Introduction to Inverse Problems of Fourier Synthesis

Lecture notes

Pierre Maréchal

Contents

Chapter 1

Introduction

1.1 Inverse problems

Let F and G be normed spaces, and 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 be *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 which 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

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

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 always 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 which 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 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:

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

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 nonnegativity. 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 which 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 be some 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

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

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 $(e_\alpha)_{\alpha\in A} \subset F$ be 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 $\operatorname{vect}\{e_1, \ldots, e_n\}$.

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: $x = \displaystyle\sum_{\alpha \in A} x_\alpha$.

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

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

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 $x \in F$, $\|x\|^2 = \sum |\langle x, e_\alpha\rangle|^2$.

The latter equality is referred to as *Parseval's identity*.

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

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

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

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

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

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

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

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

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

Proof. Since A is hermitian, we have:

$$\begin{cases} \langle A(x+y), x+y\rangle &= \langle Ax, x\rangle + 2\mathcal{R}e(\langle Ax, y\rangle) + \langle Ay, y\rangle , \\ \langle A(x-y), x-y\rangle &= \langle Ax, x\rangle - 2\mathcal{R}e(\langle Ax, y\rangle) + \langle Ay, y\rangle . \end{cases}$$

Thus $4\mathcal{R}e(\langle Ax, y\rangle) = \langle A(x+y), x+y\rangle - \langle A(x-y), x-y\rangle$. Let $M := \sup_{\|x\|=1} |\langle Ax, x\rangle|$. It is clear that $M \leq \|A\|$. Let us prove the opposite inequality. For all $x \in F$, $|\langle Ax, x\rangle| \leq M\|x\|^2$, so that

$$\begin{aligned} 4\mathcal{R}e(\langle Ax, y\rangle) &\leq |\langle A(x+y), x+y\rangle| + |\langle A(x-y), x-y\rangle| \\ &\leq M\big(\|x+y\|^2 + \|x-y\|^2\big) \\ &= 2M\big(\|x\|^2 + \|y\|^2\big) \qquad \text{(see Proposition 2.1)}. \end{aligned}$$

Thus $\mathcal{R}e(\langle Ax, y\rangle) \leq M$ for all x, y of unit norm. Consequently, $|\langle Ax, y\rangle| \leq M$ for all x, y of unit norm. The choice $y = Ax/\|Ax\|$ then shows that $\|Ax\| \leq M$ for all x of unit norm, which says that $\|A\| \leq M$. ■

3.2 Compact operators

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 $(x_k) \subset F$, the sequence (Ax_k) 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 $AE \subset E$. A 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 be an *eigenvalue* of A, and the set of all eigenvalues of A is denoted by $\Lambda(A)$.

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 $E^\perp$;

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

(c) $\Lambda(A) \subset \mathbb{R}$;

(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* $E_\lambda := \{x \in F \mid Ax = \lambda x\}$ 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 which consists of eigenvectors of A.

PROOF. Assume that E is A-stable. Given $x \in E^\perp$, we see that

$$\forall y \in E, \quad \langle y, Ax\rangle = \langle Ay, x\rangle = 0$$

(since $Ay \in E$). Thus $Ax \in E^\perp$, and (a) is established. As for (b) it follows immediate from (a) and the fact that $\operatorname{cl} E = E^{\perp\perp}$. Next, let $\lambda \in \Lambda(A)$ and let x be a nonzero vector in the corresponding eigenspace. Then

$$\lambda \langle x, x\rangle = \langle Ax, x\rangle = \langle x, Ax\rangle = \overline{\lambda} \langle x, x\rangle ,$$

so that $\overline{\lambda} = \lambda$, which proves (c). Statement (d) is immediate from (a). Let $\lambda \in \Lambda(A)$, let $x \in \operatorname{cl} E_\lambda$ and let $(x_k) \subset E_\lambda$ converge to x. We have:

$$\|Ax - \lambda x\| \le \|Ax - Ax_k\| + \|\lambda x_k - \lambda x\| \le \big(\|A\| + \lambda\big) \|x_k - x\| .$$

Since $\|x_k - x\| \to 0$ as $k \to \infty$, we must have $Ax = \lambda x$, and (e) is established. Let λ_1, λ_2 be as in (f), and let $x_1 \in E_{\lambda_1}$, $x_2 \in E_{\lambda_2}$. Then

$$\lambda_1 \langle x_1, x_2\rangle = \langle Ax_1, x_2\rangle = \langle x_1, Ax_2\rangle = \lambda_2 \langle x_1, x_2\rangle .$$

Since $\lambda_1 \ne \lambda_2$, we must have $\langle x_1, x_2\rangle = 0$. Finally, let E be as in (g). From the theory of hermitian matrices, E contains an eigenvector x of unit norm. We may then consider the orthogonal complement of $\langle x\rangle$ in E, and proceed by induction to produce an orthonormal basis of E. ■

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

Proof. By Theorem 3.1, there exists a sequence (x_k) in the unit sphere such that $|\langle Ax_k, x_k\rangle| \to \|A\|$. Taking a subsequence if necessary, we can assume that $\langle Ax_k, x_k\rangle \to \alpha$, where $\alpha := \|A\|$ or $\alpha = -\|A\|$. We then have:

$$\begin{aligned} \|Ax_k - \alpha x_k\|^2 &= \|Ax_k\|^2 - 2\alpha \langle Ax_k, x_k\rangle + \alpha^2 \|x_k\|^2 \\ &\leq 2\alpha^2 - 2\alpha \langle Ax_k, x_k\rangle \\ &= 2\alpha(\alpha - \langle Ax_k, x_k\rangle). \end{aligned}$$

Since A is compact, we see that, taking a subsequence if necessary, (Ax_k) converges to some vector y. Since $\alpha - \langle Ax_k, x_k\rangle$ tends to 0, the above inequalities show that (αx_k) converges to y as well. If $\alpha = 0$, then $\|A\| = 0$ (that is to say, $A = 0$) and we are done. If $\alpha \neq 0$, then (x_k) converges to $x := y/\alpha$, and we have

$$\begin{aligned} \|Ax - \alpha x\| &\leq \|Ax - Ax_k\| + \|Ax_k - \alpha x_k\| + \|\alpha x_k - \alpha x\| \\ &\leq \big(\|A\| + |\alpha|\big) \|x - x_k\| + \|Ax_k - \alpha x_k\|, \end{aligned}$$

which shows that $Ax = \alpha x$. ∎

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

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

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

Proof. The space $F' := \operatorname{cl}\operatorname{vect}\{E_\lambda \mid \lambda \in \Lambda(A)\}$ is clearly A-stable. By Proposition 3.3(b), $F'^{\perp}$ is also A-stable. Thus A induces a compact hermitian operator on $F'^{\perp}$ which has no eigenvalue. Lemma 3.1 then implies that $F'^{\perp} = \{0\}$. Since F' is closed, $F' = F'^{\perp\perp} = F$. ∎

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 $\Lambda(A)$.

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

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

Proof. By Parseval's identity, we have:

$$\begin{aligned}\sum_{j\in\mathbb{N}^*} \|Af_j\|^2 &= \sum_{j\in\mathbb{N}^*}\left(\sum_{k\in\mathbb{N}^*} |\langle Af_j, g_k\rangle|^2\right) \\ &= \sum_{k\in\mathbb{N}^*}\left(\sum_{j\in\mathbb{N}^*} |\langle f_j, A^\star g_k\rangle|^2\right) \\ &= \sum_{k\in\mathbb{N}^*} \|A^\star g_k\|^2 . \blacksquare\end{aligned}$$

Theorem 3.6 shows that $\sum_j \|Af_j\|^2$ is independent of the choice of 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 $L^2_{V\times W}(\mathbb{R}^{2n})$.

Proof. By Fubini's Theorem, we have:

$$\begin{aligned}\langle f_j\overline{g}_k, f_l\overline{g}_m\rangle &= \int_{V\times W} f_j(x)\overline{g_k(y) f_l(x)} g_m(y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \int_V f_j(x)\overline{f_l(x)}\,\mathrm{d}x\cdot\int_W g_m(y)\overline{g_k(y)}\,\mathrm{d}y\\ &= \langle f_j, f_l\rangle\,\langle g_m, g_k\rangle\,.\end{aligned}$$

This shows that $(f_j\overline{g_k})_{j,k\in\mathbb{N}^*}$ is an orthonormal family. In order to show that the family $(f_j\overline{g_k})_{j,k\in\mathbb{N}^*}$ spans $L^2_{V\times W}(\mathbb{R}^{2n})$, we shall prove that, if $h\in L^2_{V\times W}(\mathbb{R}^{2n})$ is orthogonal to $\mathrm{vect}(f_j\overline{g_k})$, then necessarily $h=0$.
Let $h\in L^2_{V\times W}(\mathbb{R}^{2n})$. Clearly, $h(\cdot,y)\in L^2_V(\mathbb{R}^n)$ for almost all y. Thus

$$\varphi_j(y) := \int_V f_j(x)\overline{h(x,y)}\,\mathrm{d}x = \langle f_j, h(\cdot,y)\rangle \tag{3.1}$$

is defined for almost all y. The Schwarz inequality then says that $|\varphi_j(y)|^2\le\|f_j\|^2\|h(\cdot,y)\|^2$ for almost all y. Consequently,

$$\int_W|\varphi_j(y)|^2\,\mathrm{d}y\le\|f_j\|^2\cdot\int_W\left(\int_V|h(x,y)|^2\,\mathrm{d}x\right)\mathrm{d}y=\|f_j\|^2\,\|h\|^2\,,$$

so that $\varphi_j\in L^2_W(\mathbb{R}^n)$, and we have, by Parseval's identity,

$$\begin{aligned}\|\varphi_j\|^2 &= \sum_{k\in\mathbb{N}^*}|\langle g_k,\varphi_j\rangle|^2\\ &= \sum_{k\in\mathbb{N}^*}\left|\int_W g_k(y)\overline{\varphi_j(y)}\right|^2\\ &= \sum_{k\in\mathbb{N}^*}\left|\int_W g_k(y)\left(\int_V h(x,y)\overline{f_j(x)}\,\mathrm{d}x\right)\mathrm{d}y\right|^2\\ &= \sum_{k\in\mathbb{N}^*}\left|\int_{V\times W} h(x,y)\overline{f_j(x)\overline{g_k(y)}}\,\mathrm{d}x\,\mathrm{d}y\right|^2\\ &= \sum_{k\in\mathbb{N}^*}|\langle h, f_j\overline{g}_k\rangle|^2\,.\end{aligned}\tag{3.2}$$

Suppose now that $h\perp f_j\overline{g}_k$ for all $j,k\in\mathbb{N}^*$. Then $\|\varphi_j\|^2=0$ for all $j\in\mathbb{N}^*$ by Equation (3.2) above. Consequently, $\langle f_j, h(\cdot,y)\rangle=0$ for all $j\in\mathbb{N}^*$ and almost all y. Since (f_j) is a Hilbert basis of F, this shows that $h(\cdot, y) = 0$ in $L^2_V(\mathbb{R}^n)$ (for almost all y), that is, that $h = 0$ in $L^2_{V\times W}(\mathbb{R}^{2n})$. ■

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

$$F(y) := \int_V \alpha(x,y) f(x)\,\mathrm{d}x \tag{3.3}$$

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

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

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

Proof. It is clear that $\varphi(y) := \left(\int_V |\alpha(x,y)|^2\,\mathrm{d}x\right)^{1/2}$ is finite for almost all y, and that φ belongs to $L^2_W(\mathbb{R}^n)$. In fact, $\|\varphi\|^2 = \|\alpha\|^2$. Since $\alpha(\cdot, y) \in L^2_V(\mathbb{R}^n)$ for almost all y, we see that $F(y) := \langle \alpha(\cdot,y), \overline{f}\rangle$ is well-defined for almost all $y \in W$. Clearly, F is measurable, and the Schwarz inequality reads

$$|F(y)| \le \|\alpha(\cdot,y)\|\,\|f\| = \varphi(y)\,\|f\|\,.$$

Consequently, F is square integrable, and

$$\|F\|^2 = \int_W |F(y)|^2\,\mathrm{d}y \le \|f\|^2 \cdot \int_W \varphi^2(y)\,\mathrm{d}y = \|f\|^2\,\|\alpha\|^2\,. \ ■$$

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

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

Proof. Suppose $A = \operatorname{Int}\alpha$, where $\alpha \in L^2_{V\times W}(\mathbb{R}^{2n})$. By Fubini's theorem, we have, for all $j,k \in \mathbb{N}^*$,

$$\begin{aligned}\langle Af_j, g_k\rangle &= \int_W \left(\int_V \alpha(x,y) f_j(x)\,\mathrm{d}x\right) \overline{g}_k(y)\,\mathrm{d}y \\ &= \int_{V\times W} \alpha(x,y) f_j(x)\overline{g}_k(y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \left\langle \alpha, \overline{f}_j g_k\right\rangle .\end{aligned}$$

Now, Parseval's identity shows that

$$\begin{aligned}\sum_{j\in\mathbb{N}^*} \|Af_j\|^2 &= \sum_{j,k\in\mathbb{N}^*} |\langle Af_j, g_k\rangle|^2 \\ &= \sum_{j,k\in\mathbb{N}^*} \left|\left\langle \alpha, \overline{f}_j g_k\right\rangle\right|^2 \\ &= \|\alpha\|^2 .\end{aligned}$$

Thus (2) implies (1). Conversely, suppose that $\sum_j \|Af_j\|^2 = \sum_{j,k} |\langle Af_j, g_k\rangle|^2$ is finite. By Theorem 2.16, $\sum_{j,k}\langle Af_j, g_k\rangle \overline{f}_j g_k$ converges to some $\alpha \in L^2_{V\times W}(\mathbb{R}^{2n})$, which satisfies $\langle \alpha, \overline{f}_j g_k\rangle = \langle Af_j, g_k\rangle$. Let $A' := \operatorname{Int}\alpha$. By the first part of the proof, $\langle A'f_j, g_k\rangle = \langle \alpha, \overline{f}_j g_k\rangle$. Consequently, $\langle A'f_j, g_k\rangle = \langle Af_j, g_k\rangle$ for all $j,k \in \mathbb{N}^*$, so that $A'f_j = Af_j$ for all $j \in \mathbb{N}^*$, which implies that $A' = A$. Finally, the Hilbert-Schmidt norm of A satisfies:

$$\|A\|_{HS} = (\operatorname{tr} A^\star A)^{1/2} = \left(\sum_j \|Af_j\|^2\right)^{1/2} = \|\alpha\| . \blacksquare$$

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.

Proof. By the previous theorem, $A = \operatorname{Int}\alpha$, where $\alpha \in L^2_{V\times W}(\mathbb{R}^{2n})$. Let us write

$$\alpha := \sum_{j,l\in\mathbb{N}^*} \alpha_{jk}\overline{f}_j g_k \quad \text{and} \quad \alpha_p := \sum_{j,k\leq p} \alpha_{jk}\overline{f}_j g_k,$$

and let us denote by A_p the integral operator of kernel α_p. Then

$$(A_p f)(y) = \int \sum_{j,k\leq p} \alpha_{jk}\overline{f}_j(x) g_k(y) f(x)\,\mathrm{d}x = \sum_{j,k\leq p} \alpha_{jk}\,\langle f, f_j\rangle\, g_k(y),$$

which shows that $\operatorname{rk} A_p \leq p$. On the other hand, Theorem 3.7 implies that $\|A - A_p\| \leq \|\alpha - \alpha_p\|$. Since the latter quantity tends to 0 as p tends to infinity, we see that A is the norm-limit of a sequence of finite rank operators. The result then follows by Theorem 3.3. ∎

Theorem 3.10 Let $\alpha \in L^2_{V\times W}(\mathbb{R}^{2n})$ 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 $\alpha^\star(y,x) = \overline{\alpha(x,y)}$.

Proof. By Fubini's theorem, we have:

$$\begin{aligned}
\langle Af, g\rangle &= \int_W \left(\int_V \alpha(x,y) f(x)\,\mathrm{d}x\right) \overline{g}(y)\,\mathrm{d}y \\
&= \int_{V\times W} \alpha(x,y) f(x) \overline{g}(y)\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_V f(x) \left(\int_W \alpha(x,y)\overline{g}(y)\,\mathrm{d}y\right)\,\mathrm{d}x \\
&= \int_V f(x) \overline{\int_W \overline{\alpha(x,y)} g(y)\,\mathrm{d}y}\,\mathrm{d}x.
\end{aligned}$$

By identification, we see that $(A^\star g)(x) = \int_W \overline{\alpha(x,y)} g(y)\,\mathrm{d}y$. ∎

Chapter 4

Convolution and approximation

4.1 Useful facts from integration theory

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

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

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

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

Theorem 4.3 [Fubini] Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, and let f be an $(\mathcal{A} \times \mathcal{B})$-measurable function on $X \times Y$.

(1) Suppose that f takes its values in $[0,\infty]$, and let

$$\phi(x) := \int_Y f(x,y) \, \mathrm{d}\nu(y) \quad \text{and} \quad \psi(y) := \int_Y f(x,y) \, \mathrm{d}\mu(x). \tag{4.1}$$

Then ϕ is $\mathcal{A}$-measurable, ψ is $\mathcal{B}$-measurable, and we have:

$$\int_X \phi(x) \, \mathrm{d}\mu(x) = \int_{X \times Y} f(x,y) \, \mathrm{d}(\mu \times \nu)(x,y) = \int_Y \psi(y) \, \mathrm{d}\nu(y). \tag{4.2}$$

(2) Suppose that f takes its values in $\mathbb{C}$, and that $h(x) := \int_Y |f(x,y)|\,\mathrm{d}\nu(y)$ is integrable. Then f is integrable.

(3) Suppose that f is integrable. Then $f(x,\cdot)$ is (measurable and) integrable for almost all x and $f(\cdot,y)$ is (measurable and) integrable for almost all y. Furthermore, the functions ϕ and ψ defined by Equation (4.1) are integrable and satisfy Equation (4.2).

Theorem 4.4 [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 $f\colon X \times Y \to \mathbb{R}^+$ be measurable. Then

$$\left(\int_X \left(\int_Y f\,\mathrm{d}\nu\right)^p \mathrm{d}\mu\right)^{1/p} \le \int_Y \left(\int_X f^p\,\mathrm{d}\mu\right)^{1/p} \mathrm{d}\nu. \tag{4.3}$$

Proof. Let $\varphi(x) := \int_Y f(x,y)\,\mathrm{d}\nu(y)$ and let $\psi(y) := \int_X f^p(x,y)\,\mathrm{d}\mu(x)$. By virtue of Fubini's Theorem, both φ and ψ are measurable. Therefore, the integrals in (4.3) are licit. If $f = 0$ $(\mu \times \nu)$-almost everywhere, the result is clear. Let us then assume that f is positive on a set of positive measure. This implies in particular that

$$\int_X \varphi^p(x)\,\mathrm{d}\mu(x) > 0,$$

as may be easily checked. We first assume that the left hand side of (4.3) is finite, that is, that $\int_X \varphi^p(x)\,\mathrm{d}\mu(x) < \infty$. We have:

$$\begin{aligned}\int_X \varphi^p(x)\,\mathrm{d}\mu(x) &= \int_X \left(\int_Y f(x,y)\,\mathrm{d}\nu(y)\right) \varphi^{p-1}(x)\,\mathrm{d}\mu(x)\\ &= \int_X \left(\int_Y f(x,y)\varphi^{p-1}(x)\,\mathrm{d}\nu(y)\right) \mathrm{d}\mu(x)\\ &= \int_Y \left(\int_X f(x,y)\varphi^{p-1}(x)\,\mathrm{d}\mu(x)\right) \mathrm{d}\nu(y)\end{aligned}$$

by Fubini's Theorem. Now, Hölder's Inequality shows that

$$\begin{aligned}&\int_X f(x,y)\varphi^{p-1}(x)\,\mathrm{d}\mu(x)\\ &\le \left(\int_X f^p(x,y)\,\mathrm{d}\mu(x)\right)^{1/p} \left(\int_X \varphi^{q(p-1)}(x)\,\mathrm{d}\mu(x)\right)^{1/q}.\end{aligned}$$

Since $q(p-1)=p$, we obtain:

$$\begin{aligned}&\int_X \varphi^p(x)\,\mathrm{d}\mu(x)\\ &\quad\le \int_Y \left(\int_X f^p(x,y)\,\mathrm{d}\mu(x)\right)^{1/p}\left(\int_X \varphi^p(x)\,\mathrm{d}\mu(x)\right)^{(p-1)/p}\mathrm{d}\nu(y).\end{aligned}$$

Dividing both sides of the above inequality by $\left(\int_X \varphi^p(x)\,\mathrm{d}\mu(x)\right)^{(p-1)/p}$ gives rise to (4.3). We now assume that $\int_X \varphi^p(x)\,\mathrm{d}\mu(x)=\infty$. We shall use a monotone convergence argument. Let (M_k) be a sequence of positive numbers converging to infinity, and let f_{M_k} be defined by

$$f_{M_k}(x,y):=\begin{cases} f(x,y) & \text{if } f(x,y)\le M_k,\\ 0 & \text{otherwise.}\end{cases}$$

Since $(X,\mathcal{A},\mu)$ and $(Y,\mathcal{B},\nu)$ are σ-finite, we can find increasing sequences $(A_k)\subset X$ and $(B_k)\subset Y$ such that $X\times Y=\cup_k(A_k\times B_k)$ and

$$(\mu\times\nu)(A_k\times B_k)=\mu(A_k)\cdot\nu(B_k)<\infty.$$

We then let $f_k:=1\!\!1_{A_k\times B_k}\cdot f_{M_k}$ for all k. The first part of the proof shows that, for all k,

$$\left(\int_X\left(\int_Y f_k\,\mathrm{d}\nu\right)^p\mathrm{d}\mu\right)^{1/p}\le\int_Y\left(\int_X f_k^p\,\mathrm{d}\mu\right)^{1/p}\mathrm{d}\nu. \tag{4.4}$$

It is clear that, for all $(x,y)\in X\times Y$ and all k, $f_k(x,y)\le f_{k+1}(x,y)$, and that $f_k(x,y)\to f(x,y)$ as $k\to\infty$. By virtue of the Beppo-Levi Theorem,

$$\varphi_k(x):=\int_Y f_k(x,y)\,\mathrm{d}\nu(y)\longrightarrow\int_Y f(x,y)\,\mathrm{d}\nu(y)=\varphi(x)$$

for all $x\in X$. Now, for all $x\in X$ and all k, $\varphi_k^p(x)\le\varphi_{k+1}^p(x)$ by monotony of the integral, and $\varphi_k^p(x)\to\varphi^p(x)$ as $k\to\infty$. By the Beppo-Levi Theorem again,

$$\int_X\varphi_k^p(x)\,\mathrm{d}\mu(x)\longrightarrow\int_X\varphi^p(x)\,\mathrm{d}\mu(x)=\infty.$$

We proceed likewise for the right hand side of (4.4). For all $(x,y)\in X\times Y$ and all k, $f_k^p(x,y)\le f_{k+1}^p(x,y)$, and $f_k^p(x,y)\to f^p(x,y)$ as $k\to\infty$, so that

$$\psi_k(y):=\int_X f_k^p(x,y)\,\mathrm{d}\mu(x)\longrightarrow\int_X f^p(x,y)\,\mathrm{d}\mu(x)=\psi(y).$$

For all $y \in Y$, $\psi_k^{1/p}(y) \leq \psi_{k+1}^{1/p}(y)$, and $\psi_k^{1/p}(y) \to \psi^{1/p}(y)$, so that

$$\int_Y \psi_k^{1/p}(y)\,\mathrm{d}\nu(y) \longrightarrow \int_Y \psi^{1/p}(y)\,\mathrm{d}\nu(y).$$

Passing to the limit in (4.4) then yields the result. ■

4.2 Convolution of L^p functions

Let $f, g\colon \mathbb{R}^n \to \mathbb{C}$ be measurable functions. Whenever the function $y \mapsto f(x-y)g(y)$ is integrable, we write:

$$(f * g)(x) := \int f(x-y)g(y)\,\mathrm{d}y. \tag{4.5}$$

We remark immediately that $y \mapsto f(x-y)g(y)$ is integrable if and only if $y \mapsto g(x-y)f(y)$ is integrable, and that

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

in this case. In other words, $(f * g)(x)$ exists if and only if $(g * f)(x)$ exists and both numbers are equal.

Under certain circumstances to be specified later on, $(f * g)(x)$ exists for almost every $x \in \mathbb{R}^n$. Equation (4.5) then defines a function almost everywhere. Note that the function $F(x, y) := f(x-y)g(y)$ satisfies $F = (f \circ u)(g \circ v)$, where $u(x, y) := x - y$ and $v(x, y) := y$. Since f, g, u and v are Borel functions, so are $(f \circ u), (g \circ v)$ and F.

Theorem 4.5 Let $f \in L^1(\mathbb{R}^n)$ and $g \in L^p(\mathbb{R}^n)$, where $p \in [1, \infty]$. Then, the function $\varphi_x\colon y \to f(x-y)g(y)$ 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 $\|f * g\|_p \leq \|f\|_1 \|g\|_p$.

PROOF. We first consider the case where $p = 1$. Since the function $F(x, y) := f(x-y)g(y)$ is measurable, Theorem 4.3(1) shows that $h(x) := \int |F(x, y)|\,\mathrm{d}y$ is measurable, and that

$$\int h(x)\,\mathrm{d}x = \iint |f(x-y)g(y)|\,\mathrm{d}x\,\mathrm{d}y = \int |g(y)| \int |f(x-y)|\,\mathrm{d}x\,\mathrm{d}y.$$

Since the Lebesgue measure is invariant under translations, we obtain:

$$\int h(x)\,\mathrm{d}x = \|f\|_1\,\|g\|_1\,.$$

This shows that h is integrable. Theorem 4.3(2) then implies that $F(x,y)$ is integrable, and Theorem 4.3(3) shows that $\varphi_x = F(x,\cdot)$ is integrable for almost all x. Finally,

$$\|f*g\|_1 \leq \iint |f(x-y)g(y)|\,\mathrm{d}y\,\mathrm{d}x = \int h(x)\,\mathrm{d}x,$$

and the desired result follows.

Suppose now that $p\in(1,\infty)$, and let $h_p(x) := \int |f(x-y)||g(y)|^p\,\mathrm{d}y$. Then h_p is measurable and

$$\begin{aligned}\int h_p(x)\,\mathrm{d}x &= \int\left(\int |f(x-y)|\,|g(y)|^p\,\mathrm{d}y\right)\mathrm{d}x\\ &= \int |g(y)|^p\left(\int |f(x-y)|\,\mathrm{d}y\right)\mathrm{d}x\\ &= \|f\|_1\,\|g\|_p^p\,, \end{aligned}\tag{4.6}$$

in which we have used again the invariance under translations of the Lebesgue measure. This shows that h_p is integrable (so that h_p is almost everywhere finite). Note that $y\mapsto |f(x-y)||g(y)|^p$ is measurable, which shows that $y\mapsto |f(x-y)||g(y)|$ is measurable, which implies in turn that φ_x is mesurable. Furthermore, Hölder's inequality gives rise to:

$$\begin{aligned}&\int |f(x-y)g(y)|\,\mathrm{d}y = \int |f(x-y)|^{1/q}\left(|f(x-y)|\,|g(y)|^p\right)^{1/p}\mathrm{d}y\\ &\leq \left(\int |f(x-y)|\,\mathrm{d}y\right)^{1/q}\cdot\left(\int |f(x-y)|\,|g(y)|^p\,\mathrm{d}y\right)^{1/p}\\ &= \|f\|_1^{1/q}\cdot h_p(x)^{1/p},\end{aligned}$$

so that φ_x is integrable for almost all x. Now,

$$\begin{aligned}
\|f * g\|_p &= \left(\int \left|\int f(x-y)g(y)\,\mathrm{d}y\right|^p \mathrm{d}x\right)^{1/p} \\
&\le \left(\int \left(\int |f(x-y)g(y)|\,\mathrm{d}y\right)^p \mathrm{d}x\right)^{1/p} \\
&\le \left(\int \|f\|_1^{p/q} \cdot h_p(x)\,\mathrm{d}x\right)^{1/p} \\
&= \|f\|_1^{1/q} \cdot \left(\int h_p(x)\,\mathrm{d}x\right)^{1/p} \\
&= \|f\|_1 \cdot \|g\|_p .
\end{aligned}$$

Finally, suppose that $p = \infty$. Clearly,

$$\int |f(x-y)g(y)|\,\mathrm{d}y \le \|g\|_\infty \cdot \int |f(x-y)|\,\mathrm{d}y = \|g\|_\infty \cdot \|f\|_1 < \infty.$$

Thus, for all $x \in \mathbb{R}^n$, φ_x is integrable and

$$\left|\int f(x-y)g(y)\,\mathrm{d}y\right| \le \|g\|_\infty \cdot \|f\|_1 ,$$

whence the inequality $\|f * g\|_\infty \le \|g\|_\infty \|f\|_1$. ∎

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 4.6 Let $f \in L^p(\mathbb{R}^n)$, where $p \in [1, \infty)$, and let $f_{[x]}$ be the *translate* of f by x: $f_{[x]}(y) := f(y-x)$. Then $\|f_{[x]} - f\|_p \to 0$ as $\|x\| \to 0$.

PROOF. We first prove the theorem for functions g in the space $\mathcal{C}_0(\mathbb{R}^n)$ of all continuous functions having compact support. Recall that such functions are uniformly continuous:

$$\forall \varepsilon > 0,\ \exists \eta > 0\colon\ \big\|y' - y\big\| \le \eta \Rightarrow \big|g(y') - g(y)\big| \le \varepsilon. \tag{4.7}$$

Fix $\varepsilon > 0$. Let x be such that $\|x\| \le \eta$, where η is provided by (4.7). Then, for all $y \in \mathbb{R}^n$, $|g_{[x]}(y) - g(y)| \le \varepsilon$. Let $K_0 := \mathrm{supp}(g)$. Then the set

$$K := \mathrm{supp}(g_{[x]} - g) \subset (K_0 + \{x\}) \cup K_0$$

is compact. Thus

$$\left\|g_{[x]}-g\right\|_p=\left(\int\left|g_{[x]}-g\right|^p\right)^{1/p}\leq(\varepsilon^p\operatorname{meas}K)^{1/p}\leq\varepsilon\,(2\operatorname{meas}K_0)^{1/p}.$$

Suppose now that f is any function in $L^p(\mathbb{R}^n)$, and let $\varepsilon>0$. Since $\mathcal{C}_0(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$, there exists $g\in\mathcal{C}_0(\mathbb{R}^n)$ such that $\|f-g\|_p<\varepsilon/3$. By the first part of the proof, $\|g_{[x]}-g\|_p<\varepsilon/3$ if $\|x\|$ is small enough, so that

$$\left\|f_{[x]}-f\right\|_p\leq\left\|f_{[x]}-g_{[x]}\right\|_p+\left\|g_{[x]}-g\right\|_p+\|g-f\|_p<\varepsilon. \blacksquare$$

Theorem 4.7 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

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

PROOF. Using the change of variable $x=\varepsilon x'$, we obtain

$$\int\varphi_\varepsilon(x)\,\mathrm{d}x=\int\varphi(x')\,\mathrm{d}x'=a.$$

We have:

$$\begin{aligned}(f*\varphi_\varepsilon)(x)-af(x)&=\int f(x-y)\varphi_\varepsilon(y)\,\mathrm{d}y-\int\varphi_\varepsilon(y)\,\mathrm{d}y\cdot f(x)\\&=\int(f(x-y)-f(x))\varphi_\varepsilon(y)\,\mathrm{d}y\\&=\int(f(x-\varepsilon y')-f(x))\varphi(y')\,\mathrm{d}y'\\&=\int(f_{\varepsilon y'}-f)(x)\varphi(y')\,\mathrm{d}y',\end{aligned}$$

where we have used the change of variable $y = \varepsilon y'$. Consequently,

$$\begin{aligned}
\|f * \varphi_\varepsilon - af\| &= \left(\int \left|\int (f_{\varepsilon y'} - f)(x)\varphi(y')\,\mathrm{d}y'\right|^p \mathrm{d}x\right)^{1/p} \\
&\leq \left(\int \left(\int \left|(f_{\varepsilon y'} - f)(x)\varphi(y')\right| \mathrm{d}y'\right)^p \mathrm{d}x\right)^{1/p} \\
&\leq \int \left(\int \left|(f_{\varepsilon y'} - f)(x)\varphi(y')\right|^p \mathrm{d}x\right)^{1/p} \mathrm{d}y' \\
&= \int |\varphi(y')| \cdot \left(\int \left|(f_{\varepsilon y'} - f)(x)\right|^p \mathrm{d}x\right)^{1/p} \mathrm{d}y' \\
&= \int |\varphi(y')| \cdot \left\|f_{\varepsilon y'} - f\right\|_p \mathrm{d}y',
\end{aligned}$$

in which the second inequality results from Minkowski's integral inequality. On the one hand, $|\varphi(y')| \cdot \|f_{\varepsilon y'} - f\|_p \leq 2\|f\|_p \cdot |\varphi(y')|$, and on the other hand $|\varphi(y')| \cdot \|f_{\varepsilon y'} - f\|_p$ goes to zero as $\varepsilon \to 0$ for almost all y'. Lebesgue's Dominated Convergence Theorem then implies that

$$\int |\varphi(y')| \cdot \left\|f_{\varepsilon y'} - f\right\|_p \mathrm{d}y' \longrightarrow 0 \quad \text{as } \varepsilon \to 0.$$

The theorem follows. ∎

Theorem 4.8 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 $\alpha \to 0$.

Proof. We have:

$$|(f * \psi_\alpha)(x) - f(x)| \leq \int |f(x-y) - f(x)|\,\psi_\alpha(y)\,\mathrm{d}y.$$

Let $\varepsilon > 0$ be fixed. Since f is continuous at x, there exists $\eta > 0$ such that $|x' - x| \leq \eta$ implies $|f(x') - f(x)| \leq \varepsilon/(3a)$. We then have:

$$\begin{aligned}
\int |f(x-y) - f(x)|\,\psi_\alpha(y)\,\mathrm{d}y &= \int_{\|y\|\leq\eta} |f(x-y) - f(x)|\,\psi_\alpha(y)\,\mathrm{d}y \\
+ \int_{\|y\|>\eta} |f(x-y) - f(x)|\,\psi_\alpha(y)\,\mathrm{d}y &\leq \frac{\varepsilon}{3} + I_1 + I_2,
\end{aligned}$$

in which $I_1 = \int_{\|y\|>\eta} |f(x-y)|\psi_\alpha(y)\,\mathrm{d}y$ and $I_2 = \int_{\|y\|>\eta} |f(x)|\psi_\alpha(y)\,\mathrm{d}y$. On the one hand,

$$I_1 \leq \sup_{\|y\|>\eta} \psi_\alpha(y) \cdot \int_{\|y\|>\eta} |f(x-y)|\,\mathrm{d}y \leq \sup_{\|y\|>\eta} \psi_\alpha(y) \cdot \|f\|_1\,,$$

and (b) implies that $I_1 \leq \varepsilon/3$ for α sufficiently small. On the other hand,

$$I_2 = |f(x)| \cdot \int_{\|y\|>\eta} \psi_\alpha(y)\,\mathrm{d}y = |f(x)| \cdot \int_{\|y'\|>\alpha^{-1}\eta} \psi(y')\,\mathrm{d}y',$$

in which the last integral tends to 0 as $\alpha \to 0$ by (a), so that $I_2 < \varepsilon/3$ for α sufficiently small. ∎

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 4.9 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,

$$D^\alpha(f * \varphi) = f * D^\alpha \varphi. \tag{4.8}$$

Proof. By Hölder's inequality, we have:

$$\begin{aligned}&\int |f(x-y)|\,|D^\alpha\varphi(y)|\,\mathrm{d}y\\ \leq\;& \left(\int |f(x-y)|^p\,\mathrm{d}y\right)^{1/p}\left(\int |D^\alpha\varphi(y)|^q\,\mathrm{d}y\right)^{1/q}\\ =\;& \|f\|_p\cdot\|D^\alpha\varphi\|_q\,,\end{aligned}$$

Since $D^\alpha\varphi \in \mathcal{S}(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$, the above inequality shows that the function $y \mapsto f(x-y)D^\alpha\varphi(y)$ is integrable for all x (and all $\alpha \in \mathbb{N}^n$). This shows that the right hand side of Equation (4.8) is well-defined for all α. The assumption on f and φ allow to interchange differentiation and integration as follows:

$$\begin{aligned}f * D^\alpha\varphi(x) &= \int D^\alpha\varphi(x-y)f(y)\,\mathrm{d}y\\ &= D^\alpha\int \varphi(x-y)f(y)\,\mathrm{d}y\\ &= D^\alpha(f*\varphi)(x). \ \blacksquare \qquad (4.9)\end{aligned}$$

Theorem 4.10 Let $f, g \in \mathcal{C}_0(\mathbb{R}^n)$. Then $\mathrm{supp}((f*g)) \subset \mathrm{supp}(f)+\mathrm{supp}(g)$. In particular, $f*g$ has compact support.

Proof. Let $x \in \mathrm{supp}((f*g))$. By definition of $\mathrm{supp}((f*g))$, we have:

$$\forall\varepsilon > 0, \exists x' \in B(x,\varepsilon)\colon (f*g)(x') \neq 0.$$

Clearly, we must have $g(y) \neq 0$ and $f(x'-y) \neq 0$ for some y, which shows that $x' = x' - y + y \in \mathrm{supp}(f) + \mathrm{supp}(g)$. This shows that x belongs to $\mathrm{cl}(\mathrm{supp}(f) + \mathrm{supp}(g))$, where the closure operation can be omitted since $\mathrm{supp}(f)$ and $\mathrm{supp}(g)$ are compact subsets of $\mathbb{R}^n$. $\blacksquare$

Theorem 4.11 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 $p \in [1,\infty)$.

Proof. Let $f \in L^p(\mathbb{R}^n)$, and fix $\delta > 0$. Since $\mathcal{C}_0(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$, we can find $g \in \mathcal{C}_0(\mathbb{R}^n)$ such that $\|f-g\|_p < \delta/2$. We shall prove that there

exists $h \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ such that $\|g-h\|_p < \delta/2$, which will complete the proof by the triangle inequality.
Let $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ be nonnegative and such that $\int \varphi(x)\,\mathrm{d}x = 1$. For example,

$$\varphi(x) := \begin{cases} c \cdot \exp\left(-\dfrac{1}{1-\|x\|^2}\right) & \text{if } x \in B(0,1), \\ 0 & \text{otherwise,} \end{cases}$$

where c is chosen so that the normality condition is satisfied. For $\varepsilon > 0$, let φ_ε be defined as in Theorem 4.7. By Theorems 4.9 and 4.10, $g * \varphi_\varepsilon$ belong to $\mathcal{C}_0^\infty(\mathbb{R}^n)$. By Theorem 4.7, $\|g * \varphi_\varepsilon - g\|_p \to 0$ as $\varepsilon \to 0$. ■

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

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

Proof. Suppose $P := \{p \in [1,\infty) \mid f \in L^p(\mathbb{R}^n)\} \neq \emptyset$ (for otherwise, the theorem holds vacuously). Let φ and φ_ε be as in the proof of Theorem 4.11. For all $k \in \mathbb{N}^*$ and all $p \in P$, $f\mathbb{1}_{B_k}$ belongs to L^p and has compact support. Since $\varphi_{1/k} \in \mathcal{C}_0^\infty(\mathbb{R}^n)$, the function $f_k := (f\mathbb{1}_{B_k}) * \varphi_{1/k}$ belongs to $\mathcal{C}_0^\infty(\mathbb{R}^n)$ by Theorems 4.9 and 4.10. In addition, $f * \varphi_{1/k} \in L^p(\mathbb{R}^k)$ by Theorem 4.5. For all $p \in P$, we have:

$$\begin{aligned} \|f_k - f\|_p &\leq \left\|(f\mathbb{1}_{B_k}) * \varphi_{1/k} - f * \varphi_{1/k}\right\|_p + \left\|f * \varphi_{1/k} - f\right\|_p \\ &= \left\|(f\mathbb{1}_{B_k} - f) * \varphi_{1/k}\right\|_p + \left\|f * \varphi_{1/k} - f\right\|_p. \end{aligned}$$

Theorem 4.5 implies that $\|(f\mathbb{1}_{B_k} - f) * \varphi_{1/k}\|_p \leq \|f\mathbb{1}_{B_k} - f\|_p \|\varphi_{1/k}\|_1$, in which $\|\varphi_{1/k}\|_1 = 1$. We therefore have:

$$\|f_k - f\|_p \leq \|f\mathbb{1}_{B_k} - f\|_p + \left\|f * \varphi_{1/k} - f\right\|_p.$$

Now, $\|f\mathbb{1}_{B_k} - f\|_p \to 0$ as $k \to \infty$ by Lebesgue's Dominated Convergence Theorem, $\|f * \varphi_{1/k} - f\|_p \to 0$ as $k \to \infty$ by Theorem 4.7, and the result follows. ■

We now give a generalization of Theorem 4.5.

Theorem 4.13 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 $\|h\|_r \leq \|f\|_p \|g\|_q$.

Proof. In the case where $q = \infty$, the conditions $p^{-1} \in [0, 1]$ and $r^{-1} \in [0, 1]$ imply that $r = \infty$ and $p = 1$, and the result is then contained in Theorem 4.5. By symmetry, the case where $p = \infty$ is also clear. From now on, we assume that p and q belong to $[1, \infty)$.
Observe that, if $r = \infty$, p and q are conjugate exponents. By Hölder's inequality, we then have:

$$\begin{aligned}\int |f(x-y)g(y)|\,\mathrm{d}y &\leq \left(\int |f(x-y)|^p\,\mathrm{d}y\right)^{1/p} \left(\int |g(y)|^q\,\mathrm{d}y\right)^{1/q} \\ &= \|f\|_p \|g\|_q \,. \end{aligned} \tag{4.10}$$

Thus, for all x, φ_x is integrable and

$$|h(x)| \leq \int |f(x-y)g(y)|\,\mathrm{d}y \leq \|f\|_p \|g\|_q \,.$$

Suppose now that $r \in [1, \infty)$. Let $s := p(1-q^{-1})$ and let q' be the conjugate exponent of q. We then have:

$$sq' = p\left(1 - \frac{1}{q}\right)\left(1 - \frac{1}{q}\right)^{-1} = p.$$

By Hölder's inequality, we have

$$\begin{aligned}\int |f(x-y)g(y)|\,\mathrm{d}y &= \int |f(x-y)|^{1-s} |g(y)|\, |f(x-y)|^s\,\mathrm{d}y \\ &\leq \left(\int |f(x-y)|^{(1-s)q} |g(y)|^q\,\mathrm{d}y\right)^{1/q} \left(\int |f(x-y)|^{sq'}\,\mathrm{d}y\right)^{1/q'} \qquad (4.11) \\ &= \left(\int |f(x-y)|^{(1-s)q} |g(y)|^q\,\mathrm{d}y\right)^{1/q} \|f\|_p^s \,.\end{aligned}$$

In order to prove that φ_x is integrable for almost all x, we need only prove that $\int |f(x-y)|^{(1-s)q}|g(y)|^q \,\mathrm{d}y$ is finite for almost all x. This will be a consequence of the inequality

$$\left(\int \left(\int |f(x-y)|^{(1-s)q}\,|g(y)|^q \,\mathrm{d}y\right)^{r/q} \mathrm{d}x\right)^{q/r} \leq \|f\|_p^{pq/r}\cdot\|g\|_q^q\,, \quad (4.12)$$

which we now establish.

Since $r^{-1} \in (0,1]$, we see that $p^{-1}+q^{-1} > 1$, so that $1-s>0$. By Minkowski's integral inequality, we have, for all $\alpha \geq 1$,

$$\begin{aligned} &\left(\int \left(\int |f(x-y)|^{(1-s)q}\,|g(y)|^q \,\mathrm{d}y\right)^{\alpha} \mathrm{d}x\right)^{1/\alpha} \\ \leq\;& \int \left(\int |f(x-y)|^{(1-s)q\alpha}\,|g(y)|^{q\alpha} \,\mathrm{d}x\right)^{1/\alpha} \mathrm{d}y \\ =\;& \int |g(y)|^q \cdot \left(\int |f(x-y)|^{(1-s)q\alpha} \,\mathrm{d}x\right)^{1/\alpha} \mathrm{d}y \\ =\;& \|f\|_{(1-s)q\alpha}^{(1-s)q}\cdot\|g\|_q^q\,. \end{aligned}$$

Since $rq^{-1} = 1+r-rp^{-1}$ and since $p \geq 1$, we see that $rq^{-1} \geq 1$, and we can chose $\alpha = rq^{-1}$. Inequality (4.12) then follows from the fact that $(1-s)q\alpha = p$.

Finally, let us establish the inequality $\|h\|_r \leq \|f\|_p\|g\|_q$. We have:

$$\begin{aligned} \int |h(x)|^r \,\mathrm{d}x &= \int \left|\int f(x-y)g(y)\,\mathrm{d}y\right|^r \mathrm{d}x \\ &\leq \int \left(\int |f(x-y)g(y)| \,\mathrm{d}y\right)^r \mathrm{d}x \\ &\leq \|f\|_p^{rs}\cdot\int \left(\int |f(x-y)|^{(1-s)q}\,|g(y)|^q \,\mathrm{d}y\right)^{r/q} \mathrm{d}x \\ &\leq \|f\|_p^{rs+p}\cdot\|g\|_q^r\,, \qquad (4.13) \end{aligned}$$

in which the second and third inequalities result from Inequalities (4.11) and (4.12), respectively. Thus, $\|h\|_r \leq \|f\|_p^{(rs+p)/r}\cdot\|g\|_q$, where

$$\frac{rs+p}{r} = p\left(1-\frac{1}{q}\right)+\frac{p}{r} = p\left(1+\frac{1}{r}-\frac{1}{q}\right) = 1. \;\blacksquare$$

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 throughout.

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 4.5), and

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

Proof. We have:

$$\begin{aligned}
U(f*g)(\xi) &= \int e^{-2i\pi\langle x,\xi\rangle}\left(\int f(x-y)g(y)\,\mathrm{d}y\right)\mathrm{d}x \\
&= \iint e^{-2i\pi\langle x,\xi\rangle} f(x-y)g(y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \iint e^{-2i\pi\langle x-y,\xi\rangle} f(x-y)\cdot e^{-2i\pi\langle y,\xi\rangle} g(y)\,\mathrm{d}y\,\mathrm{d}x \\
&= \int e^{-2i\pi\langle y,\xi\rangle} g(y)\left(\int e^{-2i\pi\langle x-y,\xi\rangle} f(x-y)\,\mathrm{d}x\right)\mathrm{d}y \\
&= \hat{f}(\xi)\cdot\hat{g}(\xi). \ \blacksquare \qquad (5.1)
\end{aligned}$$

Lemma 5.1 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 $x \mapsto (-2i\pi x)$;

Proof. For simplicity, assume $n = 1$. Integrating by parts yields

$$\begin{aligned}
&\int e^{-2i\pi x\xi} D^\alpha\varphi(x)\,\mathrm{d}x \\
&= \left[e^{-2i\pi x\xi} D^{\alpha-1}\varphi(x)\right]_{-\infty}^{\infty} - \int (-2i\pi\xi)e^{-2i\pi x\xi} D^{\alpha-1}\varphi(x)\,\mathrm{d}x \\
&= (2i\pi\xi)UD^{\alpha-1}\varphi(\xi),
\end{aligned}$$

and (1) follows by induction. On the other hand we have:

$$\begin{aligned}
U\left((-2i\pi x)^\beta\varphi\right)(\xi) &= \int e^{-2i\pi x\xi}(-2i\pi x)^\beta\varphi(x)\,\mathrm{d}x \\
&= \int D(e^{-2i\pi x\xi})(-2i\pi x)^{\beta-1}\varphi(x)\,\mathrm{d}x \\
&= D\int e^{-2i\pi x\xi}(-2i\pi x)^{\beta-1}\varphi(x)\,\mathrm{d}x \\
&= DU\left((-2i\pi x)^{\beta-1}\varphi\right)(\xi),
\end{aligned}$$

and (2) follows again by induction. ∎

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

Proof. Let α and β be two multi-indices. Then,

$$\begin{aligned}\left|\xi^\alpha D^\beta \hat{\phi}(\xi)\right| &= \left|\xi^\alpha U\left((-2i\pi x)^\beta \varphi\right)(\xi)\right| \\ &= \left|(2i\pi)^{-|\alpha|}(2i\pi\xi)^\alpha U\left((-2i\pi x)^\beta \varphi\right)(\xi)\right| \\ &= (2\pi)^{-|\alpha|}\left|UD^\alpha\left((-2i\pi x)^\beta \varphi\right)(\xi)\right|.\end{aligned}$$

Since the function $D^\alpha\left((-2i\pi x)^\beta \varphi\right)$ belongs to $\mathcal{S}(\mathbb{R}^n)$ (and is therefore integrable), we can write

$$\begin{aligned}\sup_\xi \left|\xi^\alpha D^\beta \hat{\varphi}(\xi)\right| &= (2\pi)^{-|\alpha|}\sup_\xi \left|\int e^{-2i\pi\langle x,\xi\rangle} D^\alpha\left((-2i\pi x)^\beta \varphi\right)(x)\,\mathrm{d}x\right| \\ &\leq (2\pi)^{-|\alpha|}\int \left|D^\alpha\left((-2i\pi x)^\beta \varphi\right)(x)\right|\,\mathrm{d}x \\ &= (2\pi)^{-|\alpha|}\left\|D^\alpha\left((-2i\pi x)^\beta \varphi\right)\right\|_1 . \blacksquare\end{aligned}$$

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.

Proof. By definition, we have, for all $\xi \in \mathbb{R}^n$,

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

This proves (1). Note that, if f belongs to $\mathcal{S}(\mathbb{R}^n)$, then so does $\hat{f}$ by the previous theorem. Thus (2) is certainly satisfied in this case. If f is any member of $L^1(\mathbb{R}^n)$, then we can find a sequence (f_k) in $\mathcal{S}(\mathbb{R}^n)$ such that

$\|f_k - f\|_1 \to 0$ as $k \to \infty$. By (1), $\|\hat{f}_k - \hat{f}\|_\infty \to 0$. Therefore $\hat{f}$ is continuous, as the uniform limit of a sequence of continuous functions. Finally $\hat{f}(\xi) \to 0$ as $\|\xi\| \to \infty$ by Exercise 5.1 above. ▪

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) $UT_y f(\xi) = M_y Uf(\xi)$;

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

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

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

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

PROOF. By Theorem 5.3, $\hat{f}$ and $\hat{g}$ are bounded, which shows that both $\hat{f}g$ and $f\hat{g}$ belong to $L^1(\mathbb{R}^n)$. On the other hand, by Fubini's Theorem,

$$\begin{aligned}\int \hat{f}(x)g(x)\,\mathrm{d}x &= \int\left(\int e^{-2i\pi\langle x,y\rangle}f(y)\,\mathrm{d}y\right)g(x)\,\mathrm{d}x\\ &= \int f(y)\int e^{-2i\pi\langle x,y\rangle}g(x)\,\mathrm{d}x\,\mathrm{d}y\\ &= \int f(x)\hat{g}(x)\,\mathrm{d}x. \blacksquare\end{aligned}$$

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 $U^{-1} = \overline{U}$.

PROOF. By Theorem 5.2(3), $\hat{f} \in \mathcal{S}(\mathbb{R}^n)$, so that the integral is well-defined. Let $\psi(x) := e^{-\pi\|x\|^2}$, and let $g(\xi) := e^{2i\pi\langle x,\xi\rangle}\psi(\alpha\xi)$. By Lebesgue's Dominated Convergence Theorem,

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

as $\alpha \to 0$. But Theorem 5.4 shows that $\int g(\xi)\hat{f}(\xi)\,\mathrm{d}\xi = \int \hat{g}(y)f(y)\,\mathrm{d}y$, where

$$\begin{aligned}\hat{g}(y) &= UM_{-x}S_\alpha\psi(y)\\ &= T_xUS_\alpha\psi(y)\\ &= T_x(\alpha^{-n}S_{1/\alpha}U\psi)(y)\\ &= \alpha^{-n}\psi((y-x)/\alpha) = \psi_\alpha(x-y),\end{aligned}$$

in which $\psi_\alpha(x) := \alpha^{-n}\psi(x/\alpha)$ as usual. Thus $\int \hat{g}(y)f(y)\,\mathrm{d}y = (f*\psi_\alpha)(x) \to f(x)$ as $\alpha \to 0$, by Theorem 4.8. $\blacksquare$

Remark 5.1 The mapping

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

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

$$\begin{aligned}\langle Uf, g\rangle &= \int\left(\int e^{-2i\pi\langle x,\xi\rangle} f(x)\right)\overline{g}(\xi)\,\mathrm{d}\xi \\ &= \int f(x)\left(\int e^{-2i\pi\langle x,\xi\rangle}\overline{g}(\xi)\,\mathrm{d}\xi\right)\mathrm{d}x \\ &= \int f(x)\overline{\int e^{2i\pi\langle x,\xi\rangle} g(\xi)\,\mathrm{d}\xi}\,\mathrm{d}x \\ &= \left\langle f, U^{-1}g\right\rangle, \end{aligned} \tag{5.2}$$

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.

Proof. Since $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$, we can find a sequence $(f_k) \subset \mathcal{S}(\mathbb{R}^n)$ such that $\|f_k - f\|_1 \to 0$ as $k \to \infty$. By Theorem 5.5, $f_k = \check{\hat{f}}_k$ for all k. Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$. For all k, we have

$$\int f_k(x)\varphi(x)\,\mathrm{d}x = \int \check{\hat{f}}_k(x)\varphi(x)\,\mathrm{d}x = \int \hat{f}_k(y)\check{\varphi}(y)\,\mathrm{d}y$$

by Theorem 5.4. On the one hand,

$$\begin{aligned}\left|\int f_k(x)\varphi(x)\,\mathrm{d}x - \int f(x)\varphi(x)\,\mathrm{d}x\right| &\leq \int |f_k(x) - f(x)|\,|\varphi(x)|\,\mathrm{d}x \\ &\leq \sup_x |\varphi(x)| \cdot \|f_k - f\|_1\,, \end{aligned}$$

which shows that $\int f_k(x)\varphi(x)\,\mathrm{d}x$ tends to $\int f(x)\varphi(x)\,\mathrm{d}x$ as k tends to infinity. On the other hand,

$$\begin{aligned}\left|\int \hat{f}_k(y)\check{\varphi}(y)\,\mathrm{d}y - \int \hat{f}(y)\check{\varphi}(y)\,\mathrm{d}y\right| &\leq \int \left|\hat{f}_k(y) - \hat{f}(y)\right|\check{\varphi}(y)\,\mathrm{d}y \\ &\leq \sup_y \left|\hat{f}_k(y) - \hat{f}(y)\right| \int \check{\varphi}(y)\,\mathrm{d}y, \end{aligned}$$

so that $\int \hat{f}_k(y)\check{\varphi}(y)\,\mathrm{d}y$ tends to $\int \hat{f}(y)\check{\varphi}(y)\,\mathrm{d}y$ as k tends to infinity by Theorem 5.3. But the latter integral is equal to $\int \check{\hat{f}}(y)\varphi(y)\,\mathrm{d}y$ by Theorem 5.4, since both $\hat{f}$ and φ belong to $L^1(\mathbb{R}^n)$. Thus we have shown that

$$\int f(x)\varphi(x)\,\mathrm{d}x = \int \check{\hat{f}}(x)\varphi(x)\,\mathrm{d}x$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$, which implies that f and $\check{\hat{f}}$ coincide almost everywhere. ∎

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 $\alpha \to 0$.

Proof. Clearly, $f * \psi_\alpha \in L^1(\mathbb{R}^n)$ and $U(f * \psi_\alpha) = Uf \cdot U\psi_\alpha$ (see Theorem 5.1). On the other hand, $\hat{\psi} \in \Gamma(\mathbb{R}^n)$ by Theorem 5.3. Consequently, the function $\hat{f} \cdot \hat{\psi}(\alpha\cdot)$ satisfies:

(a) $|\hat{f} \cdot \hat{\psi}(\alpha\cdot)| \le K|\hat{f}|$, where $K := \sup|\hat{\psi}| < \infty$;

(b) for all $\xi \in \mathbb{R}^n$, $\hat{f}(\xi) \cdot \hat{\psi}(\alpha\xi) \to \hat{f}(\xi)$ as $\alpha \to 0$.

Lebesgue's Dominated Convergence Theorem then implies that $\|\hat{f} \cdot \hat{\psi}(\alpha\cdot) - \hat{f}\|_1 \to 0$ as $\alpha \to 0$. By Theorem 5.3, this implies in turn that

$$\left\|\overline{U}(\hat{f} \cdot \hat{\psi}(\alpha\cdot)) - \overline{U}\hat{f}\right\|_\infty \to 0 \quad \text{as} \quad \alpha \to 0.$$

Note that $\overline{U}\hat{f} = f$ almost everywhere by Theorem 5.6. Now, $f * \psi_\alpha$ belongs to $L^1(\mathbb{R}^n)$ and so does $U(f * \psi_\alpha) = \hat{f} \cdot \hat{\psi}(\alpha\cdot)$ (see Theorem 5.1). This implies that $f * \psi_\alpha \overset{\text{ae}}{=} \overline{U}U(f * \psi_\alpha) = \overline{U}(\hat{f} \cdot \hat{\psi}(\alpha\cdot))$ by Theorem 5.6 again, and the result follows. ∎

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 5.2 [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 $\alpha! := \alpha_1! \cdots \alpha_n!$.

Proof. We have: $(a_1 + \cdots + a_n)^N = \sum_{|\alpha|=N} \nu_\alpha a^\alpha$, in which

$$\begin{aligned}\nu_\alpha &= C_N^{\alpha_1} C_{N-\alpha_1}^{\alpha_2} \cdots C_{N-(\alpha_1+\cdots+\alpha_{n-1})}^{\alpha_n} \\ &= \frac{N!}{\alpha_1!(N-\alpha_1)!} \frac{(N-\alpha_1)!}{\alpha_2!(N-(\alpha_1+\alpha_2))!} \cdots \frac{(N-(\alpha_1+\cdots+\alpha_{n-1}))!}{\alpha_n!\,0!} \\ &= \frac{N!}{\alpha_1!\cdots\alpha_n!}. \ \blacksquare\end{aligned}$$

Lemma 5.3 For all $x \in \mathbb{R}^n$ and all $\alpha \in \mathbb{N}^n$, $|x^\alpha| \le \|x\|^{|\alpha|}$.

Proof. For all a, b, u, v in $\mathbb{R}_+^*$, the inequality

$$\frac{u}{u+v}a + \frac{v}{u+v}b \le a + b$$

is obvious. Since the function ln is concave and increasing, we deduce that

$$\frac{u}{u+v}\ln a + \frac{v}{u+v}\ln b \le \ln\left(\frac{u}{u+v}a + \frac{v}{u+v}b\right) \le \ln(a+b),$$

so that $(u+v)\ln(a+b) \ge u\ln a + v\ln b$. We conclude, using a continuity argument, that

$$\forall a, b, u, v \in \mathbb{R}_+, \quad (a+b)^{u+v} \ge a^u b^v. \tag{5.3}$$

Now, if we substitute $x_1^2, x_2^2, \alpha_1/2$ and $\alpha_2/2$ to a, b, u and v, respectively, in (5.3), we obtain:

$$\forall x_1, x_2 \in \mathbb{R}, \quad \forall \alpha_1, \alpha_2 \in \mathbb{R}_+, \quad \left(x_1^2 + x_2^2\right)^{\frac{\alpha_1+\alpha_2}{2}} \ge |x_1|^{\alpha_1} |x_2|^{\alpha_2}.$$

The result follows by induction. $\blacksquare$

Lemma 5.4 Let $\gamma\colon \mathbb{R}^n \to [0, \infty]$ be any function. The following statements are equivalent:

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

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

Proof. The fact that (ii) implies (i) is immediate from Lemma 5.3. Suppose that (i) holds. Then $\sup_{x\in\mathbb{R}^n}\{\|x\|^N\gamma(x)\} = \max\{S_1, S_2\}$, in which

$$S_1 := \sup_{\|x\|\leq 1}\left\{\|x\|^N\gamma(x)\right\} \quad \text{and} \quad S_2 := \sup_{\|x\|>1}\left\{\|x\|^N\gamma(x)\right\}.$$

On the one hand, $S_1 \leq \sup_{\|x\|\leq 1}\gamma(x)$, latter number being finite under Condition (i). On the other hand,

$$S_2 \leq \sup_{\|x\|>1}\left\{\|x\|^{2N}\gamma(x)\right\} \leq \sum_{|\alpha|=N}\frac{N!}{\alpha!}\sup_{\|x\|>1}\left\{x^{2\alpha}\gamma(x)\right\}$$

by Lemma 5.2 and the fact the supremum of a sum is less than or equal to the sum of the suprema. Thus S_2 is also finite under Condition (i), and Condition (ii) is established. ∎

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

Proof. Let α and β be two multi-indices. By Lemma 5.3, we have:

$$\begin{aligned}\left|x^\alpha D^\beta(f*g)(x)\right| &= \left|x^\alpha D^\beta\left(\int f(x-y)g(y)\,\mathrm{d}y\right)\right| \\ &\leq \int |x^\alpha|\left|D_x^\beta f(x-y)g(y)\right|\,\mathrm{d}y \\ &\leq \int \|x\|^{|\alpha|}\left|D_x^\beta f(x-y)g(y)\right|\,\mathrm{d}y\end{aligned}$$

Now, given any $\beta \in \mathbb{N}$, the function $x \mapsto x^\beta$ is convex on $\mathbb{R}^+$, so that

$$\forall a, b \in \mathbb{R}^+, \quad (a+b)^\beta \leq 2^{\beta-1}\left(a^\beta + b^\beta\right),$$

from which we deduce that $\|x\|^{|\alpha|} \leq 2^{|\alpha|-1}\big(\|x-y\|^{|\alpha|}+\|y\|^{|\alpha|}\big)$. Consequently, $\left|x^\alpha D^\beta(f*g)(x)\right| \leq 2^{|\alpha|-1}(I_1+I_2)$ where

$$\begin{aligned} I_1 &:= \int \|x-y\|^{|\alpha|}\left|D_x^\beta f(x-y)\right|\,|g(y)|\,\mathrm{d}y \\ \text{and} \quad I_2 &:= \int \|y\|^{|\alpha|}\left|D_x^\beta f(x-y)\right|\,|g(y)|\,\mathrm{d}y.\end{aligned}$$

Clearly, $I_1 \leq \sup_z\big\{\|z\|^{|\alpha|}|D^\beta f(z)|\big\}\cdot\int|g(y)|\,\mathrm{d}y$. The latter supremum is finite by Lemma 5.4 and the fact that f belongs to $\mathcal{S}(\mathbb{R}^n)$, and so is the integral of $|g|$. On the other hand, $I_2 \leq \sup_z \left|D^\beta f(z)\right| \cdot \int \|y\|^{|\alpha|}|g(y)|\,\mathrm{d}y$. Again the supremum is finite, and so is the integral, since

$$\begin{aligned} \int \|y\|^{|\alpha|}\,|g(y)|\;\mathrm{d}y &= \int_{\|y\|\leq 1} \|y\|^{|\alpha|}\,|g(y)|\;\mathrm{d}y + \int_{\|y\|>1} \|y\|^{|\alpha|}\,|g(y)|\;\mathrm{d}y \\ &\leq \int_{\|y\|\leq 1} |g(y)|\;\mathrm{d}y + \int_{\|y\|>1} \left|\|y\|^{2|\alpha|}\,g(y)\right|\;\mathrm{d}y, \end{aligned}$$

in which $y \mapsto \|y\|^{2|\alpha|}g(y)$ belongs to $\mathcal{S}(\mathbb{R}^n)$ and is therefore integrable. Thus both I_1 and I_2 are bounded above by finite numbers which are independent of x. Since α and β were arbitrary, we see that $f * g$ belongs to $\mathcal{S}(\mathbb{R}^n)$. ∎

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:

$$\exists k, K > 0\colon \forall x \in E,\ k\,\|x\| \le \|Ax\| \le K\,\|x\|\,. \tag{5.4}$$

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

(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 $(\operatorname{cl} A)^{-1} = \operatorname{cl}(A^{-1})$.

Proof. Let $x \in F$ and let $(x_n) \subset E$ be a sequence converging to x. The image sequence (Ax_n) is a Cauchy sequence, since $\|Ax_n - Ax_m\| \le K\|x_n - x_m\|$. Since G is complete, (Ax_n) converges to some $y \in \operatorname{cl}\operatorname{ran} A$. Now, if $(x'_n) \subset E$ is another sequence converging to x, then

$$\left\|y - Ax'_n\right\| \le \|y - Ax_n\| + \left\|Ax_n - Ax'_n\right\| \le \|y - Ax_n\| + K\left\|x_n - x'_n\right\|$$

by the triangle inequality, which implies that $Ax'_n \to y$. We then define

$$\mathbb{A}x := \begin{cases} Ax & \text{if } x \in E; \\ \lim Ax_n & \text{otherwise,} \end{cases}$$

in which (x_n) is any sequence converging to x. Let us show that $\mathbb{A}$ is linear. Given $\alpha, \beta \in \mathbb{C}$, $x, y \in F$ and $(x_n), (y_n) \subset E$ such that $x_n \to x$ and $y_n \to y$, we have:

$$\begin{aligned} \mathbb{A}(\alpha x + \beta y) &= \lim A(\alpha x_n + \beta y_n) \\ &= \alpha \lim Ax_n + \beta \lim Ay_n \\ &= \alpha \mathbb{A}x + \beta \mathbb{A}y. \end{aligned} \tag{5.5}$$

The continuity of $\mathbb{A}$ will follow from Statement (1), which we now establish. Let $x \in F$ and let $(x_n) \subset E$ be any sequence converging to x. For all $n \in \mathbb{N}$, we have $k\|x_n\| \le \|\mathbb{A}x_n\| \le K\|x_n\|$. By continuity of the norm, we can pass to the limit in the last inequalities, and (1) follows. Suppose now that $\mathbb{A}'$ is another continuous linear mapping whose restriction to E

coincides with A. Let $x, (x_n)$ be as above. By continuity of $\mathbb{A}$ and $\mathbb{A}'$, $\mathbb{A}x_n \to \mathbb{A}x$ and $\mathbb{A}'x_n \to \mathbb{A}'x$. Since $\mathbb{A}'x_n = Ax_n = \mathbb{A}x_n$ and since the limit of a converging sequence is unique, we must have: $\mathbb{A}'x = \mathbb{A}x$. Next we prove that $\operatorname{ran}\mathbb{A} = \operatorname{cl}\operatorname{ran} A$. Suppose $y \in \operatorname{ran}\mathbb{A}$. There exists $x \in F$ such that $\mathbb{A}x = y$, and we can find a sequence $(x_n) \subset E$ such that $x_n \to x$, so that $\mathbb{A}x_n = Ax_n \to y$. This shows that $\operatorname{ran}\mathbb{A} \subset \operatorname{cl}\operatorname{ran} A$. In order to obtain the opposite inclusion, let $y \in \operatorname{cl}\operatorname{ran} A$. There exists $(y_n) \subset \operatorname{ran} A$ such that $y_n \to y$. Since A is injective (see Condition (5.4)), $A^{-1}\colon \operatorname{ran} A \to E$ exists. Let $x_n := A^{-1}y_n$. Then $k\|x_n - x_m\| \leq \|y_n - y_m\|$, so that (x_n) is a Cauchy sequence. Thus (x_n) converges to some $x \in F$. We have:

$$\mathbb{A}x = \lim \mathbb{A}x_n = \lim Ax_n = \lim y_n = y.$$

Thus $y \in \operatorname{ran}\mathbb{A}$, and (2) is established. The fact that $\mathbb{A}\colon F \to \operatorname{ran}\mathbb{A}$ is bijective and bicontinuous results immediately from (1). Finally, if $y \in \operatorname{ran}\mathbb{A}$, we can find a sequence $(y_n) \subset \operatorname{ran} A$ which converges to y, so that

$$\mathbb{A}^{-1}y = \mathbb{A}^{-1}\lim y_n = \lim \mathbb{A}^{-1}y_n = \lim A^{-1}y_n,$$

which proves the last part of (3). ∎

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).

Proof. In Proposition 5.1, take $E = \mathcal{S}(\mathbb{R}^n)$, $F = G = L^2(\mathbb{R}^n)$, and $A = U$. We shall prove that $k = K = 1$ in the present case, that is, that $\|\hat{\varphi}\|^2 = \|\varphi\|^2$ for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$ and let $\psi \in \mathcal{S}(\mathbb{R}^n)$ be defined by $\psi(x) := \overline{\varphi(-x)}$. Clearly, $\psi \in \mathcal{S}(\mathbb{R}^n)$. We have

$$\|\varphi\|^2 = \int \varphi(x)\overline{\varphi(x)}\,\mathrm{d}x = \int \varphi(x)\psi(-x)\,\mathrm{d}x = (\varphi * \psi)(0).$$

Note that $\varphi * \psi \in \mathcal{S}(\mathbb{R}^n)$ by Theorem 5.8, so that

$$(\varphi * \psi)(0) = \overline{U}U(\varphi * \psi)(0) = \overline{U}(\hat{\varphi}\hat{\psi})(0) = \int \hat{\varphi}(\xi)\hat{\psi}(\xi)\,\mathrm{d}\xi.$$

Now, for all $\xi \in \mathbb{R}^n$, we have

$$\hat{\psi}(\xi) = \int e^{-2i\pi\langle x,\xi\rangle}\overline{\varphi(-x)}\,\mathrm{d}x = \overline{\int e^{-2i\pi\langle x',\xi\rangle}\varphi(x')\,\mathrm{d}x'} = \overline{\hat{\varphi}(\xi)},$$

where we have used the change of variable $x' = -x$. We finally obtain:

$$\|\varphi\|^2 = \int \hat{\varphi}(\xi)\hat{\psi}(\xi)\,\mathrm{d}\xi = \int \hat{\varphi}(\xi)\overline{\hat{\varphi}(\xi)}\,\mathrm{d}\xi = \int |\varphi(\xi)|^2\,\mathrm{d}\xi = \|\hat{\varphi}\|^2\,. \ ■$$

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

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

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

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

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

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

Proof. Let $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. By Theorem 4.12, one can find a sequence $(f_k) \subset \mathcal{C}_0^\infty$ such that $\|f_k - f\|_1 \to 0$ and $\|f_k - f\|_2 \to 0$ as $k \to \infty$. On the one hand $\hat{f}_k \to \hat{f}$ uniformly (and thus pointwise) as $k \to \infty$ (see the proof of Theorem 5.3), and on the other hand $\|\hat{f}_k - \mathbb{U}f\|_2 \to 0$ as $k \to \infty$ by Theorem 5.9. Statement (1) follows. Let now f be any member of $L^2(\mathbb{R}^n)$. By the first part of the proof, $\Phi_R = \mathbb{U}(f\mathbb{1}_{B_R})$, where B_R denotes the (closed) ball of radius R centered at the origin. Thus

$$\|\Phi_R - \mathbb{U}f\|_2 = \|\mathbb{U}(f\mathbb{1}_{B_R} - f)\|_2 = \|f\mathbb{1}_{B_R} - f\|_2$$

by Theorem 5.9, and the first half of (2) follows. Now, if we substitute $\mathbb{U}^{-1}f$ to f in the definition of Φ_R, we obtain:

$$\int e^{-2i\pi\langle\cdot,\xi\rangle}\mathbb{1}_{B_R}(\xi)(\mathbb{U}^{-1}f)(\xi)\,\mathrm{d}\xi \xrightarrow{\|\cdot\|_2} f \quad \text{as } R\to\infty.$$

Clearly, $(\mathbb{U}^{-1}f)(\xi)=(\mathbb{U}f)(-\xi)$, and the change of variable $\xi'=-\xi$ shows that

$$\int e^{-2i\pi\langle x,\xi\rangle}\mathbb{1}_{B_R}(\xi)(\mathbb{U}^{-1}f)(\xi)\,\mathrm{d}\xi=\int e^{2i\pi\langle x,\xi'\rangle}\mathbb{1}_{B_R}(\xi')(\mathbb{U}f)(\xi')\,\mathrm{d}\xi'=F_R(x).$$

The second half of (2) follows. ∎

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

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 the theory of analytic functions.

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

$$\|f(x)-f(x_0)-L(x-x_0)\|_F=o(\|x-x_0\|). \tag{5.6}$$

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 (5.6), 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:

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

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

(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 $x = (x_1,\dots,x_n) \in \mathbb{K}^n$. 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

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

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 $C_k^\alpha = C_{k_1}^{\alpha_1} \dots C_{k_n}^{\alpha_n}$. 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 $j \in \{1, \dots, n\}$.

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.

Proof. By Theorem 5.14, $f - g \in A(\Omega, F)$. Since $(f-g)^{-1}(0) \supset \Omega_1$, the function $f - g$ must be constant on Ω by Theorem 5.18. Thus $f = g$ 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

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

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 $\breve{f} \in A(\mathbb{C}^n, \mathbb{C})$.

Proof. For all $j \in \{1, \ldots, n\}$ and all $\zeta \in \mathbb{C}^n$, we have:

$$\begin{aligned} \frac{\partial \breve{f}}{\partial \xi_j}(\zeta) &= \int (-2i\pi x_j) e^{-2i\pi\langle x,\xi\rangle} e^{2\pi\langle x,\eta\rangle} f(x)\,\mathrm{d}x \\ \text{and} \quad \frac{\partial \breve{f}}{\partial \eta_j}(\zeta) &= \int (2\pi x_j) e^{-2i\pi\langle x,\xi\rangle} e^{2\pi\langle x,\eta\rangle} f(x)\,\mathrm{d}x. \end{aligned}$$

Clearly, $\partial \breve{f}/\partial \xi_j$ and $\partial \breve{f}/\partial \eta_j$ are continuous on $\mathbb{C}^n$, and the Cauchy-Riemann conditions

$$\frac{\partial \breve{f}}{\partial \eta_j}(\zeta) = i \frac{\partial \breve{f}}{\partial \xi_j}(\zeta), \quad j \in \{1, \ldots, n\}$$

are satisfied at every $\zeta \in \mathbb{C}^n$. By Theorem 5.11, the partial $\mathbb{C}$-derivatives $\partial \breve{f}/\partial \zeta_j$ exist. Now, there is no doubt that $\breve{f}$ is continuous on $\mathbb{C}^n$. Therefore, Theorem 5.19 implies that $\breve{f} \in A(\mathbb{C}^n, \mathbb{C})$. ∎

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

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

is injective.

Proof. The fact that $\hat{f}$ is analytic is immediate from the theorem. Now, suppose that $Af = 0$, that is, that the function $\hat{f}$ agrees with the function identically equal to 0 (which is obviously in $A(\mathbb{R}^n, \mathbb{C})$) on Ω. Corollary 5.1 then shows that $\hat{f} = 0$ on $\mathbb{R}^n$, and the result follows. ∎

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.

Proof. Since A is continuous, $\ker A$ is closed. Theorem 2.8 then shows that $F = \ker A \oplus (\ker A)^\perp$. Let us prove that $(\ker A)^\perp = \operatorname{cl}\operatorname{ran} A^\star$. We have:

$$\begin{aligned} x \in (\operatorname{ran} A^\star)^\perp &\iff \forall u \in \operatorname{ran} A^\star,\ \langle x, u\rangle = 0 \\ &\iff \forall y \in G,\ \langle x, A^\star y\rangle = 0 \\ &\iff \forall y \in G,\ \langle Ax, y\rangle = 0 \\ &\iff Ax \in G^\perp = \{0\} \\ &\iff x \in \ker A. \end{aligned}$$

Thus $(\operatorname{ran} A^\star)^\perp = \ker A$. Consequently, $(\ker A)^\perp = (\operatorname{ran} A^\star)^{\perp\perp} = \operatorname{cl}\operatorname{ran} A^\star$, and (1) is established. Since $A^\star$ is continuous and $A^{\star\star} = A$, (2) is immediate from (1).

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

SOLUTION. The fact that $\ker A \subset \ker A^\star A$ is obvious. Conversely,

$$A^\star Ax = 0 \Longrightarrow \langle x, A^\star Ax\rangle = 0 \Longrightarrow \|Ax\|^2 = 0 \Longrightarrow Ax = 0.$$

The second equality follows by symmetry. ∎

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 $\|y - Ax\|$;

(3) x_0 solves the *normal equation*: $A^\star y = A^\star Ax_0$.

PROOF. Note first that the existence of such an x_0, which is not assumed here, is pending the condition $\tilde{y} \in \operatorname{ran} A$. Let us prove the equivalence of (1) and (2). We have:

$$\begin{aligned}
\tilde{y} = Ax_0 &\iff \|y - Ax_0\| = \inf\{\|y - z\| \mid z \in \operatorname{cl}\operatorname{ran} A\} \\
&\iff \|y - Ax_0\| = \inf\{\|y - z\| \mid z \in \operatorname{ran} A\} \\
&\iff \|y - Ax_0\| = \inf\{\|y - Ax\| \mid x \in F\}
\end{aligned}$$

in which the second equivalence results from the continuity of the norm. Let us now prove the equivalence of (1) and (3). Recall that, by Theorem 2.18, $\tilde{y} = Py$ if and only if

$$\tilde{y} \in \operatorname{cl}\operatorname{ran} A \quad \text{and} \quad y - \tilde{y} \in (\operatorname{cl}\operatorname{ran} A)^\perp = \ker A^\star.$$

Thus $Ax_0 = \tilde{y}$ if and only if $y - Ax_0 \in \ker A^\star$, and the result follows. ∎

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 $\mathcal{D}(A^+) = \operatorname{ran} A + (\operatorname{ran} A)^\perp \subsetneq G$.

Proof. Let (f_k) be a Hilbert basis of F. The fact that $\operatorname{tr} A^\star A = \sum_k \|Af_k\|^2 < \infty$ implies that $\|Af_k\|$ tends to 0 as k tends to infinity. This shows that

$$\inf_{\|f\|=1} \|Af\| = 0,$$

and (1) is established. Now, it is clear that $A^\star A$ is hermitian, compact, nonnegative and injective. Therefore, (2) follows immediately from the spectral theorem for compact hermitian operators (see Theorem 3.5 and its corollaries). Finally, if $\operatorname{ran} A$ were closed, then $\operatorname{ran} A$ would be a Banach space, and $A^{-1}\colon \operatorname{ran} A \to F$ would be continuous by the Open Mapping Theorem, in contradiction with (1). Thus (3) is clear. ∎

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 $\operatorname{cl}\operatorname{ran} A$.

Proof. For all $k, l \in \mathbb{N}^*$, we have:

$$\langle g_k, g_l\rangle = \frac{1}{\sqrt{\lambda_k \lambda_l}} \langle A f_k, A f_l\rangle = \frac{1}{\sqrt{\lambda_k \lambda_l}} \langle f_k, A^\star A f_l\rangle = \sqrt{\frac{\lambda_l}{\lambda_k}} \langle f_k, f_l\rangle = \delta_{kl},$$

in which δ_{kl} denote the Kronecker symbol. Thus (g_k) is an orthonormal subset of G. Let us prove that $\mathrm{cl}\,\mathrm{vect}(g_k) = \mathrm{cl}\,\mathrm{ran}\,A$. Clearly, $\mathrm{vect}(g_k) \subset \mathrm{ran}\,A$, so that

$$\mathrm{cl}\,\mathrm{vect}(g_k) \subset \mathrm{cl}\,\mathrm{ran}\,A.$$

On the other hand, recall that $F = \{\sum_k \alpha_k f_k \mid \sum_k |\alpha_k|^2 < \infty\}$, so that

$$\mathrm{ran}\,A = \left\{ \sum_k \alpha_k \sqrt{\lambda_k} g_k \;\middle|\; \sum_k |\alpha_k|^2 < \infty \right\} = \left\{ \sum_k \beta_k g_k \;\middle|\; \sum_k \frac{|\beta_k|^2}{\lambda_k} < \infty \right\}.$$

Since the condition $\sum_k |\beta_k|^2/\lambda_k < \infty$ implies the condition $\sum_k |\beta_k|^2 < \infty$, we see that

$$\mathrm{ran}\,A \subset \left\{ \sum_k \beta_k g_k \;\middle|\; \sum_k |\beta_k|^2 < \infty \right\} = \mathrm{cl}\,\mathrm{vect}(g_k),$$

and the result follows. ∎

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 $\mathrm{cl}\,\mathrm{ran}\,A \setminus \mathrm{ran}\,A$. Since A is a Hilbert-Schmidt opeartor,

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

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

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

PROOF. Let $f \in F$. By continuity of A,

$$Af = \sum_k \langle f, f_k\rangle\, Af_k = \sum_k \sqrt{\lambda_k}\, \langle f, f_k\rangle\, g_k.$$

Let now $g \in \operatorname{cl}\operatorname{ran} A$. By continuity of $A^\star$,

$$A^\star g = \sum_k \langle g, g_k\rangle A^\star g_k = \sum_k \sqrt{\lambda_k}\, \langle g, g_k\rangle f_k.$$

Since $G = \operatorname{cl}\operatorname{ran} A \oplus \ker A^\star$, the formula remains true for all $g \in G$. Finally, let $g \in \mathcal{D}(A^+) = \operatorname{ran} A + (\operatorname{ran} A)^\perp$. Then $g = Af + g^\perp$, in which $f \in F$ and $g^\perp \in (\operatorname{ran} A)^\perp = \ker A^\star$. We have:

$$\langle g, g_k\rangle = \left\langle Af + g^\perp, g_k \right\rangle = \langle f, A^\star g_k\rangle = \sqrt{\lambda_k}\, \langle f, f_k\rangle .$$

Therefore, $f = \sum_k \langle f, f_k\rangle f_k = \sum_k \lambda_k^{-1/2}\langle g, g_k\rangle f_k$. In particular, the latter series is convergent. Let us now check that f satisfies the normal equation. By continuity of $A^\star A$,

$$\begin{aligned} A^\star A f &= \sum_k \lambda_k^{-1/2} \langle g, g_k\rangle A^\star A f_k \\ &= \sum_k \lambda_k^{1/2} \langle g, g_k\rangle f_k \\ &= A^\star g. \end{aligned}$$

Thus (3) follows from the fact that the normal equation has at most one solution, since $\ker A = \{0\}$. ∎

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

Proof. For all $g \in \mathcal{D}(A^+)$, we have

$$\left\|A^+ g\right\|^2 = \sum_k \lambda_k^{-1} \left|\langle g, g_k\rangle\right|^2$$

by Parseval's identity. In particular, $\|A^+ g_l\|^2 = \lambda_l^{-1} \to \infty$ as $l \to \infty$. ∎

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 $\|g_\varepsilon - g\| \le \varepsilon$. 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

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

(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

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

More explicitly, the function Af is given by

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

Since V and W are bounded, the kernel $\alpha(x,y) := e^{-2i\pi\langle x,y\rangle} 1\!\!1_V(x) 1\!\!1_W(y)$ 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

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

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

Proof. The spectral theorem attests the existence of an orthonormal sequence (f_k) and an decreasing sequence $\lambda_1 \geq \lambda_2 \geq \cdots$ of positive numbers such that (f_k) is an orthonormal basis of $L^2_V(\mathbb{R}^n)$ and the λ_k runs through the spectrum of $A^\star A$. Here,

$$\begin{aligned}
\lambda_1^2 &= \|A^\star A\|^2 \\
&= \|A^\star A f_1\|^2 \\
&= \int_V \left|(U^{-1}\mathbb{1}_W U f_1)(x)\right|^2 \,\mathrm{d}x \\
&\leq \int_{\mathbb{R}^n} \left|(U^{-1}\mathbb{1}_W \hat{f}_1)(x)\right|^2 \,\mathrm{d}x \\
&= \int_{\mathbb{R}^n} \left|(\mathbb{1}_W \hat{f}_1)(y)\right|^2 \,\mathrm{d}y.
\end{aligned}$$

Consequently, $\lambda_1^2 \leq \int_W |\hat{f}_1(y)|^2 \,\mathrm{d}y < 1$, for otherwise $\hat{f}_1$ would vanish on W^c, which would imply $f_1 = 0$. ■

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

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

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 $\|A^\star A\| \leq 1$ 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,

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

Proof. Since $\|T\| < 1$, the series $\sum \|T\|^k$ converges. But $\|T^k\| \leq \|T\|^k$, so that $\sum T^k$ converges in $L(F)$, since the latter space is complete. Let $R := \sum_{k=0}^{\infty} T^k$. We have:

$$TR = RT = \sum_{k=0}^{\infty} T^{k+1} = R - I.$$

Thus $I = R - TR = (I - T)R$ and $I = R - RT = R(I - T)$. Consequently, $I - T$ is invertible and $(I - T)^{-1} = R$. Finally,

$$\left\|(I - T)^{-1}\right\| = \left\|\sum_{k=0}^{\infty} T^k\right\| \leq \sum_{k=0}^{\infty} \left\|T^k\right\| = \frac{1}{1 - \|T\|}. \ \blacksquare$$

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

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

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 $\mu_1 \leq \mu_2 \leq \ldots$ such that $\mu_1 > 0$ and $\mu_k \to 1$ as $k \to \infty$.

Proof. Since $B\colon L^2_V(\mathbb{R}^n) \to \operatorname{ran} B$ is invertible, the restriction of B^+ to $\operatorname{ran} B$, also denoted by B^+, satisfies: $B^+ = (B^\star B)^{-1} B^\star = B^{-1}(B^\star)^{-1} B^\star = B^{-1}$. Therefore, the continuity of B^+ implies that of B^{-1}, and (1) is proved. Now, let $(g'_k) \subset \operatorname{ran} B$ be converging to an element $g' \in L^2_W(\mathbb{R}^n)$. Put $f'_k := B^{-1} g'_k$. Since B^{-1} is continuous, (f'_k) is a Cauchy sequence, and since $L^2_V(\mathbb{R}^n)$ is complete, (f'_k) converges to some $f' \in L^2_V(\mathbb{R}^n)$. By continuity of B, $Bf'_k = g'_k \to Bf'$. By uniquenes of the limit, $g' = Bf'$. Thus $g' \in \operatorname{ran} B$ and (2) is established. Finally, there exists a Hilbert basis (f_k) of $L^2_V(\mathbb{R}^n)$ and a decreasing sequence $(\lambda_k) \subset \mathbb{R}^*_+$ which converges to 0 such that, for all k, $A^\star A f_k = \lambda_k f_k$. Thus

$$\forall k, \quad B^\star B f_k = (I - A^\star A) f_k = (1 - \lambda_k) f_k,$$

which shows that the eigenspace of $B^\star B$ are the same as those of $A^\star A$, and that $\Lambda(B^\star B) = 1 - \Lambda(A^\star A)$. ■

As a conclusion, the inverse problem of Fourier interpolation 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 $g \in L^2_W(\mathbb{R}^n)$, $T_\beta g$ minimize the functional

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

in which

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

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

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

We then have:

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

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

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

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

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

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

(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.

SOLUTION.

(1) Since $k \in L^1(\mathbb{R}^n)$ and $f \in L^2_V(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$, we have: $k * f \in L^2(\mathbb{R}^n)$ and $\|Af\|_2 = \|k * f\|_2 \leq \|k\|_1 \|f\|_2$. Thus $\|A\| \leq \|k\|_1$.

(2) Since V is bounded, $L^2_V(\mathbb{R}^n) \subset L^1(\mathbb{R}^n)$, so that $f \in L^1(\mathbb{R}^n)$. Since both k and f belong to $L^1(\mathbb{R}^n)$, so does $k * f$, and

$$U(k * f) = Uk \cdot Uf.$$

Thus $Af := k * f = \overline{U}U(k * f) = \overline{U}(Uk \cdot Uf) = \overline{U}A_0 f$, where

$$\begin{array}{rccc} A_0\colon & L^2_V(\mathbb{R}^n) & \longrightarrow & L^2(\mathbb{R}^n) \\ & f & \longmapsto & A_0 f := \hat{k} \cdot \hat{f}. \end{array}$$

(3) We have:

$$\begin{aligned} (A_0 f)(y) &= \hat{k}(y)\hat{f}(y) \\ &= \hat{k}(y) \int_V e^{-2i\pi\langle x,y\rangle} f(x)\, \mathrm{d}x \\ &= \int e^{-2i\pi\langle x,y\rangle} \hat{k}(y) \mathbb{1}_V(x) f(x)\, \mathrm{d}x \\ &= \int \alpha(x,y) f(x)\, \mathrm{d}x, \end{aligned}$$

where $\alpha(x,y) := e^{-2i\pi\langle x,y\rangle}\hat{k}(y)\mathbb{1}_V(x)$. Since k belongs to $L^2(\mathbb{R}^n)$, so does $\hat{k}$. Consequently,

$$\int |\alpha(x,y)|^2\, \mathrm{d}x\, \mathrm{d}y = \operatorname{meas}(V) \cdot \int \left|\hat{k}(y)\right|^2\, \mathrm{d}y < \infty.$$

Thus $\alpha \in L^2(V \times \mathbb{R}^n)$ and A_0 is a Hilbert-Schmidt operator. Now, $A^\star A = A_0^\star U \overline{U} A_0 = A_0^\star A_0$. Thus $\operatorname{tr} A^\star A = \operatorname{tr} A_0^\star A_0 < \infty$, and A is also a Hilbert-Schmidt operator.

(4) Let Ω be an open subset of $\operatorname{supp}(Uk)$. Observe that $\|Af\|^2 = \|A_0 f\|^2 = \int |\hat{k}\hat{f}|^2$. Consequently,

$$\begin{aligned} Af = 0 \iff & \int \left|\hat{k}\hat{f}\right|^2 = 0 \\ \iff & \left|\hat{k}\hat{f}\right|^2 = 0 \text{ almost everywhere} \\ \implies & \left|\hat{f}\right|^2 = 0 \text{ almost everywhere in } \Omega \\ \iff & f = 0. \end{aligned}$$

Recall indeed that,

$$\begin{array}{cccc} A_\Omega\colon & L^2_V(\mathbb{R}^n) & \longrightarrow & L^2_\Omega(\mathbb{R}^n) \\ & f & \longmapsto & 1\!\!1_\Omega \hat{f} \end{array}$$

is injective. We conclude that the problem of *deconvolution*, which consists in recovering f from the knowledge of Af, is an ill-posed problem. ■

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

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

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 $k = (k_1, \ldots, k_2) \in \mathbb{Z}^n$, 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,

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

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

Proof. We have:

$$(U^{-1} 1\!\!1_{C_a})(x) = \int_{C_a} e^{2i\pi\langle x,\xi\rangle} \,\mathrm{d}\xi = \prod_{j=1}^{n} \int_{-a/2}^{a/2} e^{2i\pi x_j \xi_j} \,\mathrm{d}\xi_j.$$

If $x_j \neq 0$, then

$$\int_{-a/2}^{a/2} e^{2i\pi x_j \xi_j} \,\mathrm{d}\xi_j = \left[\frac{e^{2i\pi x_j \xi_j}}{2i\pi x_j} \right]_{-a/2}^{a/2} = \frac{1}{\pi x_j} \frac{e^{i\pi a x_j} - e^{-i\pi a x_j}}{2i} = a \frac{\sin \pi a x_j}{\pi a x_j}.$$

If $x_j = 0$, then $\int_{-a/2}^{a/2} e^{2i\pi x_j \xi_j} \,\mathrm{d}\xi_j = a$. Thus $(U^{-1} 1\!\!1_{C_a})(x) = \prod_j a \operatorname{sinc} \pi a x_j = a^n \operatorname{sinc} \pi a x$. $\blacksquare$

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

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

Proof. Note that L^2-convergence is meant in Equations (7.1) and (7.2), and that the analytical member of the class of f is assumed in expressions such as $f(x)$. We first prove Equation (7.2). By Theorem 7.2,

$$\hat{f} = \sum_{k\in\mathbb{Z}^n} \left\langle \hat{f}, \varphi_k \right\rangle \varphi_k.$$

We have:

$$\begin{aligned}
\left\langle \hat{f}, \varphi_k \right\rangle &= \int \hat{f}(\xi)\overline{\varphi_k(\xi)}\,\mathrm{d}\xi \\
&= a^{-n/2}\int_{C_a} \hat{f}(\xi)\exp\left[2i\pi\left\langle \frac{k}{a},\xi\right\rangle\right]\mathrm{d}\xi \\
&= a^{-n/2}\int_{\mathbb{R}^n} \hat{f}(\xi)\exp\left[2i\pi\left\langle \frac{k}{a},\xi\right\rangle\right]\mathrm{d}\xi \\
&= a^{-n/2} f\left(\frac{k}{a}\right),
\end{aligned}$$

in which the third and fourth equalities are due to the fact that $\hat{f} \in L^2_{C_a}(\mathbb{R}^n) \subset L^1_{C_a}(\mathbb{R}^n)$. Consequently,

$$\begin{aligned}
\hat{f}(\xi) &= \sum_{k\in\mathbb{Z}^n} a^{-n/2} f\left(\frac{k}{a}\right)\cdot a^{-n/2}\exp\left[-2i\pi\left\langle \frac{k}{a},\xi\right\rangle\right] \\
&= \frac{1}{a^n}\sum_{k\in\mathbb{Z}^n} f\left(\frac{k}{a}\right)\exp\left[-2i\pi\left\langle \frac{k}{a},\xi\right\rangle\right].
\end{aligned}$$

Next, we prove Equation (7.3). Since U is an isometry,

$$\langle f, g\rangle = \left\langle \hat{f}, \hat{g}\right\rangle = \sum_{k\in\mathbb{Z}^n}\left\langle \hat{f},\varphi_k\right\rangle \overline{\langle \hat{g},\varphi_k\rangle}$$

by Parseval's theorem. Consequently,

$$\langle f, g\rangle = \sum_{k\in\mathbb{Z}^n} a^{-n/2} f\left(\frac{k}{a}\right) a^{-n/2}\overline{g}\left(\frac{k}{a}\right) = \frac{1}{a^n}\sum_{k\in\mathbb{Z}^n} f\left(\frac{k}{a}\right)\overline{g}\left(\frac{k}{a}\right).$$

Finally, let us prove the interpolation formula. The function f is given by $f(x) = U^{-1}\hat{f}(x)$, where

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

By continuity of U^{-1}, we can write:

$$U^{-1}\hat{f}(x) = \frac{1}{a^n}\sum_{k\in\mathbb{Z}^n} f\left(\frac{k}{a}\right) U^{-1}\left(1\!\!1_{C_a}\cdot\exp\left[-2i\pi\left\langle\frac{k}{a},\cdot\right\rangle\right]\right)(x).$$

But

$$\begin{aligned} & U^{-1}\left(1\!\!1_{C_a}\cdot\exp\left[-2i\pi\left\langle\frac{k}{a},\cdot\right\rangle\right]\right)(x) \\ = \ & \int 1\!\!1_{C_a}(\xi)\exp\left[-2i\pi\left\langle\frac{k}{a},\xi\right\rangle\right]\exp\left[2i\pi\left\langle x,\xi\right\rangle\right]\,\mathrm{d}\xi \\ = \ & \int 1\!\!1_{C_a}(\xi)\exp\left[2i\pi\left\langle x-\frac{k}{a},\xi\right\rangle\right]\,\mathrm{d}\xi \\ = \ & (U^{-1}1\!\!1_{C_a})\left(x-\frac{k}{a}\right) \\ = \ & a^n\operatorname{sinc}\pi a\left(x-\frac{k}{a}\right) \end{aligned}$$

by Lemma 7.1. Therefore,

$$f(x) = \frac{1}{a^n}\sum_{k\in\mathbb{Z}^n} f\left(\frac{k}{a}\right)\cdot a^n\operatorname{sinc}\pi a\left(x-\frac{k}{a}\right),$$

and Equation (7.1) is proved. ∎

In the case where b is strictly less than 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:

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

Proof. Notice first that, since γ belongs $\mathcal{S}(\mathbb{R}^n)$, so does $\check{\gamma}$, and the *interpolation function* $x\mapsto\check{\gamma}(x)\operatorname{sinc}\pi ax$ decays at infinity much faster than

$x \mapsto \operatorname{sinc} \pi a x$. Therefore, then series in Equation (7.4) must converge faster than the series in Equation (7.1). Let $\psi := \gamma * \mathbb{1}_{C_a}$ and let $c := 2a - b$. By Theorem 4.9, ψ is a C^∞-function. If $\xi \in C_b$ and $\|\eta\| \leq (a-b)/2$, then $\xi - \eta \in C_a$. Thus, for all $\xi \in C_b$, we have:

$$\psi(\xi) = \int \gamma(\eta) \mathbb{1}_{C_a}(\xi - \eta) \, \mathrm{d}\eta = \int \gamma(\eta) \, \mathrm{d}\eta = 1.$$

Similarly, if $\xi \in \mathbb{R}^n \setminus C_c$ and $\|\eta\| \leq (a-b)/2$, then $\xi - \eta \in \mathbb{R}^n \setminus C_a$, so that

$$\psi(\xi) = \int \gamma(\eta) \mathbb{1}_{C_a}(\xi - \eta) \, \mathrm{d}\eta = 0.$$

By Theorem 7.3, the formula

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

holds for all $\xi \in C_a$. The a-periodic extension of $\hat{f}$ has support in the set $\cup_{k \in \mathbb{Z}^n}(ka + C_b)$. Therefore, the formula

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

holds for all $\xi \in \mathbb{R}^n$. Taking inverse Fourier transforms yields:

$$f(x) = \frac{1}{a^n} \sum_{k \in \mathbb{Z}^n} f\left(\frac{k}{a}\right) U^{-1}\left(\psi \exp\left[-2i\pi \left\langle \frac{k}{a}, \cdot \right\rangle\right]\right)(x).$$

But

$$\begin{aligned}
& U^{-1}\left(\psi \exp\left[-2i\pi \left\langle \frac{k}{a}, \cdot \right\rangle\right]\right)(x) \\
= & \int (\gamma * \mathbb{1}_{C_a})(\xi) \exp\left[-2i\pi \left\langle \frac{k}{a}, \xi \right\rangle\right] \exp\left[2i\pi \langle x, \xi \rangle\right] \mathrm{d}\xi \\
= & \int (\gamma * \mathbb{1}_{C_a})(\xi) \exp\left[2i\pi \left\langle x - \frac{k}{a}, \xi \right\rangle\right] \mathrm{d}\xi \\
= & (U^{-1}(\gamma * \mathbb{1}_{C_a})(\xi))\left(x - \frac{k}{a}\right) \\
= & (U^{-1}\gamma \cdot U^{-1}\mathbb{1}_{C_a})\left(x - \frac{k}{a}\right) \\
= & \check{\gamma}\left(x - \frac{k}{a}\right) \cdot a^n \operatorname{sinc} \pi a \left(x - \frac{k}{a}\right),
\end{aligned}$$

whence the result. ∎