(Z) := \left\{ g \in L^2(Z) \,\middle|\, (1+\rho^2)^{\alpha/2}(U_r g)(\rho) \in L^2(Z) \right\}.$$

The equation

$$\langle g_1, g_1 \rangle_\alpha := \int_Z (U_r g_1)(\theta,\rho)\overline{(U_r g_2)(\theta,\rho)}(1+\rho^2)^\alpha \,\mathrm{d}\rho$$

defines an inner product in $H^\alpha(Z)$, and the corresponding norm is given by

$$\|g\|_\alpha = \left(\int_Z |(U_r g)(\theta,\rho)|^2 \,(1+\rho^2)^\alpha \,\mathrm{d}\rho \right)^{1/2}.$$

Let $Z_1 := S^{n-1} \times [-1,1]$. We shall denote by $H_0^\alpha(Z_1)$ the closure in $H^\alpha(Z)$ of $\mathcal{C}_0^\infty(Z_1) := \{g \in \mathcal{C}^\infty(Z) | \operatorname{supp}(g) \subset Z_1\}$, and by $H_{HL}^\alpha(Z_1)$ the closure in $H^\alpha(Z)$ of the space $\mathcal{C}_{HL}^\infty(Z_1) \subset \mathcal{C}_0^\infty(Z_1)$ of the functions satisfying the Helgason-Ludvig conditions.

Theorem 8.7 Let $\alpha \geq 0$. There exists positive constants c and C such that

$$\forall f \in \mathcal{C}_0^\infty(B), \quad c\,\|f\|_\alpha \leq \|Rf\|_{\alpha+(n-1)/2} \leq C\,\|f\|_\alpha.$$

Corollary 8.4 Let $\mathcal{S}_{HL}(Z_1) := \{g \in \mathcal{S}_{HL}(Z) | \operatorname{supp}(g) \subset Z_1\}$. The operator $R\colon \mathcal{C}_0^\infty(B) \to \mathcal{S}_{HL}(Z_1)$ can be extended to a unique operator, denoted likewise, and

$$R\colon H_0^\alpha(B) \longrightarrow H_{HL}^{\alpha+(n-1)/2}(Z_1)$$

is an isomorphism.