Lecture notes, Harmonic Analysis minicourse

Joaquim Bruna, UAB

April 19, 2016

Chapter 1

Preliminaries in Real and Functional Analysis

1.1 Lebesgue Integrable functions.

Let (X, μ) a measure space, μ being a σ-finite positive measure defined on a σ-algebra. Most of the time,$X = \mathbf{R^d}$ with $d\mu = dm$, the Lebesgue measure (more generally a locally compact topological space, σ-compact, and $d\mu$ a Borel measure on X). Also it is to keep in mind the discrete setting, $X = \mathbf{Z^d}$ with counting measure, and the periodic setting $X = \mathbf{R^d}/\mathbf{Z^d}$. Through the map

$$(x_1, x_2, \dots, x_d) \mapsto (e^{2\pi i x_1}, e^{2\pi i x_2}, \dots, e^{2\pi i x_d}),$$

we identify $\mathbf{R^d}/\mathbf{Z^d}$ with $\mathbf{T^d}$, the d-dimensional torus, where $\mathbf{T} = \{\mathbf{z} \in \mathbf{C} : |\mathbf{z}| = \mathbf{1}\}$.

Using approximation by simple functions (linear combinations of indicator functions of measurable sets), every measurable function defined on X taking values in $[0, +\infty]$ has a well defined integral $\int_X f\, d\mu$, finite or $+\infty$. For $p > 0$, the space $L^p(X)$ consists of the measurable complex-valued functions such that

$$\|f\|_p = (\int_X |f(x)|^p\, dm(x))^{\frac{1}{p}} < +\infty.$$

It is thus a question of the size $|f|$. Functions in $L^1(X)$ are called integrable.

Functions in L^p are finite a.e. We identify functions which are equal a.e. The space $L^\infty(X)$ consists of complex-valued functions which are bounded a.e., $\|f\|_\infty$ being the least essential supnorm.

When X is a topological space, $L^p_{loc}(X)$ consists of functions whose restrictions to compact sets Y belong to $L^p(Y)$.

1.2 Convergence theorems.

Monotone convergence theorem: if f_n are positive and a.e. non-decreasing, and $f(x) = \lim_n f_n(x) \leq +\infty$, then $\int_X f d\mu = \lim_n \int_X f_n d\mu$. For general sequences, $\int_X \liminf f_n d\mu \leq \liminf \int_X f_n d\mu$. As a consequence, $\int_X (\sum_n g_n) d\mu = \sum_n \int_X g_n d\mu$ for arbitrary positive functions.
Dominated convergence theorem: if $f(x) = \lim_n f_n(x)$ a.e. x , $|f_n(x)| \leq g(x)$ a.e. x with $\int_X g d\mu < +\infty$, then f is integrable and $\int_X f d\mu = \lim_n \int_X f_n d\mu$. As a consequence, if $\sum_n \int_X |g_n| d\mu < +\infty$, the series $\sum_n g_n(x)$ absolutely converges a.e. and $\int_X (\sum_n g_n(x)) d\mu = \sum_n \int_X g_n d\mu$.

1.3 Fubini's theorem

If $(X, \mu), (Y, \nu)$ are measure spaces there is a natural estructure of measure space in $X \times Y$ with the product measure $\mu \times \nu$. In case $X = \mathbf{R^d}, \mathbf{Y} = \mathbf{R^m}$, of course the product measure of Lebesgue measures is Lebesgue measure.

Fubini's theorem has two parts: a) for a positive measurable function, the double integral $\int\int_{X \times Y} f(x, y) d\mu(x) d\nu(y)$ and the two iterated integrals $\int_X (\int_Y f(x, y) d\nu(y)) d\mu(x)$, $\int_Y (\int_X f(x, y) d\mu(x)) d\nu(y)$ are equal, finite or not. b) If $f(x, y)$ is integrable in $X \times Y$, then both iterated integrals make sense and are finite (meaning that for a.e. $x, f(x, y)$ is integrable in Y and the function so defined a.e. in X is integrable in X, and the other way around) and equal to the double integral.

Normally we use both parts in the following form: if one of the two iterated integrals of $|f(x, y)|$ is finite, then all three integrals are finite and equal. In particular the iterated integrals are a.e. absolutely convergent.

1.4 Holder's and Minkowski's inequalities

Holder's inequality reads as follows: If $1 \leq p, q, r \leq +\infty, \frac{1}{p} + \frac{1}{q} = \frac{1}{r}$, and $f \in L^p, g \in L^q$, then $fg \in L^r$ and $\|fg\|_r \leq \|f\|_p \|g\|_q$. For $p = q = 2$ this is known as Cauchy-Schwartz inequality.

Recall that this depends on Young's inequality and that there is a version with n indexes, i.e., if $\frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_n} = \frac{1}{r}$, then

$$\|f_1 f_2 \dots f_n\|_r \leq \|f_1\|_{p_1} \cdots \|f_n\|_{p_n}.$$

In the opposite direction, if f is measurable and

$$\sup\{\int_X |fg| d\mu : \|g\|_{p'} \leq 1\} < +\infty,$$

p' being the conjugate exponent of p, $\frac{1}{p} + \frac{1}{p'} = 1$, then $f \in L^p$ and

$$\|f\|_p = \sup\{\left|\int_X fg d\mu\right| : \|g\|_{p'} \leq 1\}.$$

We call this the *reverse Hlder inequality*. In many occasions we already know that $f \in L^p$ and use this to estimate its norm.

From Holder's one obtains the classical Minkowski or triangle inequality: $\|f+g\|_p \leq \|f\|_p + \|g\|_p, 1 \leq p \leq +\infty$, so $\|\|_p$ is a norm if $p \geq 1$. For $p < 1$, since $|a+b|^p \leq |a|^p + |b|^p$, the quantity $\|\cdot\|_p^p$ is subadditive:

$$\|\sum_n f_n\|_p^p \leq \sum_n \|f_n\|_p^p.$$

Of course this goes over to finite sums and to infinite sums as well, in the form of the so-called *continuous Minkowski inequality*: if we have an infinite sum of functions, say

$$G(y) = \int_X F(x,y) d\mu(x), G = \int_X F_x d\mu(x), |G| \leq \int_X |F_x| d\mu(x),$$

then

$$\|G\|_p \leq \int_X \|F_y\|_p d\mu(x), 1 \leq p$$

that is

$$\left(\int_Y \left(\int_X |F(x,y)| d\mu(x)\right)^p d\nu(y)\right)^{\frac{1}{p}} \leq \int_X \left(\int_Y |F(x,y)|^p d\nu(y)\right)^{\frac{1}{p}} d\mu(x).$$

One can prove this by approximating integrals by finite sums. An alternative proof is as follows. For $p = 1$ it is just Fubini's theorem and for $p = +\infty$ is trivial. To prove it for $1 < p < +\infty$ we may assume that $F \geq 0$. We have, using Fubini and Holder

$$\int_Y G(y)^p d\nu(y) = \int_Y G(y)^{p-1} G(y) d\nu(y) = \int_Y G(y)^{p-1} (\int_X F(x,y) d\mu(x)) d\nu(y) =$$

$$= \int_X (\int_Y F(x,y) G(y)^{p-1} d\nu(y)) d\mu(x) \leq$$

$$\leq \int_X (\int_Y F(x,y)^p d\nu(y))^{\frac{1}{p}} (\int_Y G(y)^p d\nu(y))^{\frac{1}{p'}} d\mu(x) =$$

$$= (\int_Y G(y)^p d\nu(y))^{\frac{1}{p'}} \int_X (\int_Y F(x,y)^p d\nu(y))^{\frac{1}{p}} d\mu(x).$$

The same is true replacing the L^1- norm in x above by an L^r-norm with $1 \leq r \leq p$:

$$\left(\int_Y \left(\int_X |F(x,y)|^r d\mu(x)\right)^{\frac{p}{r}} d\nu(y)\right)^{\frac{1}{p}} \leq \left(\int_X \left(\int_Y |F(x,y)|^p d\nu(y)\right)^{\frac{r}{p}} d\mu(x)\right)^{\frac{1}{r}}.$$

In fact this follows from the above applied to $|F|^r$ with p replaced by $\frac{p}{r}$.

1.5 The $L^p(X)$ spaces as Banach spaces

The $L^p(X)$-spaces, $1 \leq p \leq +\infty$ equipped with the L^p norm are complete spaces, meaning that every Cauchy sequence, i.e. $\|f_n - f_m\| \to 0$ as $n, m \to +\infty$, is convergent to some $f \in L^p(X)$, that is $\|f_n - f\| \to 0$ as $n \to +\infty$. If $p < +\infty$, convergence of f_n to f in L^p implies pointwise a.e. convergence of a partial subsequence; if $p = +\infty$, we have of course uniform convergence a.e. For $p = 2$, $L^p(X)$ is a Hilbert space with the scalar product defined by

$$\langle f, g\rangle = \int_X f(x)\overline{g(x)}d\mu(x).$$

If X is an open set U in $\mathbf{R^d}$, the space $C_c(U)$ of continuous compactly supported functions is dense in $L^p(U), 1 \leq p < +\infty$, meaning that an arbitrary $f \in L^p(U)$ can be approximated by a sequence $g_n \in C_c(U)$: $\|f - g_n\|_p \to 0$ as $n \to +\infty$. In case $p = +\infty$ the space in which $C_c(U)$ is dense is the space $C_0(U)$ of continuous functions vanishing at infinity, meaning that for each ε there is a compact set in U out of which $|f| < \varepsilon$. In the periodic setting $C(\mathbf{T^n})$ is dense in the $L^p(\mathbf{T^n})$ for $1 \leq p < +\infty$ and is already complete with the sup-norm.

Recall that the linear maps $T : E \to F$ between linear spaces that are continuous are the bounded ones, that is, those for which $\|T(f)\|_F \leq C\|f\|_E$ for some constant $C > 0$, the infimum of such C being the norm of T. The density of a subspace G of E is useful to define linear continuous maps $T : E \to F$ between Banach spaces; namely every linear continuous maps $T : G \to F$ extends uniquely to E. This is so because T maps Cauchy sequences to Cauchy sequences, hence if f is approximated by $g_n \in G$, Tf can be defined by the limit of the Cauchy sequence $T(g_n)$.

We define the translation operator $\tau_x, x \in \mathbf{R^d}$ acting on functions f defined on $\mathbf{R^d}$ by $(\tau_x f)(y) = f(y - x)$. Obviously $\|\tau_x f\|_p = \|f\|_p$. For fixed $f \in L^p(\mathbf{R^d}), \mathbf{1} \leq \mathbf{p} < +\infty$, the map $x \to \tau_x f$ is continuous. It is enough proving continuity at zero. This is clear for $g \in C_c(\mathbf{R^d})$ because $\tau_x g \to g$ uniformly as $x \to 0$; since the $\tau_x g$ have a common compact support, one has then $\|\tau_x g - g\|_p \to 0$ as well. Once we have established this for g in the dense space $C_c(\mathbf{R^d})$ the general case follows through

$$\|\tau_x f - f\| \leq \|\tau_x f - \tau_x g\| + \|\tau_x g - g\| + \|g - f\| = 2\|f - g\| + \|\tau_x g - g\|.$$

In the periodic situation, translation becomes rotation, of course.

The reverse Hlder inequality can be used to prove that the dual space of $L^p(X), 1 \leq p < +\infty$ is $L^{p'}(X)$, meaning that the general form of a continuous linear form $\omega : L^p(X) \to \mathbf{C}$ is of the form

$$f \mapsto \int_f g \, d\mu,$$

for some $g \in L^{p'}(X)$.

1.6 Riesz representation theorems

There are several versions of the Riesz representation theorem, the contents of each of them being the identification of the dual of some space of continuous functions.

Let X be a locally compact topological space which is σ-compact, such as $\mathbf{R^d}$ or an open set. A *Radon measure* μ on X is a positive Borel measure on X which is locally finite and both inner and outer regular. Exactly as with Lebesgue measure, $C_c(X)$ is dense in $L^p(X, \mu)$ for $1 \leq p < +\infty$, while its completion with the sup-norm is the space $C_0(X)$ of continuous functions vanishing at infinity.

The most simple example of Radon measure is the one having mass one at x and zero elsewhere.

$$\int_X f d\delta_x = f(x).$$

If μ is a Radon measure on X then

$$L(f) = \int_X f d\mu$$

is a linear functional on $C_c(X)$ which is positive, that is $L(f) \geq 0$ for $f \geq 0$. Conversely, the first Riesz representation states that every positive linear functional is of this type; more precisely, given L there is a unique Radon measure such that L is given as above.

A first consequence of this result is the description of the dual of $C_c(X)$ equipped with the sup-norm, or what amounts to the same, the description of the dual of $C(X)$, the space of complex-valued continuous functions on a compact topological space, equipped with the sup-norm. The Riesz representation theorem establishes that there is a one-to-one correspondence between the dual space of $C(X)$ and the regular complex Borel measures μ on X. Namely, the general form of a continuous linear functional L on $C(X)$ is the one defined by such a measure through

$$L(f) = \int_X f d\mu,$$

and $\|L\| = \|\mu\|$, the total variation mass of μ, which is always finite for complex measures. Since $d\mu = hd|\mu|$ for some h with $|h| = 1$, we may think that

$$\int_X f d\mu = \int_X f h d|\mu|,$$

is the definition of the integral with respect to $d\mu$.

Sometimes $C_c(X)$ is viewed as equipped with the so called inductive topology (convergence of a sequence $f_n \to f$ means that all f_n have their support within a fixed compact set and f_n converges uniformly there), for which it is complete. The Riesz representation theorem can be used to show that the general form of a continuous linear functional L in this case is

$$L(f) = \int_X f h d\rho, f \in C_c(X),$$

where $d\rho$ is a Radon measure on X and $h \in L^1_{loc}(X, \rho)$.

Along the same lines, the general form of a continuous linear functional on the space $C(U)$ of all continuous functions in an open set with the topology of uniform convergence on compacts subsets of U is

$$L(f) = \int_U f d\mu,$$

where $d\mu$ is a regular complex measure with compact support.

The statement that $L^{p'}(X)$ is the dual space of $L^p(X), 1 \leq p < +\infty$ is also known as a Riesz-type representation theorem.

In all these cases an important result is the Banach-Alouglu theorem: if one has a uniformly bounded family f_ε of $L^{p'}(X)$ functions, then there exists a subsequence ε_n such that f_{ε_n} has a weak limit $f \in L^{p'}$, that is

$$\lim_n \int_X g(x) f_{\varepsilon_n}(x)\, d\mu(x) = \int_X g(x) f(x) d\mu(x), g \in L^p(X).$$

Similarly, if μ_ε is a family of complex valued measures with uniformly bounded mass, then there exists a complex valued (finite) measure μ and a sequence $\varepsilon_n \to 0$ such that

$$\lim_n \int_X g(x) d\mu_{\varepsilon_n}(x) = \int_X g(x) d\mu(x), g \in C_0(X).$$

The Banach-Alaouglu theorem is a powerful existence theorem.

1.7 Operators between the L^p spaces

We will be dealing with linear maps among the $L^p(X)$, that is, operators T that map functions $f(x)$ to functions $T(f)(y)$ linearly. In case of finite dimensional dimensional signals, that is, $x = 1, 2, \ldots, n, f$ is a vector

$(u_1, u_2, \ldots, u_n)$ in $\mathbf{C^n}$, and Tf is a vector $(v_1, \ldots, v_m)$ in $\mathbf{C^m}$, T is given by an $m \times n$ matrix,

$$v_i = \sum_{j=1}^{n} a_{ij} u_j, i = 1, \ldots, m.$$

1.7.1 Sequence spaces

Similarly, a linear operator mapping infinite sequences $u = (u_j)$ to infinite sequences $Tu = v = (v_i)$ is, loosely speaking, given by a doubly infinite matrix a_{ij},

$$v_i = \sum_{j=1}^{\infty} a_{ij} u_j.$$

Of course here we should take care of convergence questions. Let us give us a precise example. Suppose that $T : l^p(\mathbf{Z}) \to \mathbf{l^q}(\mathbf{Z})$ is linear and continuous, that is, $\|Tu\|_q \leq C\|u\|_p$ for some constant C and all sequences $u \in l^p(\mathbf{Z})$. Denote by δ_j the sequence that has 1 in position j and zero elsewhere. Since $\sum_j |u_j|^p$ is finite, the representation $u = \sum_j u_j \delta_j$ is convergent in $l^p(\mathbf{Z})$ and so $Tu = \sum_j u_j T(\delta_j)$, with convergence in $l^q(\mathbf{Z})$. This means that T is determined by the $T(\delta_j)$. If $T(\delta_j) = (a_{ij})$, the convergence in $l^q(\mathbf{Z})$ of $Tu = \sum_j u_j T(\delta_j)$ implies pointwise convergence and so we must have

$$v_i = \sum_j a_{ij} u_j, i \in \mathbf{Z}.$$

This shows that T must be given by a matrix (a_{ij}) satisfying

$$\sup_j (\sum_i |a_{ij}|^q)^{\frac{1}{q}} = C < +\infty, \tag{1.1}$$

(that is the columns are uniformly in l^q)and

$$|v_i| = |\sum_{j=-\infty}^{+\infty} a_{ij} u_j| \leq \|Tu\|_q \leq M\|u\|_p.$$

The later implies that all files must be uniformly in $l^{p'}$:

$$\sup_i (\sum_j |a_{ij}|^{p'})^{\frac{1}{p}'} \leq M < +\infty. \tag{1.2}$$

This does not mean that a matrix (a_{ij}) satisfying conditions (1.1) and (1.2) defines a linear continuous map T from $l^p(\mathbf{Z})$ to $l^q(Z)$; it is in general a difficult problem to characterize exactly those matrices. In case $p = 1$, however, these two necessary conditions are also sufficient. Observe first that condition (2) reads in this case $\sup_j |a_{ij}| \leq \|T\|$ and follows from (1). Since

$$(\sum_i |v_i|^q)^{\frac{1}{q}} \leq (\sum_i (\sum_j |a_{ij}||u_j|)^q)^{\frac{1}{q}},$$

the continuous Minkowski inequality gives that this is bounded by

$$\sum_j (\sum_i |a_{ij}|^q |u_j|^q)^{\frac{1}{q}} = \sum_j |u_j| (\sum_i |a_{ij}|^q)^{\frac{1}{q}} \leq C\|u\|_1.$$

In an analogous way, the case $q = +\infty$ can be described as well. Here condition (1) reads $\sup_{i,j} |a_{ij}| \leq C$ and follows from condition (2), and since by Holder's inequality

$$|v_i| \leq \sum_j |a_{ij}||u_j| \leq (\sum_j |a_{ij}|^{p'})^{\frac{1}{p}'} \|u\|_p \leq M\|u\|_p,$$

we are done.

1.7.2 $L^p(X)$-spaces

Thus, continuous linear maps between the $l^p(\mathbf{Z})$ spaces are given by infinite matrices. In an analogous way, an operator from $L^p(X)$ to $L^q(Y)$ for general X, Y is also given in loose terms by an infinite matrix or kernel $K(x, y)$ in the sense that

$$T(f)(y) = \int_X K(x, y) f(x) d\mu(x), y \in Y.$$

The informal and non-rigorous argument is as before. If δ_x denotes the "function equal to 1 at x and zero elsewhere", we may think that

$$f = \int_X f(x)\delta_x d\mu(x)$$

and so by linearity and continuty $Tf = \int_X f(x)(T\delta_x)d\mu(x)$. If $T(\delta_x)(y) = K_x(y) = K(x, y)$, we have

$$Tf(y) = \int_X f(x)K(x, y)d\mu(x).$$

Generically speaking, they are called *integral operators.* We will see later rigorous arguments along this lines, but by now it should be clear that to do that one needs to enlarge the notion of function; for instance, the identity operator should correspond to the "function" $\delta(x, y)$ which equals 1 when $x = y$ and 0 otherwise.

Very often we will be considering the following situation: a dense subspace D_X of $L^p(X)$ (usually the space $\mathbf{C_c}(\mathbf{R^d})$ of continuous with compact suport when $X = \mathbf{R^d}$ or the space $C_c^\infty(\mathbf{R^d})$ of compactly supported infinitely differentiable functions) and an integral operator

$$Tf(y) = \int_X K(x,y)f(x)d\mu(x),$$

that makes sense for $f \in D_X$ and a.e. y (or all y). If we prove that $\|Tf\|_q \leq C\|f\|_p$, and $1 \leq p < +\infty$, then T extends to a continuous linear mapping defined on the whole of $L^p(X)$. In some cases, as e.g. with the Fourier transform in L^2 or with CZO integrals, it is hard to prove that the extension is given a.e. by the same expression.

Duality is a very important tool to deal with inequalities. In the above situation, usually there is a similar space D_Y in $L^{q'}(Y)$ such that

$$(Tf,g) = \int_X \int_Y K(x,y)f(x)g(y)d\mu(x)d\nu(y)$$

is well defined for $f \in D_X, g \in D_Y$ and one can use Fubini to write $(Tf,g) = (f, T^*g)$ with

$$T^*(g)(x) = \int_Y K(x,y)g(y)d\nu(y).$$

This is the dual map. If $1 \leq p < +\infty$, then by the reverse Holder inequality, $\|Tf\|_q \leq C\|f\|_p$ implies $\|T^*g\|_{p'} \leq C\|g\|_{q'}$, and so both are equivalent if $1 < p < +\infty$. Note however that in case $p = +\infty$ if we know a priori that $T^*g \in L^1(X)$ it is still true that $\|Tf\|_q \leq C\|f\|_\infty$ implies $\|T^*g\|_1 \leq C\|g\|_{q'}$,

1.7.3 Integral operators depending only on size

To finish this section we consider conditions on the kernel K that ensure that T is bounded from $L^p(X)$ to $L^q(Y)$. We emphasize that all these criteria depend just on the size $|K|$, that is to say they are in fact properties of the operator $T_{|K|}$ defined by $|K|$, and consequence of the trivial estimate $|T_K(f)(x)| \leq T_{|K|}(|f|)(x)$. Along the same lines, all these criteria provide operators for which

$$Tf(y) = \int_Y K(x,y)f(x)d\mu(x)$$

converges absolutely for a.e. y. In later chapters we will encounters much subtle situations in which boundedness of T depends on cancellation properties of K and pointwise convergence becomes a delicate issue.

First, as with sequence spaces, bounded mappings from $L^1(X)$ to Ł$^q(Y)$ can be characterized by the condition

$$\sup_x \left(\int_Y |K(x,y)|^q d\nu(y)\right)^{\frac{1}{q}} \leq C.$$

If this holds true, then Tf is a.e. defined for all $f \in L^1(X)$ and T is bounded from $L^1(X)$ to $L^q(Y)$ as a consequence of Minkowski continuous inequality

$$\| \int_X f(x)K(x,y)d\mu(x)\|_{L^q(Y)} \leq \int_X \|f(x)K(x,y)\|_{L^q(Y)}d\mu(x) = \int_X |f(x)|\|K(x,y)\|_{L^q(Y)}d\mu(x) \leq C\|f\|_1.$$

In fact one can prove that

$$\|T\| = \sup_x (\int_Y |K(x,y)|^q d\nu(y))^{\frac{1}{q}}.$$

Also as before, bounded maps on $L^\infty(Y)$ are described by the condition

$$\|T\| = \sup_y (\int_X |K(x,y)|^{p'} d\mu(x))^{\frac{1}{p'}}$$

by using simply the Holder inequality.

In case $p = +\infty$, then

$$|Tf(y)| \leq \|f\|_\infty \int_X |K(x,y)|d\mu(x)$$

and so T is bounded from $L^\infty(X)$ to Ł$^q(Y)$ if

$$\int_Y (\int_X |K(x,y)|d\mu(x))^q d\nu(y) < +\infty$$

the converse being true if $K \geq 0$.

In case $q = 1$, since

$$\int_Y |Tf(y)|d\nu(y) \leq \int_X |f(x)|(\int_Y |K(x,y)|d\nu(y))d\mu(x)$$

T is bounded from $L^p(X)$ to $L^1(Y)$ if

$$\int_X (\int_Y |K(x,y)|d\nu(y))^{p'} d\mu(x) < +\infty,$$

the converse being true again if $K \geq 0$.

In other cases there are no explicit necessary and sufficient conditions in terms of K so that T is bounded from $L^p(X)$ to $L^q(Y)$. We collect here a number of criteria.

Theorem 1. (Schur's) Suppose $1 \leq p, q, r \leq +\infty$, $\frac{1}{p} + \frac{1}{r} = 1 + \frac{1}{q}$ and

$$\sup_x (\int_Y |K(x,y)|^r d\nu(y))^{\frac{1}{r}} \leq C$$
$$\sup_y (\int_X |K(x,y)|^r d\mu(x))^{\frac{1}{r}} \leq C.$$

Then T is bounded from $L^p(X)$ to $L^p(Y)$ with constant $\|T\| \leq C$.

Proof. If $q = +\infty$ it follows from Holder's inequality, while the cases $r = +\infty$ or $p = +\infty$ are trivial; so we assume that all indexes are finite.

The hypothesis imply that

$$\frac{1}{r'} + \frac{1}{q} + \frac{1}{p'} = 1, \frac{p}{q} + \frac{p}{q'} = 1, \frac{r}{q} + \frac{r}{p'} = 1.$$

Using Hlder's inequality with r', q, p',

$$|Tf(y)| \leq \int_X |f(x)|^{\frac{p}{r'}} |f(x)|^{\frac{p}{q}} |K(x,y)|^{\frac{r}{q}} |K(x,y)|^{\frac{r}{p'}} d\mu(x) \leq$$

$$\leq \|f\|_p^{\frac{p}{r'}} \left(\int_X |f(x)|^p |K(x,y)|^r d\mu(x) \right)^{\frac{1}{q}} \left(\int_X |K(x,y)|^r d\mu(x) \right)^{\frac{1}{p'}} \leq$$

$$\leq C^{\frac{r}{p'}} \|f\|_p^{\frac{p}{r'}} \left(\int_X |f(x)|^p |K(x,y)|^r d\mu(x) \right)^{\frac{1}{q}}$$

Raising to q and integrating in y gives

$$\|Tf\|_q \leq C^{\frac{r}{p'}} \|f\|_p^{\frac{p}{r'}} \left(\int_X \int_Y |f(x)|^p |K(x,y)|^r d\mu(x) \right)^{\frac{1}{q}} =$$

$$= C^{\frac{r}{p'}} \|f\|_p^{\frac{p}{r'}} C^{\frac{r}{q}} \|f\|_p^{\frac{p}{q}} = C\|f\|_p$$

□

Theorem 2. (Schur's) Suppose $1 < p < +\infty$ and that there exists functions h_X, h_Y such that

$$\int_X |K(x,y)| h_X(x)^{p'} d\mu(x) \leq C h_Y(y)^{p'}$$

$$\int_Y |K(x,y)| h_Y(y)^p d\nu(y) \leq C h_X(x)^p$$

Then T is bounded from $L^p(X)$ to $L^p(Y)$ with norm at most C.

Proof. Assuming $K \geq 0$, we write

$$|Tf(y)| \leq \int_X K(x,y) |f(x)| h_X(x) h_X^{-1}(x) d\mu(x) =$$

$$= \int_X K^{\frac{1}{p'}} h_X(x) K^{\frac{1}{p}} |f(x)| h_X^{-1}(x) d\mu(x) \leq$$

$$\leq (\int_X K(x,y) h_X(x)^q d\mu(x))^{\frac{1}{q}} (\int_X K(x,y) h_X(x)^{-p} |f(x)|^p d\mu(x))^{\frac{1}{p}} \leq$$

$$\leq C^{\frac{1}{q}} h_Y(y) (\int_X K(x,y) h_X(x)^{-p} |f(x)|^p d\mu(x))^{\frac{1}{p}}$$

so that

$$|Tf(y)|^p \leq C^{\frac{p}{q}} h_Y(y)^p \int_X K(x,y) h_X(x)^{-p} |f(x)|^p d\mu(x).$$

Now, integrating in y and using the hypothesis again the proof is finished. □

Note than when $X = Y$ and K is symmetric, $K(x,y) = K(y,x)$, T is bounded in $L^2(X)$ if there exists $h > 0$ a.e. such that $T(h) \leq Ch$ for some constant C. The converse is true if K is positive; indeed if T is bounded and $\lambda > \|T\|$, and $f > 0$ a.e. is in $L^2(X)$ consider

$$h(x) = \sum_n \frac{T^n f}{\lambda^n}.$$

Since $\|T^n f\|_2 \leq \|T\|^n \|f\|_2$, the series is convergent in $L^2(X)$ and defines $h \in L^2(X)$ such that $Th \leq \lambda h$.

The operators satisfying the hypothesis of next theorem are called *Hilbert-Schmidt operators*

Theorem 3. If $K \in L^2(X \times Y)$, i.e.

$$\int_X \int_Y |K(x,y)|^2 d\mu(x) d\nu(y) = C < +\infty,$$

then T is bounded from $L^2(X)$ to $L^2(Y)$ with norm at most C.

Proof. Just notice that by Hlder's inequality

$$|Tf(y)|^2 \leq \|f\|_2^2 \int_Y |K(x,y)|^2 d\mu(x).$$

□

1.8 Hilbert spaces. Different notions of basis

The Euclidean space of countable or continuum dimensions is formally introduced with the notion of *H*ilbert space. Recall that a Hilbert space H is a linear space over $\mathbf{C}$ endowed with a positive definite hermitian product $\langle u, v\rangle$ so that with the norm $\|u\| = \sqrt{\|u, u\|}$ is complete. Completeness ensures that the familiar results in finite dimensional linear algebra still hold in this context. In particular, if F is a closed subspace of H, every $u \in H$ has a well defined projection $P_F(u) \in F$ such that $\|u - P_F(u)\|$ realizes the distance from u to F and $u - P_F(u)$ is orthogonal to F; if $v = u - P_F(u), u = P_F(u) + v$ is the unique decomposition of u as a sum of a vector in F and a vector in F^o, the orthogonal of F, and $\|u\|^2 = \|P_F(u)\|^2 + \|v\|^2$.

Besides the finite dimensional spaces $\mathbf{C^n}$ (mappings $\{1,2,...,n\} \to \mathbf{C}$), we single out the space of countable dimension $l^2(\mathbf{Z})$ consisting of sequences $u = (u_n)_{n\in\mathbf{Z}}$ such that $\|u\|^2 = \sum_n |u_n|^2 < +\infty$. This is the space $L^2(X)$ above when $X = \mathbf{Z}$ is equipped with the counting measure (we could use as well $\mathbf{N}$ instead of $\mathbf{Z}$ as index set). Analogously, the space $l^p(\mathbf{Z})$ consists of sequences such that $\sum_n |u_n|^p < +\infty$.

$L^2(X)$ should be thought as the Euclidean space $\mathbf{C^X}$ of as many dimensions as the cardinal of X. Finite n-dimensional vectors (signals) $v = (v_i)_{i=1,...,n}$ are replaced by $f = (f(x))_{x\in X}$, the finite sums $\|v\|^2 = \sum_{i=1}^n |v_i|^2$ by the infinite sum (integral) $\|f\|^2$, and the scalar product $\sum_i v_i\overline{u_i}$ replaced by $\int_X f(x)\overline{g(x)}d\mu(x)$. As with vectors, the quantity

$$\rho(f,g) = \frac{|\langle f,g\rangle|}{\|f\|_2\|g\|_2},$$

which by Schwarz inequality satisfies $0 \le \rho(f,g) \le 1$ measures the degree of linear dependence (correlation) between f, g: if $\rho(f,g) = 1, f, g$ are linearly dependent.

A family $(e_i)_{i\in I}$ of vectors in a Hilbert space H is called an *orthonormal basis* if: a) $\langle e_i, e_j\rangle = \delta_{i,j}$ and b) their finite linear combinations are dense in H. For the time being we may think that I is countable, say $I = \mathbf{Z}$ or $I = N$.

If $(e_i)_{i\in I}$ is an o.b., then every vector $u \in H$ has a series expansion

$$u = \sum_i \langle u, e_i\rangle e_i$$

which is convergent in H, and

$$\|u\|^2 = \langle u,u\rangle = \sum_i \langle\langle u,e_i\rangle e_i, e_i\rangle = \sum_i \overline{\langle u,e_i\rangle}\langle u,e_i\rangle = \sum_i |\langle u,e_i\rangle|^2,$$

which is known as Parseval's equality. Conversely, if $\sum_i |\lambda_i|^2 < +\infty, u = \sum_i \lambda_i e_i$ defines a vector $u \in H$ with coefficients λ_i.

An orthonormal basis thus establishes a linear isometry $H \longrightarrow l^2(I), u \mapsto (\langle u, e_i\rangle)_i$.

The most important example to have in mind is $H = L^2(\mathbf{T^d})$ and the orthonormal basis consisting of the complex exponentials

$$e^{2\pi ik\cdot t}, k \in \mathbf{Z^d}$$

Here $k \cdot t = k_1t_1 + k_2t_2 + \cdots + k_dt_d$. They are easily seen to be an orthonormal family; their finite linear combinations are dense in the space of continuous functions $C(\mathbf{T^d})$, by Weierstrass theorem, whence they are dense in $L^2(\mathbf{T^d})$ too.

A *frame* in H is a family of vectors $e_i, i \in I$ such that

$$m\sum_i |\langle u, e_i\rangle|^2 \leq \|u\|^2 \leq M\sum_i |\langle u, e_i\rangle|^2, u \in H,$$

for some constants m, M.

If the index set I is a continuum provided with a measure, the infinite sums are replaced with integrals,

$$m\int_I |\langle u, e_i\rangle|^2 d\mu(i) \leq \|u\|^2 \leq M\int_I |\langle u, e_i\rangle|^2 d\mu(i), u \in H.$$

When $m = M$ the frame is called *rigid*. These notions are interesting even in finite dimension. The above means that the corrrelations $\langle u, e_i\rangle$ still code u in a stable way. The mapping $u \to (\langle u, e_i\rangle)_i$ is one-to-one (meaning that the linear combinations of the e_i are dense) onto the space of coefficients, a closed subspace F on $l^2(I)$. The fact that $F \neq l^2(I)$ is due to the existence of possible linear dependencies among the e_i. Thus, a frame corresponds, in linear algebra, to a set of generating vectors which are not necessarily linearly independent. In any event, if (e_i) is a frame, the coefficients $\langle u, e_i\rangle$ determine u uniquely in a stable way. The way they do is through the *dual frame*; namely, if (e_i) is a frame there exists another frame (f_i), termed the dual frame, such that $u = \sum_i \langle u, e_i\rangle f_i$. In fact, also $u = \sum_i \langle u, f_i\rangle e_i$. Among all possible ways of writing u as a (infinite) linear combination of the e_i, this last one minimizes the l^2 norm of the coefficients.

If the frame is rigid with constant M, then $f_i = \frac{1}{M}e_i$. An example of a rigid frame in the plane is provided (in complex notation) by the three unit vectors at angles of 120 degrees between them.

If there is no redundancy and the above map is onto the whole of $l^2(I)$, that is, for every $(\lambda_i), \sum_i |\lambda_i|^2 < +\infty$, there exists $u \in H$ such that $\langle u, e_i\rangle = \lambda_i$, then the (e_i) are called a *Riesz basis*. In this case, every $u \in H$ has a unique expansion $u = \sum_i \lambda_i e_i$ with

$$m\|u\|^2 \leq \sum_i |\lambda_i|^2 \leq M\|u\|^2,$$

for some constants m, M. Riesz basis are as good as orthonormal basis in applications and easier to construct.

Chapter 2

Fourier analysis in some LCA groups

2.1 Sines and cosines, complex exponentials. Why these? The notion of character

Historically, Fourier analysis arose as a way to deal with periodic functions in the real line. A periodic function f has a group of periods, which is an additive subgroup of the real line $\mathbf{R}$, that is either discrete or dense. In the later case f is trivial, and in the former case the group of periods consists of integer multiples of a basic, fundamental period.

The simplest and elementary periodic functions with a fundamental period a are the sine and cosine functions $\sin \frac{2\pi}{a}x, \cos \frac{2\pi}{a}x$. We choose manipulating them in terms of the complex exponential $e^{i\frac{2\pi}{a}x}$ and its conjugate $e^{-i\frac{2\pi}{a}x}$.

In his pioneer work dealing with heath diffusion Fourier stated that an arbitrary a- periodic function f can be written as a superposition (infinite sum) of the elementary a- periodic functions, that is, those with fundamental period $\frac{a}{n}$:

$$f(x) = \sum_{n\in\mathbf{Z}} c_n(f) e^{i\frac{2\pi}{a}nx}.$$

We call $\frac{n}{a}$ the n-th frequency and $c_n(f)$ the n-th harmonic of f.

The theory of Fourier series deals with several aspects related to this representation that we will review in the next sections. Normalizing to $a = 1$, we deal with functions defined on the unit circle $\mathbf{T}$. To deal with general, non-periodic functions, one needs to consider all frequencies ξ and all sines and cosines, and we are led to the theory of Fourier integrals.

This the heart of Fourier analysis, expressing functions in terms of simpler elementary pieces, expressions that make easier the study of a certain aspect.

The fact that the "simpler elementary pieces" are sines and cosines, complex exponentials in modern notation, is justified historically but also by a very important fact that we emphasize here, which is their behavior under the rotation/ translation operators.

To explain this and for further reference we place ourselves in the general context of a locally compact abelian group G, that is, an abelian group (with additive notation) equipped with a topology that makes all group operations continuous. In this setting there always exists a translation invariant measure, unique aup to constants, called the Haar measure. Our two main examples to keep in mind are of course $G = \mathbf{R^d}/\mathbf{Z^d} = \mathbf{T^d}$ and $G = \mathbf{R^d}$, with Lebesgue measure, but it is also worth noting the cases $G = \mathbf{Z^d}$ or the cyclic group $G = \mathbf{Z_N}$ of the N-th roots of unity, with the counting measure. A fifth example is the multiplicative group $G = \mathbf{R}^+$, with Haar measure $\frac{dt}{t}$, but we will just consider here the first four examples.

The translation operator $\tau_x, x \in G$ acts on functions f defined on G by $(\tau_x f)(y) = f(y - x)$. It is then natural to consider *translation-invariant spaces* of functions on G, that is, spaces E such that $\tau_x f \in E$ whenever $f \in E$, and $\tau_x f$ is continuous in x. For instance, all L^p spaces are.

If E is translation-invariant, an operator $T : E \to E$ is said to *commute with translations* if $\tau_x(Tf) = T(\tau_x f), x \in G$. Note that this is a very natural assumption to do when dealing for instance with functions of time or space that describe physical phenomena and operations T among them that are time invariant, i.e. do not depend on the choice of origin in time or space. For instance, a differential operator T is time-invariant if it has constant coefficients.

We look now at functions f such that the smallest translation-invariant space containing f has dimension one. This means that for $x \in G, \tau_x f$ must be a scalar multiple of f; denoting by convenience $\chi(-x)$ this scalar factor, this means that

$$\tau_x f = \chi(-x)f, f(y - x) = \chi(-x)f(y), x, y \in G. \tag{2.1}$$

In particular, χ is continuous and so is f. Also, since $\tau_x\tau_y = \tau_{x+y}$, χ must satisfy

$$\chi(x + y) = \chi(x)\chi(y), x, y \in G. \tag{2.2}$$

Specializing (2.1) to $y = 0$ yields $f(-x) = \chi(-x)f(0)$. Therefore f is a scalar multiple of χ.

We are thus lead to the notion of a *character* of G, a continuous non-zero homomorphism $\chi : G \to \mathbf{C}$, that is, satisfying (2.2).

All functions whose translation-invariant span has dimension one are multiples of a character. Note that this implies $\chi(0) = 1$ and that if χ is bounded then $|\chi| = 1$, that is the bounded characters are the continuous homomorphisms from G to T.

The set of characters of G has a natural group structure and constitute the so-called *dual group* $\widehat{G}$.

Now, if χ is a character and T is a translation invariant operator acting on a space containing χ, we can repeat the argument above. Namely, a character χ satisfies, as function of y,

$$\tau_x \chi = \chi(-x)\chi, x \in G.$$

If T commutes with translations, it follows that

$$\tau_x(T\chi) = T(\tau_x \chi) = T(\chi(-x)\chi),$$

and since T is linear, this equals $\chi(-x)T(\chi)$. Hence

$$(T\chi)(y-x) = \chi(-x)T(\chi)(y), x, y \in G.$$

If we set $y = 0$ we find that $T\chi = \lambda\chi$ with $\lambda = T(\chi)(0)$, that is, we have proved

Theorem 4. The characters of a group are eigenvectors of all translation invariant operators.

Let us now compute what are the characters of $\mathbf{R}$. Integrating (2.2) in y over $[0, h]$ for h small, we obtain

$$\int_x^{x+h} \chi(z)dz = \int_0^h \chi(x+y)dy = \chi(x)\int_0^h \chi(y)dy.$$

Since χ is continuous and equals 1 at zero we may choose h small enough so that the last integral is non-zero. The left-hand side is differentiable because χ is continuous, whence χ is differentiable. Then, using (2.2) again we get

$$\chi'(x) = \lim_{y \to 0} \frac{\chi(x+y) - \chi(x)}{y} = \chi'(0)\chi(x).$$

Hence $\chi(x) = e^{\alpha x}$ for some $\alpha \in \mathbf{C}$. These are all the characters in $\mathbf{R}$.

From this it is immediate that the characters in $\mathbf{R^d}$ are of the same form with α multicomplex, if we interpret that $\alpha \cdot x = \alpha_1 x_1 + \cdots + \alpha_d x_d$.

Obviously, the characters in $\mathbf{T^d} = \mathbf{R^d}/\mathbf{Z^d}$ are those of the above that are 1-periodic, that is, $\alpha = 2\pi i n$ for some $n \in \mathbf{Z^d}$.

In $\mathbf{R^d}$, if we want the characters to be bounded α must be pure imaginary. For convenience we write $\alpha = 2\pi i \xi$. We are thus lead to

$$e_\xi(x) = e^{2\pi i \xi \cdot x}, \xi \in \mathbf{R^d}, \xi = (\xi_\mathbf{1}, \dots, \xi_\mathbf{d}).$$

These are all the bounded characters on $\mathbf{R^d}$, and $e_n, n \in \mathbf{Z^d}$ are all the characters in $\mathbf{T^d}$. Thus the dual group of $\mathbf{R^d}$ is identified with $\mathbf{R^d}$ and the dual group of $\mathbf{T^d}$ is identified with $\mathbf{Z^d}$.

It is worth mentioning here that the complex exponentials are linearly independent, that is, whenever we have different multifrequencies ξ_k,

$$\sum_k c_k e^{2\pi i \xi_k \cdot x} = 0,$$

then $c_k = 0$ for all k.

It is immediate to check that the characters in $\mathbf{Z}$ are of the form $n \mapsto z^n$ for some $z \in \mathbf{C}$, and so the bounded characters correspond to $|z| = 1$, that is

$$\chi_z(n) = z^n = e^{2\pi i t n}.$$

Thus the dual group of $\mathbf{Z}$ is identified with $\mathbf{T}$.

We identify the cyclic group $\mathbf{Z_N}$ with $\{0, 1, \ldots, N-1\}$ and functions there with N- periodic sequences $x = (x_n)$ indexed by $n \in \mathbf{Z}$. If $\omega_N = e^{2\pi i/N}$ denotes the primitive root of unity, it is immediate to check that the dual group is $\mathbf{Z_N}$ itself through

$$\psi_m(n) = \omega_N^{nm}, n \in \mathbf{Z}, \mathbf{m = 0, 1, \ldots, N-1}.$$

2.2 Translation invariant operators. Convolution. Impulse function.

We have seen above that the characters of a group are eigenvectors of all translation invariant operators T. We wish now to see how this helps to understand translation-invariant operators.

First let us precise in what spaces we consider these operators. In general, every locally compact abelian group has a Haar measure $d\mu$, which is the unique measure, up to constants, invariant by translations. In our four examples, Haar measure is the counting measure in the two discrete examples and the Lebesgue measure dm in the others. We consider linear operators T betwen the $L^p(G)$ spaces. Recall we said in subsection 1.7.2 that at the formal level T is given by a kernel $K(x, y)$,

$$Tf(y) = \int_G K(x, y) f(x) d\mu(x), y \in G,$$

where formally $K_x(y) = K(x, y) = T(\delta_x)(y)$. Suppose now that T commutes with translations, and set $g(y) = K_0(y) = T(\delta_0)(y)$, Then, since formally $\delta_x = \tau_x(\delta_0)$ we will have $K_x = T(\delta_x) = T(\tau_x \delta_0) = \tau_x T(\delta_0) = \tau_x g$. That is, $K(x, y) = g(y - x)$. Hence, formally, all translations invariant operators are given by

$$Tf(y) = (f * g)(y) = \int_G g(y - x) f(x) d\mu(x),$$

with a fixed g. This is called the *convolution* of f, g. The function $g = T(\delta_0)$ is known, specially in the engineering community, by the "impulse response" of T.

Definition 1. The convolution of two functions f, g is defined by

$$(f * g)(y) = \int_G g(y - x) f(x) d\mu(x),$$

whenever this makes sense. More generally, the convolution of a measure ν and g is defined by

$$(\nu * g)(y) = \int_G g(y - x) d\nu(x).$$

It is worth emphasizing that $f * g$ is an (infinite) linear combination of translates of g, namely we can write that as a whole,

$$f * g = \int_G \tau_x(g) f(x) d\mu(x).$$

Note also that $f * g = g * f$. One can check that convolution is also associative, that is $f * (g * h) = (f * g) * h$.

In case $g \in L^1(G)$ is positive and $\int_G g d\mu = 1$,, the convolution

$$(f*)g(y) = \int_G f(y - x) g(x) d\mu(x)$$

can be seen as a weighted average of f. For instance, if g is the characteristic function of a ball B (divided by $\mu(B)$) then $(f * g)(y)$ is the mean value of f in the ball $y + B$. If we think in g as being the density of a random variable X, then $f * g$ is the expected value of $f(y - X)$.

If we specialize theorem 1 to the case $K(x, y) = g(x - y)$ we obtain

Theorem 5. (Young's inequality) Suppose $1 \leq p, q, r \leq +\infty, \frac{1}{p} + \frac{1}{r} = 1 + \frac{1}{q}, f \in L^p(G), g \in L^r(G)$. Then

$$(f * g)(y) = \int_G g(y - x) f(x) d\mu(x),$$

converges absolutely for a.e y, $f * g \in L^q(G)$ and $\|f * g\|_q \leq \|f\|_p \|g\|_r$.

Note that the case $p = r = 1$ extends trivially to measures, namely the convolution $\nu * g$ of a finite measure ν and $g \in L^1(G)$ is a.e defined and $\nu * g \in L^1(G)$.

In case $G = \mathbf{R^d}$ we can state a local version of Young's inequality, in which one of the functions has compact support while the other is locally in the corresponding L^p-space.

Theorem 6. (Young's inequality) Suppose $1 \leq p,q,r \leq +\infty, \frac{1}{p}+\frac{1}{r} = 1+\frac{1}{q}, f \in L^p_{loc}(G), g \in L^r_c(G)$. Then

$$f * g(y) = \int_G g(y-x)f(x)d\mu(x),$$

converges absolutely for a.e y, and $f * g \in L^q_{loc}(G)$.

In case $G = \mathbf{R^d}, \mathbf{G} = \mathbf{T^d}$, it makes sense to look at the regularity properties of $f * g$.

Theorem 7. Suppose $1 \leq p,r \leq +\infty, \frac{1}{p}+\frac{1}{r} = 1$, and that either $f \in L^p(G), g \in L^r(G)$, $f \in L^p_{loc}(G), g \in L^r_c(G)$ or $f \in L^p_c(G), g \in L^r_{loc}(G)$. Then $f * g$ is a continuous function. Assume that f (resp. g) is differentiable at every point and that its partial derivatives $\frac{\partial f}{\partial x_i}$ (respectively $\frac{\partial g}{\partial x_i}$) satisfy the same hypothesis of f (resp. g). Then $f * g$ is differentiable and

$$\frac{\partial}{\partial x_i}(f * g) = \frac{\partial f}{\partial x_i} * g, (\text{respectively} = f * \frac{\partial g}{\partial x_i}).$$

In general, $f * g$ inherits the regularity properties of both f, g. For instance, using the notation

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}},$$

we can state that the rule

$$D^\alpha(f * g) = (D^\alpha f) * g, (\text{respectively} = f * D^\alpha g),$$

holds whenever one the the right terms makes sense.

2.3 The Fourier transform in G. Multipliers

Now, we know that formally the characters are eigenvectors of T. It is quite natural then to try to express f in terms of characters. This is analogous of what is done in elementary linear algebra; to deal with a linear operator T on $\mathbf{C^d}$ (look at vectors as functions defined on $1, 2, \dots n$) we try to diagonalize it in a basis of eigenvectors. If the characters of G constitute a basis of some sort in E then we will have that *all translation-invariant operators on E diagonalize in a basis of characters*. We emphasize the value of this fact; in finite dimensional linear algebra, a single operator (matrix) may diagonalize in a certain basis, for instance when it is symmetric. Here we find a basis that does the job not for one but for a *whole range of operators*, all translation-invariant operators.

That's why we will consider the correlations of f with the characters.

Definition 2. The Fourier transform of a function $f \in L^1(G)$ is the function $\hat{f}$ on $\widehat{G}$ defined by

$$\hat{f}(\chi) = \langle f, \chi \rangle = \int_G f\overline{\chi}d\mu.$$

The map $f \mapsto \hat{f}$ is called the Fourier transform in G.

Note that since the characters are bounded this is perfectly defined for $f \in L^1(G)$. and $\hat{f} \in L^\infty(\widehat{G})$.

Now, note that in our four examples $\widehat{G}$ also has a natural structure of group and a Haar measure $d\nu$. Our hope is that, as in finite linear algebra, the χ behave like an orthonormal basis or continuous rigid frame

$$f = \int_{\widehat{G}} \hat{f}(\chi)\chi \, d\nu(\chi) = \int_{\widehat{G}} \langle f, \chi \rangle \chi \, d\nu(\chi)$$

that is

$$f(x) = \int_{\widehat{G}} \hat{f}(\chi)\chi(x) \, d\nu(\chi), x \in G.$$

If this is so, since T commutes with infinite sums we will have

$$Tf = \int_{\widehat{G}} \hat{f}(\chi)T(\chi) \, d\nu(\chi).$$

Next we use that χ is a eigenvector of T, say $T(\chi) = m(\chi)\chi$, to get

$$Tf = \int_{\widehat{G}} \hat{f}(\chi)m(\chi)\chi \, d\nu(\chi).$$

Another formal way of writing the above is that

$$\widehat{Tf}(\chi) = m(\chi)\hat{f}(\chi).$$

This means that in "the base of characters" T diagonalizes, its infinite matrix being the diagonal one with entries $m(\chi)$.

There is a relation between the formally introduced g and the formally introduced m. This follows from the fact the Fourier transform of a convolution is the product of Fourier transforms

$$\hat{Tf}(\chi) = (\hat{f * g})(\chi) = \int_G (f * g)(y)\overline{\chi(y)}d\mu(y) =$$
$$= \int_G \int_G g(y-x)f(x)\overline{\chi(y-x)\chi(x)}d\mu(x)d\mu(y) =$$
$$= \hat{f}(\chi)\hat{g}(\chi)$$

Hence, formally $m = \hat{g}$. The function m is called the transference function or *multiplier* or *symbol* of T.

In the next sections we will see the precise form of these (formal) facts in each of four examples. If we think in the object "g" as being a function in some L^p- space then in some cases this formal argument breaks down, and g must be considered in larger spaces.

2.4 The discrete Fourier transform

For a function in $\mathbf{Z_N}$, that is, a N-periodic sequence $(x_n)_n$, the N translations are the shifted sequences $(x_{n+m})_n$. A linear operator acting on these sequences is given by a $N \times N$ matrix $A = (a_{i,j})$ that accordingly we think as a doubly infinite matrix whose rows and columns are N periodic, $a_{i+N,j} = a_{i,j}, a_{i,j+N} = a_{i,j}$. The operator is translation-invariant when the matrix A is a *circulant matrix*, meaning that

$$a_{i+1,j+1} = a_{i,j}, i, j \in \mathbf{Z}.$$

We consider $\mathbf{Z_N}$ endowed with the counting measure, the L^2-norm of a N-periodic sequence x being

$$\|x\|^2 = \sum_{n=0}^{N-1} |x_n|^2.$$

Of course this identifies $L^2(\mathbf{Z_N})$ with $\mathbf{C^N}$. It is immediate to check that the characters ψ_m are pairwise orthogonal:

$$\sum_{n=0}^{N-1} \psi_m(n)\overline{\psi_k(n)} = \sum_{n=0}^{N-1} \omega_N^{n(m-k)}.$$

If $m \neq k$ this equals

$$\frac{1 - \omega_N^{N(m-k)}}{1 - \omega_N^{m-k}} = 0,$$

and N if $m = k$. Accordingly, the normalized characters $e_m = \frac{1}{\sqrt{N}}\psi_m, m = 0, 1, \ldots, N-1$, constitute an orthonormal basis of $L^2(\mathbf{Z_N})$.

This means that the N vectors in $\mathbf{C^N}$

$$e_m = \frac{1}{\sqrt{N}}(\omega_N^{mn})_{n=0}^{N-1}, m = 0, 1, \ldots, N-1,$$

constitute an orthonormal basis of $\mathbf{C^N}$ in which all circulant matrices diagonalize.

For a finite signal $x = (x_n)_{n=0}^{N-1}$ the correlations

$$\langle x, e_m \rangle = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x_n \omega_N^{-mn}, m = 0, \ldots, N-1,$$

satisfy of course

$$x = \sum_{m=0}^{N-1} \langle x, e_m \rangle e_m = \frac{1}{N} \sum_{m=0}^{N-1} \langle x, \psi_m \rangle \psi_m,$$

that is

$$x_n = \sum_{m=0}^{N-1} \hat{x}(m) \omega_N^{mn}, \tag{2.3}$$

with

$$\hat{x}(m) = \frac{1}{N} \sum_{n=0}^{N-1} x_n \omega_N^{-mk}. \tag{2.4}$$

The equations (2.4) define the discrete Fourier transform of the N-periodic signal x and (2.3) is called the discrete Fourier inversion formula.

As it is well known, the computation of the discrete Fourier transform is implemented with the famous Cooley-Tukey algorithm called the Fast Fourier transform (FFT).

2.5 The Fourier transform in Z

Identifying z with $z = e^{2\pi it}$, here the Fourier transform of a sequence $x = (x_n) \in l^1(\mathbf{Z})$ is the function $\hat{x}$ defined on $\mathbf{T}$ by

$$\hat{x}(t) = \sum_{n \in \mathbf{Z}} x_n e^{-2\pi int}.$$

Note that the series converges uniformly on $\mathbf{T}$, and so $\hat{x}$ is continuous. Now, remember that the functions $\frac{1}{\sqrt{2\pi}} e^{2\pi int}, n \in \mathbf{Z}$ form an orthonormal basis in $L^2(T)$, and so

$$\frac{1}{2\pi} \int_0^{2\pi} \hat{x}(t) e^{2\pi imt} \, dt = x_m.$$

This is the inverse Fourier transform in $\mathbf{Z}$. If x is a finite sequence, the same computation shows that

$$\int_0^{2\pi} |\hat{x}(t)|^2 \, dt = 2\pi \sum_n |x_n|^2.$$

Therefore, the Fourier transform extends continuously to $l^2(\mathbf{Z})$,

$$\hat{x}(t) = \sum_{n \in \mathbf{Z}} x_n e^{-2\pi int},$$

where the series in the right is convergent in $L^2(\mathbf{T})$, and with the same inversion formula. The fact that the functions e_n form an orthonormal basis says that the Fourier transform is up to a constant an isometry from $l^2(\mathbf{Z})$ to $L^2(\mathbf{T})$.

Here the δ sequence is indeed in all function spaces, so there is no problem in stating that the general bounded linear translation invariant operator T from $L^p(Z)$ to $L^q(Z)$ must be a convolution operator

$$(Tx)(m) = \sum_n a(m-n)x(n), m \in \mathbf{Z},$$

for some $a = T(\delta) \in L^q(\mathbf{Z})$. If $p = 1$ there is no other restriction at all, that is, convolution with a general $a \in l^q(\mathbf{Z})$ is the general bounded translation invariant operator from $l^1(\mathbf{Z})$ to $l^q(\mathbf{Z})$; this is so by the continuous Minkowski's inequality.

However, for $p \neq 1$ we cannot recognize in terms of a and its size when convolution with a is bounded. For $p = q = 2$ however, we can describe them all, using Fourier transform and the notion of multiplier described before in a general setting.

Theorem 8. The bounded translation invariant operators in $l^2(\mathbf{Z})$ are in one-to-one correspondence with the bounded multipliers $m \in L^\infty(\mathbf{T})$, through the equation

$$\widehat{Tx}(t) = m(t)\hat{x}(t), t \in \mathbf{T}.$$

Equivalently, they are given by convolution with a sequence a given by

$$a_m = \frac{1}{2\pi}\int_0^{2\pi} m(t)e^{2\pi imt}\,dt,$$

with m bounded. In fact, $\|T\| = \|m\|_\infty$.

The important point to be noticed here is that this is a criteria for boundedness that is not a matter of size, that is,it does not depend solely on $|a|$.

2.6 Approximate identities. Regularization

For the non discrete groups $G = \mathbf{T^d}$ or $G = \mathbf{R^d}$, the delta mass is not a function but a measure, so it does not belong to any L^p space. However there is a good replacement for it. The starting point is, of course, that δ is the formal unit for convolution, $f * \delta = f$. In what follows G denotes one these groups, with additive notation and dx being the Lebesgue measure.

Definition 3. An approximate identity (or approximate kernel) is a family (k_ε) of functions in $L^1(G)$ satisfying

1. $\int_G k_\varepsilon dx = 1$.

2. $\int_G |k_\varepsilon| dx \le C$, for some constant $C > 0$.

3. For any $\delta > 0$, $\int_{|x|>\delta} |k_\varepsilon(x)| dx \to 0$ as $\varepsilon \to 0$.

It is very easy to produce examples. If $k \in L^1(G)$ is arbitrary with integral one, set $k_\varepsilon(x) = \varepsilon^{-d} k(x/\varepsilon)$. The first two conditions are obvious, while for the third one

$$\int_{|x|>\delta} |k_\varepsilon(x)| dx = \int_{|x|>\frac{\delta}{\varepsilon}} |k(x)| dx \to 0$$

because the rests of an absolutely convergent integral tend to zero by dominated convergence.

The simplest example is to take as k the normalized characteristic function of the unit ball. That is, if ω_d is the volume of the unit ball B in $\mathbf{R^d}$, consider $k(x) = \frac{1}{\omega_d}$ if $x \in B$ and zero otherwise, so that k has integral one. Then k_ε is the normalized characteristic function of the ball B_ε of radious ε and $f * k_\varepsilon(x)$ is simply the mean of f in $x + B_\varepsilon$, the ball centered at x of radious ε.

It is worth mentioning the *Poisson family* in $\mathbf{R^d}$ that corresponds to

$$k(x) = c_d \frac{1}{(|x|^2 + 1)^{\frac{d+1}{2}}}, c_d = \frac{\Gamma(\frac{d+1}{2})}{\pi^{\frac{d+1}{2}}},$$

and the Gaussian family given by

$$k(x) = \frac{1}{(\sqrt{2\pi})^d} e^{-\frac{1}{2}|x|^2}.$$

On the torus $\mathbf{T^d}$ we will see later that natural examples appear when dealing with convergence of the Fourier series, namely the Fejer kernel. We may consider as well approximations of the identity indexed by $n \in \mathbf{N}$ with obvious modifications.

Theorem 9.
- If (k_ε) is an approximation of the identity and $f \in L^p(G), 1 \le p < +\infty$, then $f * k_\varepsilon \to f$ in $L^p(G)$ as $\varepsilon \to 0$.
- If $f \in C_0(G)$, then $f * k_\varepsilon \to f$ uniformly on G.
- If $f \in L^1(G)$ and f is continuous at a point x_0, then $(f * k_\varepsilon)(x_o) \to f(x_0)$.

Proof. We can write

$$f - f * k_\varepsilon = f - \int_G \tau_x(f) k_\varepsilon(x) d\mu(x) = \int_G (f - \tau_x(f)) k_\varepsilon(x) d\mu(x),$$

and hence, by the continuous Minkowski inequality

$$\|f - f * k_\varepsilon\|_p \leq \int_G \|f - \tau_x(f))\|_p |k_\varepsilon(x)| d\mu(x).$$

To estimate it we break the above in two parts, corresponding to small x, say $\|x\| \leq \delta$, and $\|x\| > \delta$. The first one is estimated by

$$C \sup_{\|x\| \leq \delta} \|f - \tau_x(f)\|_p,$$

and hence can be made arbitrarily small if δ is small enough, uniformly in ε, due to the continuity of translations in $L^p(G)$, while the second is estimated by

$$2\|f\|_p \int_{|x|>\delta} |k_\varepsilon(x)| dx.$$

□

A consequence of this is

Theorem 10. A bounded operator T from $L^p(\mathbf{R^d})$ to $L^q(\mathbf{R^d})$, $1 \leq p, q < +\infty$ commutes with translations if and only if commutes with convolution with L^1 functions, that is,

$$T(f * g) = f * Tg, f \in L^1(\mathbf{R^d}), \mathbf{g} \in \mathbf{L^p}(\mathbf{R^d}).$$

Proof. By Minkowski's continuous inequality, the right hand side of

$$f * g = \int_G (\tau_x g) f(x) d\mu(x),$$

is convergent in L^p, hence if T commutes with translations,

$$T(f * g) = \int_G f(x) T(\tau_x g)\, d\mu(x) = \int_G f(x) \tau_x T g\, d\mu(x) = f * Tg.$$

If T commutes with convolutions, we consider an approximation of the identity k_ε so that

$$T(\tau_x g) = \lim_\varepsilon T((\tau_x g) * k_\varepsilon) = \lim_\varepsilon T(g * (\tau_x k_\varepsilon))$$
$$= \lim_\varepsilon (Tg) * (\tau_x k_\varepsilon)) = \tau_x (\lim_\varepsilon (Tg) * k_\varepsilon) = \tau_x(Tg).$$

□

The above theorem has two important consequences.

Theorem 11. The space of infinitely differentiable functions $C^\infty(\mathbf{T^d})$ is dense in all $L^p(\mathbf{T^d})$ spaces, $1 \leq p < \infty$. The space $C_c^\infty(\mathbf{R^d})$ of infinitely differentiable functions with compact support is dense in all $L^p(\mathbf{R^d})$ spaces, $1 \leq p < +\infty$.

Proof. . The space of continuous functions with compact support is dense. If f is in this space, and we take an approximation of the identity $k_\varepsilon(x) = \varepsilon^{-d}k(x/\varepsilon)$, with k a C^∞ function with compact support, then $k_\varepsilon * f \in C_c^\infty(\mathbf{R^d})$ and tends to f in L^p.

□

The same proof shows that for an open set $U \subset \mathbf{R^d}$ the space $C_c^\infty(U)$ is dense in all $L^p(U)$ spaces as well. We point out the following easy consequence

Theorem 12. If $f \in L^1_{loc}(U)$ and

$$\int_U f(x)\varphi(x)\,dx = 0$$

for all $\varphi \in C_c^\infty(U)$, then $f = 0$ a.e. The same is true if

$$\int_B f(x)dx = 0$$

for all balls $B \subset U$.

.

A remark is in order here. For most of the approximations of the identity of type above, for $f \in L^1_{loc}(U)$, not only the means $f * k_\varepsilon \to f$ in $L^1_{loc}(U)$, but in fact we will see later that $f * k_\varepsilon \to f$ pointwise a.e. (Lebesgue tjeorem)

Theorem 13. The general form of a bounded translation-invariant operator in $L^1(\mathbf{T^d})$ or $L^1(\mathbf{R^d})$ is convolution with a finite complex Borel measure $d\mu$.

Proof. Obviously convolution with a complex finite Borel measure is bounded and translation invariant. Conversely, given such T, the idea is of course that $d\mu$ should be $T(\delta_0)$, and we simply replace δ_0 by an approximation of the identity k_ε. Since they are bounded in L^1, $T(k_\varepsilon)$ will be also bounded in L^1. By the Banach-Alaoglu theorem there exists a finite complex valued measure $d\mu$ and a sequence $\varepsilon_n \to 0$ such that

$$\lim_n \int g(y)T(k_{\varepsilon_n})(y)dy = \int g(y)d\mu(y), g \in C_c.$$

Now, since $g = \lim_n g * k_{\varepsilon_n}$ and T is bounded and commutes with convolutions, one has $Tg = \lim_n g * T(k_{\varepsilon_n})$. But

$$(g * Tk_{\varepsilon_n})(x) = \int g(x-y)T(k_{\varepsilon_n})(y)dy = \int g(x-y)d\mu(y) = (g * \mu)(x).$$

Hence T is convolution with μ on all functions with compact support and hence on all functions.

□

Chapter 3

The Fourier analysis of periodic functions

We will deal mainly with dimension $d = 1$ in the beginning.

3.1 The Fourier series of periodic functions

Scaling we may assume that the period is 1 and we deal with functions on $\mathbf{T}$, parametrized by $|t| \leq \frac{1}{2}$ through $e^{2\pi it}$. Here translation invariance becomes rotation invariance and convolutions are circular or periodic convolutions

$$(f * g)(t) = \int_{|t| \leq \frac{1}{2}} f(t-x)g(x)dx.$$

Motivated by the considerations above we are to consider the characters $e_n(x) = e^{i2\pi nt}, n \in \mathbf{Z}$, that is, the sines and cosines of period 2π as elementary building blocks. Recall that the e_n constitute an orthogonal basis of $L^2(\mathbf{T})$, the so-called *Fourier basis.* The expression $\sum_n \langle f, e_n \rangle e_n$ is usually written

$$\sum_n c_n(f) e^{2\pi int},$$

with

$$c_n(f) = \int_0^1 f(t) e^{-2\pi int} \, dt.$$

This makes sense for $f \in L^1(\mathbf{T})$ (and for all $\in L^p(\mathbf{T}) \subset \mathbf{L^1(T), p \geq 1}$) , $c_n(f)$ is called the n-th Fourier coefficient of f and the formal series

$$S(f) = \sum_{n \in \mathbf{Z}} c_n(f) e^{2\pi int},$$

is called the *Fourier series* of f.

The fact that the e_n constitute an orthonormal basis of $L^2(\mathbf{T})$ can be restated by saying that the map $f \to (c_n(f))_n$ is a bijection from $L^2(\mathbf{T})$ to $l^2(\mathbf{Z})$ satisfying the so-called *Plancherel's identity*

$$\sum_n |c_n(f)|^2 = \int_0^1 |f(t)|^2\, dt,$$

and its polarized version *Parseval's relation*

$$\sum_n c_n(f)\overline{c_n(g)} = \int_0^1 f(t)\overline{g(t)}\, dt.$$

3.2 Properties of Fourier coefficients

Proposition 1. The following properties hold:

- $c_n(f * g) = c_n(f)c_n(g)$.
- $c_n(\tau_x f) = e^{-2\pi inx}c_n(f)$.
- (The Riemann-Lebesgue lemma). $|c_n(f)| \leq \|f\|_1$ and $c_n(f) \to 0$ as $|n| \to \infty$.
- If f is of class C^k and 2π- periodic, then $c_n(f^{(k)}) = (2\pi in)^k c_n(f)$ and $c_n(f) = o(|n|^{-k})$.

Proof. The first two are immediate consequences of the definition (they can be guessed by formal manipulation of $S(f)$ too), as well as the boundedness of the coefficients. From the first it follows that $c_n(f - \tau_x f) = (1 - e^{-2\pi inx})c_n(f)$; choosing $x = \frac{1}{2n}$ we get $2|c_n(f)| \leq \|f - \tau_{2/n} f\|_1$, so the result is a consequence of the continuity of translations. The last property is immediate too. □

It can be seen that nothing more general can be said regarding the speed of convergence of the coefficients $c_n(f)$, meaning that given any preestablished decay one can construct $f \in L^1(\mathbf{T})$ with slower decay.

3.3 The Dirichlet, Fejer and Poisson kernels in T

To study $S(f)$ it is natural to consider the partial sums

$$S_N(f) = \sum_{-N}^{N} c_n(f)e^{2\pi int}.$$

By direct computation,

$$S_N(f)(t) = \int_0^1 f(x) \sum_{-N}^{N} e^{2\pi i n(t-x)}\, dx = (f * D_N)(t),$$

with

$$D_N(t) = \sum_{n=-N}^{N} e^{2\pi i n t} = \frac{\sin(2N+1)\pi t}{\sin \pi t}.$$

The family (D_N) is called the *Dirichlet* kernel. If it were an approximation of the identity we would have that $S_N f \to f$ for $f \in L^p(T)$, but it is easy to see that $\|D_N\|$ behaves like $\log N$.

It is a general fact in analysis that if a sequence a_N has a bad behaviour one should look at the sequence of averages $\frac{1}{N+1}(a_0 + a_1 + \cdots + a_{N+1})$ as it generally exhibits a better behaviour. That's why we consider

$$\sigma_N(f) = \frac{1}{N+1}(S_0(f) + \cdots + S_N(f)) = (f * \sigma_N),$$

where

$$\sigma_N(t) = \frac{1}{N+1}(D_0(t) + \cdots + D_N(t)) = \sum_{j=-N}^{N} \left(1 - \frac{|j|}{N+1}\right) e^{2\pi i n t} = $$

$$= \frac{1}{N+1} \left(\frac{\sin(N+1)\pi t}{\sin \pi t}\right)^2.$$

Thus σ_N is positive, has integral one and if $|t| > \delta$ then $|\sigma_N(t)| \le c_\delta \frac{1}{N}$ and so it is an approximation of the identity. Hence we have

Proposition 2.
- If $f \in L^p(\mathbf{T})$ then $\sigma_N(f) \to f$ in $L^p(\mathbf{T})$ as $N \to \infty$, $1 \le p < +\infty$.
- If f is continuous at t_0, then $\sigma_N(f)(t_0) \to f(t_0)$. If f is continuous then $\sigma_N(f) \to f$ uniformly.
- (Uniqueness theorem) If $c_n(f) = 0$ for all n then $f = 0$. Hence, if $c_n(f) = c_n(g)$ for all n, then $f = g$.
- If f is continuous at t_0 and $S(f)(t_0)$ converges, the sum must be $f(t_0)$

Note that item b) above provides a constructive proof of the Weierstrass approximation theorem on $\mathbf{T}$. On the other hand by exploiting the symmetry of the Fejer kernel one can show that in case f has lateral limits at t_0 then $\sigma_N(f)(t_0)$ converges to their mean value and so this is the only possible sum of $S(f)(t_0)$.

The *Poisson kernel* arises when dealing with the Dirichlet problem in the unit disc $\mathbf{D}$, that is, given a continuous function $f \in C(\mathbf{T})$ to find an harmonic function u on $\mathbf{D}$, continuous in $\overline{\mathbf{D}}$ such that $u = f$ on T. The later is of course equivalent to the statement that $u(re^{2\pi it}) \to f(t)$ uniformly. We solve this problem exploiting its invariance by rotations. To be precise, if for each $0 < r < 1$ we pose $u_r(t) = u(re^{2\pi it})$, the operator $f \to u_r$ is rotation invariant, hence it must be given by a circular convolution with some P_r,

$$u_r = f * P_r,$$

and must diagonalize in the Fourier basis, i.e.

$$c_n(u_r) = m_n c_n(f), m_n = c_n(P_r).$$

Since the solution of the Dirichlet problem for $f(t) = e^{2\pi int}$ is $u(re^{2\pi it}) = z^n$ if n is positive and $\overline{z}^n$ if n is negative, $z = re^{2\pi it}$, we find $m_n = r^{|n|}$. Since

$$P_r(t) = \sum_n r^{|n|} e^{2\pi int} = \frac{1 - r^2}{|1 - re^{2\pi it}|^2} = \frac{1 - r^2}{1 + r^2 - 2r\cos(2\pi t},$$

we guess that the solution should be

$$u(re^{2\pi it}) = (f * P_r)(t) = \int_0^1 f(x) \frac{1 - r^2}{1 + r^2 - 2r\cos(2\pi(t - x))} dx. \tag{3.1}$$

One can reach the same conclusion by separating variables, that is, writing the Fourier series of u_r

$$u_r(t) = \sum_n c_n(r) e^{2\pi int},$$

and imposing that u satisfies all properties. The kernel P_r is called the Poisson kernel. Note that it is positive with integral one, while if $|t| > \delta$, then $|1 - re^{2\pi it}|$ is bounded below uniformly in r, whence

$$P_r(t) \leq c_\delta(1 - r^2).$$

This shows that (P_r) is also an approximation of the identity as $r \to 1$. With this one can easily prove that indeed the above is the solution to Dirichlet problem:

Theorem 14. For $f \in L^1(\mathbf{T})$, the function u defined on the unit disc by (3.1) above is an harmonic function in D satisfying $u_r \to f$ in $L^1(\mathbf{T})$. In case $f \in C(T)$, u is continuous in the closed disc with boundary values equal to f and is thus the solution of Dirichlet's problem.

3.4 Pointwise convergence

The pointwise convergence of the Fourier series of f is a very natural question. When using sumability "a la Fejer" o " a la Poisson" the situation is quite good. Indeed, as both the Fejer and Poisson kernels are approximate identities one can prove that for $f \in L^1(\mathbf{T})$ both $F_N(f)(t)$ and $u_r(t)$ have limit $f(t)$ a.e. We will see this later as an application of the maximal function of Hardy-Littlewood.

The pointwise convergence of $S(f)$ in the classical sense is a much subtle question. To begin with, Kolmogorov constructed an $f \in L^1(\mathbf{T})$ such that $S(f)$ diverges a.e.

The poinwise convergence of $S(f)$ for $f \in L^p(\mathbf{T}), \mathbf{1} < \mathbf{p}$, was a very hard open problem in Fourier analysis till Carleson proved that $S(f)(t)$ converges to $f(t)$ for a.e. t for $f \in L^2(T)$ in a celebrated breakthrough, and this was generalized to $L^p(\mathbf{T}), \mathbf{1} < \mathbf{p} < +\infty$ by Hunt.

To state sufficient conditions at a given point it is convenient to rewrite

$$S_N(f)(t) - f(t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} (f(t-x) - f(t)) \frac{\sin(2N+1)\pi x}{\sin \pi x}\, dx =$$

$$= \int_{-\frac{1}{2}}^{\frac{1}{2}} A(t,x) \sin 2\pi N x\, dx + \int_{-\frac{1}{2}}^{\frac{1}{2}} f(t-x) \cos 2\pi N x\, dx,$$

with

$$A(t,x) = \frac{f(t-x) - f(t)}{\tan \pi x}.$$

Of the two terms above, the second one has limit zero as $N \to +\infty$ by the Riemann-Lebesgue lemma. As to the first one, the contribution of $|x| > \delta$ also tends to zero by the same reason because $A(t,x)$ is integrable in $|x| > \delta$. this shows that the statement $S_N(f)(t) \to f(t)$ is equivalent to

$$\int_{|x|<\delta} \frac{f(t-x) - f(t)}{\tan \pi x} \sin 2\pi N x\, dx \to 0,$$

for δ small. This implies that the convergence of $S(f)(t)$ to $f(t)$ is of local nature, i.e. it only depends on the local behaviour of f at t (Riemann's localization principle). For instance, the above shows that if $\frac{f(t-x)-f(t)}{x}$ is integrable, then $S(f)(t)$ converges to $f(t)$ (Dini's criterion). In particular this holds true if f satisfies a Lipschitz condition at t or if f is differentiable at t.

3.5 Convergence in norm

Regarding convergence in $L^p(\mathbf{T}), \mathbf{1} \leq \mathbf{p} < +\infty$, we have seen that $\sigma_N(f) \to f$ in $L^p(\mathbf{T})$ while we trivially know that $S_N(f) \to f$ in $L^2(\mathbf{T})$ if $f \in L^2(\mathbf{T})$.

We will see later, as an application of the CZ theory, that this holds true for $1 < p < +\infty$.

Another natural question is studying uniform convergence of $S(f)$ when f is continuous. One needs extra assumptions, for du Bois-Reymond constructed a continuous f such that $S(f)$ diverges at some point (in fact examples can be constructed where $S(f)$ diverges on a dense set).

Using the arguments above one can prove that the Fourier series of an absolutely continuous function or a Lipschitz function converges uniformly. More generally, if f is continuous and of bounded variation then $S(f)$ converges to f uniformly.

3.6 Rotation invariant operators in $L^p(\mathbf{T})$

The Fourier series helps understanding the structure of rotation invariant operators, as shown by the following result.

Theorem 15. For a bounded operator $T : L^p(\mathbf{T}) \to \mathbf{L^q(T), 1 \leq p, q < \infty}$ the following are equivalent:

- It commutes with rotations.
- It commutes with convolution with $L^1(\mathbf{T})$ functions.
- The characters e_n are eigenvectors of T, $Te_n = m_n e_n$.
- It diagonalizes in the Fourier basis: $c_n(Tf) = m_n c_n(f), n \in \mathbf{Z}$.

The same holds replacing $L^p(\mathbf{T}), \mathbf{L^q(T)}$ by $C(\mathbf{T})$ in case $p, q = +\infty$. Moreover

- In case $p = q = 1$, the general form of T is given by $Tf = f * \mu$, with a finite complex Borel measure μ, in which case $m_n = c_n(\mu)$, and $\|T\|$ equals the total variation of μ.
- The general form of a bounded translation-invariant operator $T : C(\mathbf{T}) \to \mathbf{C(T)}$ is also $Tf = f * \mu$, with a finite complex Borel measure μ.
- In case $p = q = 2$, the general form of T is given by $c_n(Tf) = m_n c_n(f)$ with m_n an arbitrary bounded sequence, in which case the norm of T as an operator in $L^2(\mathbf{T})$ equals $\sup_n |m_n|$.

.

Proof. We already know that a) implies b) implies c). To see that this implies d) we consider $f \in L^p(\mathbf{T})$. Then

$$c_n(Tf) = \lim_N c_n(\sigma_N Tf) = \lim_N c_n(T(\sigma_N f)) = m_n c_n(f).$$

If d) holds then $\tau_x Tf$ and $T(\tau_x f)$ have the same Fourier coefficients and so they are equal.

In case $p = q = 1$ we know already the statement after theorem ??. In case $p = q = 2$ the result is an easy consequence of the fact that the e_n constitute a basis of $L^2(T)$.

In the case of $C(\mathbf{T})$ the statement is almost a tautology; indeed, if T is bounded and translation invariant in $C(T)$, then $f \to Tf(0)$ is a continuous bounded functional, hence there exists a measure μ such that

$$Tf(0) = \int_{\mathbf{T}} f(x) d\mu(x).$$

If $d\nu(x) = d\mu(-x)$, then

$$T(f)(y) = \tau_{-y}(Tf)(0) = T(\tau_{-y}f)(0) = \int_{\mathbf{T}} (\tau_{-y}f)(x) d\mu(x) =$$

$$\int_{\mathbf{T}} (\tau_{-y}f)(-x) d\nu(x) = \int_{\mathbf{T}} f(y-x) d\nu(x).$$

□

Note that in the above result T is formally the convolution of f with

$$g(t) = \sum_n m_m e^{2\pi i n t}.$$

But this object g is not in general an $L^1(\mathbf{T})$ function (for instance if m_n does not tend to zero) not even a measure. That object g is in general a *distribution*. In $L^1(\mathbf{T})$ the situation is somewhat the opposite: we know precisely the object g, a measure, but we do not know exactly which multipliers m_n may arise. In the other cases we do not know the exact description of neither g nor the m_n.

3.7 The Fourier transform in $\mathbf{T^d}$

The Fourier series of a function on $\mathbf{T^d}$ is

$$Sf(t) = \sum_{k \in \mathbf{Z^d}} c_k(f) e^{2\pi i k \cdot t}, k \cdot t = k_1 t_1 + \cdots + k_d t_d,$$

with

$$c_k(f) = \int_0^1 \dots \int_0^1 f(t) e^{-2\pi i k \cdot t} \, dt_1 \dots dt_d.$$

Much of the analysis done in the previous section goes over to $N > 1$, provided that appropriate definitions are given, namely that of $S_N f(t)$. If rectangulars sums are used, that is,

$$S_N^r(f)(t) = \sum_{|k_i| \le N} c_k(f) e^{2\pi i k \cdot t},$$

and correspondingly for $\sigma_N(f)$, then the results for $S_N(f)$ and $\sigma_N(f)$ hold as well. However, if spherical sums are considered

$$S_N^e(f)(t) = \sum_{|k|\leq N} c_k(f)e^{2\pi i k\cdot t}, |k|^2 = k_1^2 + \cdots + k_d^2,$$

then the situation becomes more complicated and will not be explained here.

Finally we remark another aspect to be taken into consideration for later reference. In one variable a non trivial periodic function has a fundamental period $a > 0$ and by scaling the Fourier development reads

$$f(t) = \sum_n c_n(f)e^{\frac{2\pi}{a}int},$$

with

$$c_n(f) = \frac{1}{a}\int_0^a f(t)e^{-\frac{2\pi}{a}int}\,dt.$$

The frequencies are then the integer multiples of $\frac{2\pi}{a}$.

For a full non trivial periodic function g in $\mathbf{R^d}, \mathbf{d} > \mathbf{1}$ its group of periods is a lattice of the form $\Lambda = A(\mathbf{Z^d})$ for some $n \times n$ invertible matrix A, with fundamental region $I = A([0,1]^n)$. If $f(t) = g(At)$, f is $\mathbf{Z^d}$ periodic and has a Fourier series expansion as above. Rewriting it in terms of g one obtains, with $\Lambda^* = (A^*)^{-1}(\mathbf{Z^d})$ being the dual lattice

$$g(t) = \sum_{\rho\in\Lambda^*} c_\rho(g)e^{2\pi i\rho\cdot t},$$

where

$$c_\rho = \frac{1}{|detA|}\int_I g(t)e^{-2\pi i\rho\cdot t}dt.$$

The frequencies are then located at Λ^*.

Chapter 4

The Fourier transform in $\mathbf{R^d}$

4.1 The Fourier transform in $L^1(\mathbf{R^d})$

For $f \in L^1(\mathbf{R^d})$ we define its Fourier transform $\hat{f}$ by

$$\hat{f}(\xi) = \langle f, e_\xi \rangle = \int_{\mathbf{R^d}} f(x) e^{-2\pi i \xi \cdot x}\, dx, \xi \in \mathbf{R^d}.$$

More generally we can define the Fourier transform of a finite complex Borel measure $d\mu$ by

$$\hat{\mu}(\xi) = \int_{\mathbf{R^d}} e^{-2\pi i \xi \cdot x}\, dx.$$

Besides the translation operators $\tau_x, x \in \mathbf{R^d}$, we consider too the dilation operators $\rho_\lambda f(x) = f(\lambda x), \lambda > 0$. The following properties are elementary. They establish that translations and multiplication by characters correspond one each other under Fourier transform, dilation goes to inverse dilation, and differentiation corresponds to multiplication by polynomials, up to constants.

Proposition 3. The following properties hold:

- $\widehat{\tau_x f}(\xi) = e^{-2\pi i \xi \cdot x} \hat{f}(\xi)$
- If $g(x) = e^{2\pi i \eta \cdot x} f(x)$, then $\hat{g}(\xi) = \tau_\eta(\xi)$.
- $\widehat{\rho_\lambda f}(\xi) = \lambda^{-d} \hat{f}(\frac{\xi}{\lambda})$.
- $\widehat{D^\alpha f}(\xi) = (2\pi i \xi)^\alpha \hat{f}(\xi)$.
- $(D^\alpha \hat{f})(\xi) = ((-2\pi i x)^\alpha f(x))\widehat{\ }(\xi)$.
- $\widehat{f * g} = \hat{f}\hat{g}$.

- If A is an invertible matrix and $f_A(x) = f(Ax)$, then

$$\widehat{f_A}(\xi) = \frac{1}{|detA|}\hat{f}(A^{-1})^*\xi).$$

- (The Riemann-Lebesgue lemma). $\hat{f}$ is a continuous function vanishing at ∞.

Also note that if A is an orthogonal matrix, then the Fourier transform commutes with composition by A. It follows that if a function $f(x)$ is radial(that is, it only depends on $|x|$, or if $f(Ax) = f(x)$ for all orthogonal A), then $\hat{f}$ is also radial. As a consequence, convolution of radial functions is radial.

In particular, if $P(D) = \sum_{\alpha\in\mathbf{N^d}} c_\alpha D^\alpha$ is a differential operator with constant coefficients (and so translation invariant) one has

$$\widehat{P(D)f}(\xi) = P(2\pi i\xi)\hat{f}(\xi), (P(D)\hat{f})(\xi) = (P(-2\pi ix)f\widehat{)}(\xi),$$

whenever the left hand side makes sense.

A translation- invariant operator T has a multiplier, for instance that of $P(D)$ is $m(\xi) = P(2\pi i\xi)$. We say that T is invariant by rigid motions if moreover $T(f_A) = (Tf)_A$ for all orthogonal matrices. Then, its multiplier must be radial. For instance, the Laplace operator

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_d^2},$$

is radial and has multiplier $m(\xi) = -4\pi^2|\xi|^2$. If a differential operator $P(D)$ is invariant by rigid motions, then its multiplier is a radial polynomial, that is a polynomial in $|\xi|^2$, and hence we have

Proposition 4. A differential operator $P(D)$ is invariant by rigid motions if and only if it is a polynomial in Δ.

4.2 The Dirichlet, Fejer and Poisson kernels in $\mathbf{R^d}$

Again, by the motivation explained in a general context, we ask ourselves about the validity of

$$f = \int_{\mathbf{R^d}} \langle f, e_\xi\rangle e_\xi d\xi,$$

that is

$$f(x) = \int_{\mathbf{R^d}} \hat{f}(\xi)e^{2\pi ix\cdot\xi}\, d\xi, \tag{4.1}$$

in some sense.

The first and most natural method is to consider the Dirichlet means. Here we have two choices, we can use cubes of size R or else balls of size R, that is,

$$(S^c_R f)(x) = \int_{|\xi_i|\le R} \hat{f}(\xi)e^{2\pi i x\cdot\xi}\, d\xi$$
$$(S^b_R f)(x) = \int_{|\xi|\le R} \hat{f}(\xi)e^{2\pi i x\cdot\xi}\, d\xi$$

It turns out that these means have a different behaviour. Here we will use the rectangular means, because as we shall see their kernels are tensor products of one variable kernels. By direct computation we find that

$$(S^c_R f)(x) = (f * D_R)(x), D_R(x) = \prod_{j=1}^{d} \frac{\sin 2\pi R x_j}{\pi x_j}.$$

The kernel D_R is strictly speaking not integrable (a typical example of a conditionally convergent integral) and is not an approximation of the identity. However, if we consider their means as we did in the periodic case, we find

$$(\sigma_R f)(x) = \frac{1}{R}\int_0^R (S^c_r f)(x)\, dr = (f * F_R)(x),$$

where $F_R(x) = R^d F(Rx)$, with

$$F(x) = \prod_{i=1}^{d} \frac{1-\cos 2\pi x_j}{2\pi^2 x_j^2}.$$

One can check that F has integral one and hence F_R is an approximation of the identity.

Another way, more useful, to deal with (4.1) is to introduce other types of means. The general scheme is as follows: we consider a continuous integrable function Φ such that $\Phi(0) = 1$ and the mean

$$\int_{\mathbf{R^d}} \hat{f}(\xi)e^{2\pi i x\cdot\xi}\Phi(\varepsilon\xi)d\xi.$$

Now, Fubini's theorem implies that

$$\int_{\mathbf{R^d}} \hat{f}(\xi)g(\xi)d\xi = \int_{\mathbf{R^d}} f(y)\hat{g}(y)dy, f, g \in L^1(\mathbf{R^d}).$$

The Fourier transform of $\Phi(\varepsilon\xi)$ is $\varepsilon^{-d}\widehat{\Phi}(\frac{y}{\varepsilon}) = \hat{\Phi}_\varepsilon(y)$, whence the Fourier transform of $e^{2\pi i x\cdot\xi}\Phi(\varepsilon\xi)$ is $\hat{\Phi}_\varepsilon(y-x)$ and so

$$\int_{\mathbf{R^d}} \hat{f}(\xi)e^{2\pi i x\cdot\xi}\Phi(\varepsilon\xi)d\xi = (f * \widehat{\Phi}_\varepsilon)(x).$$

The most simple choice is $\Phi(x) = e^{-\pi|x|^2}$ (Gauss means), for which it is easy to compute that $\widehat{\Phi}(\xi) = e^{-\pi|\xi|^2}$, that is $\widehat{\Phi} = \Phi$. For this choice we then have

$$\int_{\mathbf{R^d}} \hat{f}(\xi) e^{-\varepsilon\pi|\xi|^2} e^{2\pi i x\cdot\xi}\, d\xi = (f * \Phi_\varepsilon)(x). \tag{4.2}$$

But $\int_{\mathbf{R^d}} \Phi(x)dx = \widehat{\Phi}(0) = \Phi(0) = 1$, therefore Φ_ε is an approximation of the identity and therefore $G_\varepsilon F \to f$ in $L^1(\mathbf{R^d})$. This implies the unicity theorem: if $\widehat{f} = 0$, then $f = 0$. It also implies

Theorem 16. (Inversion theorem) If $f \in L^1(\mathbf{R^d})$ and $f \in L^1(\mathbf{R^d})$ then

$$f(x) = \int_{\mathbf{R^d}} \hat{f}(\xi) e^{2\pi i x\cdot\xi}\, d\xi, a.e.x$$

and in particular f is a.e. equal to a continuous function vanishing at infinity.

The Gauss means (4.2) are connected with the heath diffusion problem: indeed one can check that $u(t,x) = f_\phi(\sqrt{t}, x)$ is the solution of the heath equation

$$\frac{\partial u}{\partial t} = \frac{1}{4}\Delta_x u(x,t), u(0,x) = f(x).$$

More generally, we can choose any continuous function Φ such that $\Phi(0) = 1$ and $\widehat{\Phi}$ is integrable. Then by the above the integral of $\widehat{\Phi}$ equals $\Phi(0) = 1$ and we can repeat the same argument. For instance, another choice is $\Phi(x) = e^{-2\pi|x|}$ leading to the Abel means. One can check that in this case

$$\widehat{\Phi}(\xi) = c_d \frac{1}{(1+|\xi|^2)^{(d+1)/2}}, c_d = \frac{\Gamma[\frac{d+1}{2}]}{\pi^{(d+1)/2}}.$$

and that in this case $u(t,x) = f_\phi(t,x)$ satisfies

$$\frac{\partial^2 u}{\partial t^2} + \Delta_x u(t,x) = 0, u(0,x) = f(x),$$

that is, is the solution of the Dirichlet problem in the half-space.

We saw that the Fourier transform converts convolution to products. We can now show that it also converts products to convolutions, when everything makes sense:

Proposition 5. If $f, g, \hat{f}, \hat{g} \in L^1(\mathbf{R^d})$, then $\widehat{fg} = \hat{f} * \hat{g}$.

Proof. By the inversion formula, both f, g are continuous functions vanishing at infinity, so fg is integrable. Checking that both terms have the same transform we are done. □

4.3 The Fourier transform in $L^2(\mathbf{R^d})$

Let us observe that the pointwise definition of $\hat{f}(\xi)$ does not make sense for $f \in L^2(\mathbf{R^d})$. We will see, however that it can be still be defined in a suitable sense, in L^2 sense. The basic fact is the following theorem:

Theorem 17. If $f \in L^1(\mathbf{R^d}) \cap \mathbf{L^2}(\mathbf{R^d})$ then $\hat{f} \in L^2(\mathbf{R^d})$ and $\|f\|_2 = \|\hat{f}\|_2$.

Proof. We observe first that (4.2) also implies that if $h \in L^1(\mathbf{R^d})$, $\hat{h} \geq 0$ and h is continuous at zero then $\hat{h} \in L^1(\mathbf{R^d})$, the inversion formula holds and $h(0) = \int_{\mathbf{R^d}} \hat{h} d\xi$. Given $f \in L^1(\mathbf{R^d}) \cap \mathbf{L^2}(\mathbf{R^d})$, let $g(x) = \overline{f(-x)}$ so that $\hat{g} = \overline{\hat{f}}$, and consider $h = f * g$. Then h is in $L^1(\mathbf{R^d})$, $\mathbf{\hat{h}} = |\mathbf{\hat{f}!^2}$; Since both f, g are in $L^2(\mathbf{R^d})$, h is continuous (by the continuity of translations). Hence

$$\int_{\mathbf{R^d}} |\hat{f}(\xi)|^2 \, d\xi = \int_{\mathbf{R^d}} \hat{h} = h(0) = \int_{Rd} |f(x)|^2 \, dx.$$

□

At this point we introduce the Schwarz class $\mathcal{S}(\mathbf{R^d})$ of C^∞ functions f such that all derivatives $D^\alpha f$ decay faster at infinity than all polynomials:

$$\lim_{|x| \to +\infty} |x^\beta D^\alpha f(x)| = 0, \alpha, \beta \in \mathbf{N^d}.$$

Obviously this space is dense in all L^p spaces, $1 \leq p < +\infty$ because it contains the space $C_c^\infty(\mathbf{R^d})$. Using the properties of the Fourier transform and the inversion theorem, it is immediate to see that the Fourier transform is a bijection from $\mathcal{S}(\mathbf{R^d})$ to itself. Now, trivially $L^1(\mathbf{R^d}) \cap \mathbf{L^2}(\mathbf{R^d})$ contains $\mathcal{S}(\mathbf{R^d})$ and so does its image under the Fourier transform. Thus we have that the Fourier transform is an isometry between a dense subspace of $L^2(\mathbf{R^d})$ an another dense subspace. We can then define the Fourier transform of an arbitrary $f \in L^2(\mathbf{R^d})$ taking an approximation in L^2 by functions in $L^1(\mathbf{R^d}) \cap \mathbf{L^2}(\mathbf{R^d})$, for instance $f(x)$ times the indicator function of a ball of radius R. Then

$$\hat{f}(\xi) = \lim_{R \to +\infty} \int_{|x| \leq R} f(x) e^{-2\pi i \xi \cdot x} \, dx,$$

exists in $L^2(\mathbf{R^d})$, defines $\hat{f}$, and

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

(Plancherel's identity). It is an isometry onto $L^2(\mathbf{R^d})$, the inverse being

$$f(x) = \lim_{R \to +\infty} \int_{|\xi| \leq R} \hat{f}(\xi) e^{2\pi i \xi \cdot x} \, d\xi,$$

also convergent in $L^2(\mathbf{R^d})$.

Being an isometry is equivalent to preserving inner products, that is, to the polarized version

$$\int_{\mathbf{R^d}} f(x)\overline{g(x)}\,dx = \int_{\mathbf{R^d}} \hat{f}(\xi)\overline{\hat{g}(\xi)}\,d\xi,$$

which is known as Parseval's relation.

With this definition, the properties of the Fourier transform listed above hold true as well in $L^2(\mathbf{R^d})$. Generally speaking, a rule that makes sense does hold. The convolution of two L^2 functions is a bounded continuous functions, not necessarily in L^1, so we cannot consider at this stage its Fourier transform. However, their product is in L^1 and the convolution $\hat{f}*\hat{g}$ is a continuous function. So the rule $\widehat{fg} = \hat{f}*\hat{g}$ makes sense and would be proved by an approximation argument. In a similar way, for instance, if $f \in L^1(\mathbf{R^d}), \mathbf{g} \in \mathbf{L^2}(\mathbf{R^d})$, then $f*g \in L^2(\mathbf{R^d})$ and has Fourier transform the L^2 function $\hat{f}\hat{g}$.

With an abuse of notation we may thus write that indeed, in the sense above,

$$f = \int_{\mathbf{R^d}} \langle f, e_\xi\rangle e_\xi d\xi. \tag{4.3}$$

Some comments are in order. The e_ξ do not even belong to $L^2(\mathbf{R^d})$ (as they have modulus one), yet they behave as if they were an orthonormal basis of $L^2(\mathbf{R^d})$, as in the periodic case. More precisely, they behave like a rigid frame. Incidentally, real orthonormal basis of $L^2(\mathbf{R^d})$ can be constructed, namely the wavelet bases, but this is another direction of Fourier analysis.

Coming back to the Fourier transform, it is clear that (4.3) is due to the existence of cancellations that we try now to explain. To begin with, we have seen that the Fourier transform is an isometry of $L^2(\mathbf{R^d})$. Observe that it is given by a kernel $K(x,\xi) = e^{2\pi i x\cdot\xi}$ of modulus one; this is by the way enough to ensure that it is bounded from $L^1(\mathbf{R^d})$ to $L^\infty(\mathbf{R^d})$, but its boundedness in L^2 depends on much more than size. To ilustrate it, let us formally manipulate

$$\int_{\mathbf{R^d}} |\hat{f}(\xi)|^2\,d\xi = \int_{\mathbf{R^d}}\int_{\mathbf{R^d}}\int_{\mathbf{R^d}} f(x)\overline{f(y)}e^{2\pi i\xi\cdot(y-x)}dxdyd\xi.$$

This being equal to $\int_{\mathbf{R^d}} |f(x)|^2\,dx$ means formally that

$$\int_{\mathbf{R^d}} e^{2\pi i\xi\cdot x}\,d\xi = \delta_0(x).$$

We arrive to the same formal conclusion if we manipulate similarly the Fourier inversion theorem. The above says that superposition of all frequencies is zero outside zero. An intuitive way to understand this is by noting that $(d = 1)$

$$\int_{R}^{R} e^{2\pi i\xi x}\, d\xi = \frac{\sin 2\pi Rx}{\pi x},$$

is zero for x an integer multiple of $\frac{1}{R}$, so when $R \to +\infty$ the zeros become more and more dense.

A final remark regarding the definition itself of the Fourier transform in $L^2(\mathbf{R^d})$ is in order. Namely we have defined

$$\hat{f}(\xi) = \lim_{R\to+\infty} \int_{|x|\le R} f(x)e^{-2\pi i\xi\cdot x}\, dx,$$

where the convergence is in the L^2-norm. This implies that, given f, a sequence $R_j \to +\infty$ exists so that

$$\hat{f}(\xi) = \lim_{j\to+\infty} \int_{|x|\le Rj} f(x)e^{-2\pi i\xi\cdot x}\, dx,$$

almost everywhere, but does not imply a.e. convergence of the means. This is in fact true, a very deep result proved by L. Carleson.

4.4 Translation invariant operators in $L^p(\mathbf{R^d})$-spaces

With the above tools we can have a better understanding of translation invariant operators, as the next theorem shows.

Theorem 18. For a bounded operator $T : L^p(\mathbf{R^d}) \to \mathbf{L^q(R^d), p, q = 1, 2}$, the following are equivalent:

- It commutes with translations.
- It commutes with convolution with $L^1(\mathbf{R^d})$ functions.
- It diagonalizes in the Fourier basis: $\widehat{Tf}(\xi) = m(\xi)\hat{f}(\xi)$.

Moreover, the general form of T is given

- in case $p = q = 1$, by $Tf = f * \mu$, with a finite complex Borel measure μ, in which case $m(\xi) = \hat{\mu}(\xi)$ is continuous.
- in case $p = q = 2$, by $\widehat{Tf}(\xi) = m(\xi)\hat{f}(\xi)$ with an arbitrary bounded function m, in which case the norm of T as an operator in $L^2(\mathbf{R^d})$ equals $\|m\|_\infty$.

Proof. We need prove only that T diagonalizes in case it is translation invariant and commutes with convolution with L^1 functions. If $f, g \in L^1(\mathbf{R^d}) \cap \mathbf{L^2(R^d)}$ we will have

$$f * Tg = T(f * g) = T(g * f) = g * Tf,$$

and hence $\hat{f}\widehat{Tg} = \hat{g}\widehat{Tf}$. Take now as g the Gaussian function for which $\hat{g} = g > 0$ to find $\widehat{Tf} = m\hat{f}$, with $m = \frac{\widehat{Tg}}{\hat{g}}$, for all $f \in L^1(\mathbf{R^d}) \cap \mathbf{L^2}(\mathbf{R^d})$ whence for all f by density.

Once we have m the statement in $L^2(\mathbf{R^d})$ is an easy consequence of the Fourier transform being an isometry in L^2. □

We can describe as well all continuous operators acting on $\mathcal{S}(\mathbf{R^d})$ which are translation invariant. They will correspond to functions m that act on $\mathcal{S}(\mathbf{R^d})$ by multiplication.

Definition 4. We say that a C^∞ function ψ has slow growth if for every $\alpha \in \mathbf{N^d}$ there exists $k \in \mathbf{N}$ such that $|D^{alpha}\psi(x)| = O(|x|^k)|$.

We call $\mathcal{B}(\mathbf{R^d})$ the space of functions of slow growth.

Theorem 19. A function m operates on $\mathcal{S}(\mathbf{R^d})$ by multiplication if and only if $m \in \mathcal{B}(\mathbf{R^d})$. Then, $\widehat{T\varphi} = m\hat{\varphi}$ is the general form of a continuous translation invariant operator on $\mathcal{S}(\mathbf{R^d})$.

Notice that in the above statements if we take an approximation of the identity $f = k_\varepsilon$ with $k \in L^(\mathbf{R^d}) \cap \mathbf{L^2}(\mathbf{R^d})$, then

$$\widehat{Tk_\varepsilon}(\xi) = m(\xi)\hat{k}(\varepsilon\xi),$$

and since $\hat{k}(0) = \int k = 1$,

$$m(\xi) = \lim_{\varepsilon} \widehat{Tk_\varepsilon}(\xi),$$

which amounts to say that T is given by its action on the "delta" mass.

Another final comment: T is formally the convolution of f with

$$g(x) = \int_{\mathbf{R^d}} m(\xi)e^{2\pi i x\cdot\xi}\, d\xi.$$

But this object g is not in general a function but a distribution, see later.

The last result describing all translation invariant operators of $L^2(\mathbf{R^d})$ serves to describe all closed translations invariant subspaces E of $L^2(\mathbf{R^d})$. Indeed, let us associate to E the projection operator P onto E, that is, $Pf \in E$ and $f - Pf$ is orthogonal to E, $P^2 = P$. If E is invariant by translations so is P, hence it has a bounded multiplier $m \in \text{Ł}^\infty(\mathbf{R^d})$. Now, $P^2 = P$ translates to $m^2 = m$, whence $m = 0$ or $m = 1$. Let A be the set where $m = 0$. A given $f \in E$ if and only if $Pf = f$, that is $m\hat{f} = \hat{f}$, whence it follows that $f \in E$ if and only if $\hat{f}$ vanishes a.e. on A. This is the general form of a closed translation invariant subspace in $L^2(\mathbf{R^d})$. In particular, the translates of a a given function $f \in L^2(\mathbf{R^d})$ span the whole of $L^2(\mathbf{R^d})$ if and only if $\hat{f} \neq 0$ a.e. (Beurling's theorem)

Chapter 5

Distributions

5.1 What is a distribution?

The basic idea of distribution is to consider that functions f are not given by their values at points but by their action on other functions by integration. That is, two functions f, g defined in an open set U of $\mathbf{R^d}$ are equal if

$$\int_U f(x)\varphi(x)dx = \int_U g(x)\varphi(x)dx, \tag{5.1}$$

for all φ in a certain test space. The smaller the test space the stronger is the statement. So we chose once for all the smallest and nicer test space, the space that we have seen is dense in all $L^p(U)$ spaces. So the starting point is the statement 12 that we recall here

Proposition 6. . If $f, g \in L^1_{loc}(U)$ and (5.1) holds for all $\varphi \in$, then $f = g$ a.e.

This means that f is completely known as soon as one knows

$$u_f(\varphi) = \int_U f(x)\varphi(x)\, dx,$$

and is the basis of the following definition.

Definition 5. A *distribution* on U is a continuous linear map $u :\to \mathbf{C}$.

We understand continuity as follows: if φ_n is a sequence in that tends to zero (this meaning that they have their supports in a fixed compact set K of U and $D^\alpha(\varphi_n) \to 0$ uniformly in K), then $u(\varphi_n) \to 0$. It is customary to write $u(\varphi) = \langle u, \varphi\rangle$.

Thus every function $f \in L^1_{loc}$ is a distribution, in particular constants. A locally finite measure $d\nu$ on U is also a distribution. The Dirac measure at a will be denoted δ_a. If Λ is a discrete set in U (hence countable), the comb

$$\sum_{a\in\Lambda} \delta_a,$$

is also a distribution.

The following is an example of a distribution that is not a function nor a measure. We define the distribution $p.v.\frac{1}{x}$ in $\mathbf{R}$ by

$$\langle p.v.\frac{1}{x}, \varphi\rangle = \lim_{\varepsilon\to 0} \int_{|x|>\varepsilon} \frac{\varphi(x)}{x}\, dx.$$

Note that the limit exists because it equals

$$\lim_{\varepsilon\to 0} \int_{\varepsilon}^{\infty} \frac{\varphi(x) - \varphi(-x)}{x}\, dx.$$

The space of distributions on $\mathbf{R^d}$ is denoted $\mathcal{D}'(\mathbf{R^d})$

5.2 Operations with distributions

When defining an operation on distributions we look for consistency, and in doing so the definition comes up in a natural way. For instance, we want to define the translation $\tau_x u$ of a distribution in $\mathbf{R^d}$. The definition should be so that $\tau_x u_f = u_{\tau_x f}$ for $f \in L^1_{loc}$. Since

$$\int_{\mathbf{R^d}} \tau_x f(y)\varphi(y)dy = \int_{\mathbf{R^d}} f(y-x)\varphi(y)dy = \int_{\mathbf{R^d}} f(z)\varphi(z+x)dz = \\ = \int_{\mathbf{R^d}} f(z)\tau_{-x}\varphi(z)\, dz,$$

we must define for a general distribution u the translation $\tau_x u$ by the rule

$$\langle \tau_x u, \varphi\rangle = \langle u, \tau_{-x}\varphi\rangle.$$

A distribution in $\mathbf{R}$ is called *periodic* with period a if $\tau_a u = u$. All a-periodic functions are, and also the Dirac comb $\Delta_a = \sum_{n\in\mathbf{Z}} \delta_{na}$.

Let us define next the product of a distribution u with a function g. This cannot be other than

$$\langle gu, \varphi\rangle = \langle u, g\varphi\rangle,$$

and we realize that makes sense for $g \in C^\infty(U)$. For instance, trivially $g\delta_a = g(a)\delta_a$. Analogously, $g\Delta_a = \sum_n g(na)\delta_{na}$. Incidentally, $g\Delta_a$ represents the sampling of g every a units.

Also

$$\langle xv.p.\frac{1}{x}, \varphi = \lim_{\varepsilon\to 0} \int_{|x|>\varepsilon} \varphi(x)dx = \int_{\mathbf{R}} \varphi(x)dx = \langle 1, \varphi\rangle,$$

that is, $x\, v.p.\frac{1}{x} = 1$.

We define now the derivative $D^\alpha u$ of a distribution in the only possible way to keep consistency, namely, in view of the integration by parts rule,

$$\langle D^\alpha u, \varphi\rangle = (-1)^{|\alpha|}\langle u, D^\alpha \varphi\rangle.$$

In $\mathbf{R}$, the integration by parts rule $\int f'\varphi = -\int f(\varphi)'$ holds for all locally absolutely continuous functions (undefinite integrals of integrable functions), so that $(u_f)' = u_{f'}$ for those.

Consider the unit step of Heaviside function H, 1 for positive x and zero for negative x. Then

$$\langle H', \varphi\rangle = -\langle H, \varphi'\rangle = -\int_0^\infty \varphi'(x)dx = \varphi(0),$$

so that $H' = \delta_0$.

it is important in some cases to distinguish between classical derivative and derivative in the sense of distributions. Consider for example a function f which is continuously differentiable in the closed intervals determined by some points $a_1, \ldots, a_N$ where it has some jump discontinuities with jumps s_i. Then $(u_f)' = u_{f'} + \sum_i s_i \delta_{a_i}$. The a- periodic function which in each interval $[na, (n+1)a]$ is linear from 0 to 1 has derivative $\frac{1}{a} - S_a$.

As a final example, let us compute the derivative of $\log|x|$. Its action on φ is

$$-\int_{\mathbf{R}} \log|x|\varphi'(x)\,dx = -\lim_\varepsilon \left(\int_{-\infty}^{-\varepsilon} + \int_\varepsilon^{+\infty}\right)\log|x|\varphi'(x)\,dx.$$

If we integrate by parts we find that this equals

$$\lim_\varepsilon(\varphi(\varepsilon) - \varphi(-\varepsilon))\log\varepsilon + \lim_\varepsilon \int_{|x|>\varepsilon} \frac{\varphi(x)}{x}\,dx,$$

so that its derivative is $p.v.\frac{1}{x}$.

One can prove, in $\mathbf{R}$, that if $u' = 0$ then u is constant and that every distribution has a primitive.

5.3 Convergence of distributions

The notion of convergence of distributions that we adopt is simply the weak convergence: $u_n \to u$ means simply $\langle u_n, \varphi\rangle \to 0$ for all φ. With this definition all operations are continuous, in particular the differentiation. In particular we can consider series of distributions. We will be interested in trigonometric series

$$\sum_{n\in\mathbf{Z}} c_n e^{2\pi i \frac{n}{a} x}.$$

The partial sums act as

$$\langle \sum_{n=-N}^{N} c_n e^{2\pi i \frac{n}{a} x}, \varphi \rangle = \sum_{n=-N}^{N} c_n \hat{\varphi}(-\frac{n}{a}).$$

Now, $\hat{\varphi}$ is in the Schwarz class, so that $\hat{\varphi}(-\frac{n}{a}) = O(|n|^{-k})$ for all k. This shows that if the coefficients c_n are slowly increasing meaning that $c_n = O(|n|^k)$ for some k then the series indeed defines a distribution. This is not necessarily the Fourier series of a periodic function.

Let us consider the a- periodic function f equal to $\frac{x}{a}$ in $[0, a]$. By direct computation we find its Fourier series

$$f(x) = \frac{1}{2} + \frac{i}{2\pi} \sum_{n \neq 0} \frac{1}{n} e^{2\pi i \frac{n}{a} x}.$$

As it is convergent in $L^2(\mathbf{T})$ is also convergent as distributions, and its derivative is

$$f'(x) = -\frac{1}{a} sum_{n \neq 0} e^{2\pi i \frac{n}{a} x}.$$

But we shaw before that $f' = \frac{1}{a} - \Delta_a$, so it follows that the periodic distribution Δ_a has also a Fourier series

$$\Delta_a = \sum_{n \in \mathbf{Z}} \delta_{na} = \frac{1}{a} \sum_{n \in \mathbf{Z}} e^{2\pi i \frac{n}{a} x}.$$

5.4 Distributions with compact support

Definition 6. A distribution with compact support is a continuous linear map $u : C^\infty(U) \to \mathbf{C}$.

Continuity means here that if $\varphi_n \in C^\infty(\mathbf{R^d})$ tend to zero (meaning that $D^\alpha \varphi_n(x) \to 0$ uniformly on compacts) then $\langle u, \varphi_n \rangle \to 0$. This amounts to the existence of a constant C, a compact K and m such that

$$|\langle u, \varphi \rangle| \leq C \sup_{x \in K, |\alpha| \leq m} |D^\alpha \varphi(x)|.$$

It is completely known by its restriction to $C_c^\infty(U)$, by density. Again, we may think that u has compact support if it is capable to act against all $C^\infty(\mathbf{R^d})$ functions. To be precise, let us define that a distribution $u \in \mathcal{D}'(\mathbf{R^d})$ has support in a compact set K if $\langle u, \varphi \rangle = 0$ for all φ supported in the complement of K. Then one can see that it extends its action to $\varphi \in C^\infty$ in a continuous way.

The following is left as an exercise:

Theorem 20. A distribution u is supported in 0 if and only if is a finite linear combination of derivatives of δ.

The space of distributions with compact support is denoted $\mathcal{E}'(\mathbf{R^d})$

5.5 Tempered distributions, Fourier transform of tempered distributions

We would like to define the Fourier transform of a distribution. For $f \in L^1(\mathbf{R^d})$ we have

$$\int_{\mathbf{R^d}} \hat{f}(x)\varphi(x)dx = \int_{\mathbf{R^d}} f(x)\hat{\varphi}(x)dx,$$

and so we should define

$$\langle \hat{u}, \varphi \rangle = \langle u, \hat{\varphi} \rangle.$$

The problem with this definition is that $\hat{\varphi}$ is no longer in $\mathcal{D}(\mathbf{R^d})$. So we cannot define the Fourier transform of an arbitrary distribution and must restrict to a particular class of distributions.

Recall that we shaw in section 4.3. that the Schwarz space is invariant by the Fourier transform, so the above would work if $\mathcal{S}(\mathbf{R^d})$ were used instead of $\mathcal{D}(\mathbf{R^d})$ to define distributions. That's why we define

Definition 7. A tempered distribution in $\mathbf{R^d}$ is a continuous linear map $u : \mathcal{S}(\mathbf{R^d}) \to \mathbf{C}$

Here continuity means that if $\varphi_n \to 0$ in $\mathcal{S}(\mathbf{R^d})$, meaning that

$$\sup_x |x|^{|\beta|} |D^\alpha \varphi_n(x)| \to 0$$

as $n \to +\infty$ for all $\alpha, \beta \in \mathbf{N^d}$, then $\langle u, \varphi_n \rangle \to 0$. Again, this amounts to the existence of C and m such that

$$|\langle u, \varphi \rangle| \leq C \sup_{x \in \mathbf{R^d}, |\alpha| \leq \mathbf{m}} |D^\alpha \varphi(x)||x|^m.$$

The restriction of u to $\mathcal{D}(\mathbf{R^d})$ is then a distribution (and in fact u is completely determined by this restriction since $\mathcal{D}(\mathbf{R^d})$ is dense in $\mathcal{S}(\mathbf{R^d})$). Here it is enough that we look at tempered distributions as those that are capable to act on the larger space $\mathcal{S}(\mathbf{R^d})$. We denote by $\mathcal{S}'(\mathbf{R^d})$ the space of tempered distributions.

For instance, among the locally integrable functions f, those that have a slow growth, meaning that $|f(x)| = O(|x|^k)$ for some integer k are tempered distributions. All L^p- functions, $1 \leq p \leq +\infty$ are as well. It is easy to see too that all periodic functions, integrable over a period, are also tempered distributions. This is so because if $\varphi \in \mathcal{S}(\mathbf{R^d})$,

$$\sum_n |\varphi(x + na)|$$

is a bounded function.

It is easy to see that $gu \in \mathcal{S}'(\mathbf{R^d})$ if $u \in \mathcal{S}'(\mathbf{R^d})$ and $g \in \mathcal{B}(\mathbf{R^d})$, that is, for all $\alpha \in \mathbf{N^d}, |\mathbf{D^\alpha g(x)}| = \mathbf{O}(|\mathbf{x}|^\mathbf{k})$ for some k, because in this case $g\varphi \in \mathcal{S}(\mathbf{R^d})$ for all $\varphi \in \mathcal{S}(\mathbf{R^d})$.

The Fourier transform $\hat{u}$ of a tempered distribution is thus defined

$$\langle \hat{u}, \varphi \rangle = \langle u, \hat{\varphi} \rangle.$$

Since the Fourier transform in $\mathcal{S}(\mathbf{R^d})$ is an isomorphism with its inverse being the same transform composed with reflection, the same happens with $\mathcal{S}'(\mathbf{R^d})$. The properties of the Fourier transform regarding translations vs multiplication by exponentials and derivatives vs multiplication by polynomials go over to $\mathcal{S}'(\mathbf{R^d})$.

Example 1. 1. $\hat{\delta}_a(\xi) = -e^{2\pi i a\xi}, \widehat{e^{2\pi i a x}} = \delta_a$. In particular, $\widehat{\delta_0} = 1, \hat{1} = \delta_0$. This is a restatement of the inversion theorem.

2. In particular $\widehat{\Delta_a} = \sum_n \widehat{\delta_{na}} = \sum_n e^{2\pi i n a\xi}$. But we shaw before that this equals $\frac{1}{a}\Delta_{\frac{1}{a}}$. Therefore

$$\hat{\Delta_a} = \frac{1}{a}\Delta_{\frac{1}{a}},$$

and in particular, Δ_1 is its own Fourier transform.

3. Let us compute the Fourier transform of $p.v.\frac{1}{x}$. Its action on φ is

$$\lim_\varepsilon \int_{\varepsilon<|xi|<1/\varepsilon} \frac{\hat{\varphi}(\xi)}{\xi}\, d\xi =$$

$$\lim_\varepsilon \int_\mathbf{R} \varphi(x) \left(\int_{\varepsilon<|xi|<1/\varepsilon} e^{-2\pi i x\xi} \frac{d\xi}{\xi} \right) dx =$$

$$= -i \lim_\varepsilon \int_\mathbf{R} \varphi(x) \left(\int_{\varepsilon<|xi|<1/\varepsilon} \sin 2\pi x\xi \frac{d\xi}{\xi} \right) dx =$$

But the last inner integral is known to be uniformly bounded in ε, x and has limit $\pi\mathrm{sign}(x)$, so the Fourier transform of $p.v.\frac{1}{x}$ is $-i\pi\mathrm{sign}(\xi)$.

5.6 The Fourier transform of a distribution with compact support

If $f \in \mathcal{D}(\mathbf{R^d})$, then $\hat{f} \in \mathcal{S}(\mathbf{R^d})$. In fact something much more precise can be said. Note first that $\hat{f}(\xi)$ makes sense for $z \in \mathbf{C^d}$,

$$\hat{f}(z) = \int_{\mathbf{R^d}} f(x) e^{-2\pi i z\cdot x}\, dx.$$

and it is an entire function in $\mathbf{C^d}$ (in particular it cannot have compact support in $\mathbf{R^d}$).

Now, if $u \in \mathcal{E}'(\mathbf{R^d})$, as it is capable to act on C^∞- functions not necessarily with compact support, we may consider as well the entire function

$$h(z) = \langle u_x, e^{-2\pi i z\cdot x}\rangle,$$

which is formally $\hat{u}(x)$ for $x \in \mathbf{R^d}$. One can check that the two definitions of $\hat{u}, u \in \mathcal{E}'(\mathbf{R^d})$ agree, that is,

$$\langle u, \hat{\varphi}\rangle = \int_{\mathbf{R^d}} h(x)\varphi(x).$$

This means that for $u \in \mathcal{E}'(\mathbf{R^d})$, $\hat{u}$ is in fact the restriction to $\mathbf{R^d}$ of an entire function.

Moreover, it is easy to see that $\hat{u} \in \mathcal{B}(\mathbf{R^d})$. In fact, the *Paley-Wiener theorem* characterizes exactly the class of entire functions F that are Fourier transforms of $u \in \mathcal{E}'(\mathbf{R^d})$. They are exactly those of exponential type, meaning that

$$|F(z)| \leq Ce^{A|Imz|}, z \in \mathbf{C^d},$$

for some constants A, C, and such that the restriction to $\mathbf{R^d}$ is in $\mathcal{B}(\mathbf{R^d})$. Among these, the ones that are Fourier transforms of function $\varphi \in \mathcal{D}(\mathbf{R^d})$ are exactly those such that the restriction to $\mathbf{R^d}$ is in $\mathcal{S}(\mathbf{R^d})$.

5.7 Convolutions among functions and distributions

Now we would like to define convolutions, among functions and distributions first, and among distributions in a second step. From

$$g * f(x) = \int g(x-y)f(y)dy$$

we see that if we want to replace f by a general distribution u we should define by consistency

$$(g * u)(x) = \langle u_y, g(x-y)\rangle,$$

This makes sense in three cases, and defines a function lying in the indicated space

$$\mathcal{D}(\mathbf{R^d}) * \mathcal{D}'(\mathbf{R^d}) \subset \mathbf{C}^\infty(\mathbf{R^d}), \mathcal{S}(\mathbf{R^d}) * \mathcal{S}'(\mathbf{R^d}) \subset \mathbf{C}^\infty(\mathbf{R^d})$$
$$C^\infty(\mathbf{R^d}) * \mathcal{E}'(\mathbf{R^d}) \subset \mathbf{C}^\infty(\mathbf{R^d}), \mathcal{D}(\mathbf{R^d}) * \mathcal{E}'(\mathbf{R^d}) \subset \mathcal{D}(\mathbf{R^d})$$

Checking that $g * u$ is a C^∞ function depends on showing of course that the rule

$$D^\alpha(g * u) = D^\alpha g * u$$

holds. For instance, in case $|\alpha| = 1$, this is a consequence of the fact that for $g \in \mathcal{D}(\mathbf{R^d})$, the incremental quotients defining $D^\alpha g$ converge not only pointwise but in $\mathcal{D}(\mathbf{R^d})$, and similarly in the other cases.

It is an immediate consequence of the definition that convolution is continuous in both arguments.

When convolving $\varphi \in \mathcal{S}(\mathbf{R^d})$ with $u \in \mathcal{S}'(\mathbf{R^d})$, one can check just using the definition of continuity of a tempered distribution that moreover one has $\varphi * u \in \mathcal{B}(\mathbf{R^d})$. This makes sense to the rules

$$\widehat{\varphi * u} = \hat{\varphi}\hat{u}, \varphi \in \mathcal{S}(\mathbf{R^d}), \mathbf{u} \in \mathcal{S}'(\mathbf{R^d})$$
$$\widehat{\varphi u} = \hat{\varphi} * \hat{u}, \varphi \in \mathcal{S}(\mathbf{R^d}), \mathbf{u} \in \mathcal{S}'(\mathbf{R^d})$$

and they can be routinely checked too.

Next we want to define the convolution of two distributions. Again from the relations

$$\langle g * f, \varphi \rangle = \int \int g(x-y) f(y) \varphi(x) dx\, dy =$$
$$= \int (\int g(z)\varphi(y+z)dz) f(y)dy = \int (\int f(y)\varphi(y+z)dy) g(z)dz$$

we see that we should define

$$\langle u * v, \varphi \rangle = \langle u_z, \langle v_y, \varphi(y+z) \rangle \rangle,$$

or else

$$\langle u * v, \varphi \rangle = \langle v_y, \langle u_z, \varphi(y+z) \rangle \rangle.$$

Written in a more compact form, using the notation $\varphi_\sigma(x) = \varphi(-x)$,

$$\langle u * v, \varphi \rangle = \langle u, (\varphi_\sigma * v)_\sigma \rangle, \varphi \in \mathcal{D}(\mathbf{R^d}).$$

Also note that for this to make sense, one of the distributions must have compact support. Fortunately one may routinely check that the two definitions agree and defines a commutative operation $u*v \in \mathcal{D}'(\mathbf{R^d})$. Again, convolution is continuous in both variables and

$$D^\alpha(u * v) = (D^\alpha u) * v = u * D^\alpha v.$$

Note that of course $\delta * u = u, \mathbf{D}^\alpha \delta * \mathbf{u} = \mathbf{D}^\alpha \mathbf{u}$ for all distributions u.

Now assume that $u \in \mathcal{E}'(\mathbf{R^d}), \mathbf{v} \in \mathcal{S}'(\mathbf{R^d})$. Then still $\varphi * v \in C^\infty(\mathbf{R^d})$, and so the definition above makes sense for $\varphi \in \mathcal{S}(\mathbf{R^d})$, and $u * v \in \mathcal{S}'(\mathbf{R^d})$. Then, as one may expect

$$\widehat{u * v} = \hat{u}\hat{v}$$

is another relation that makes sense (because $\hat{u} \in \mathcal{B}(\mathbf{R^d})$) and so does hold.

5.8 Translation invariant operators in spaces of distributions

With the language of distributions and tempered distributions one can state a number of representation theorems for continuous operators in spaces of distributions invariant by translations.

Of course the definition is the same. If T acts on distributions, we say that T is invariant by translations if $T\tau_x = \tau_x T$. We will show that generally speaking T is always given by convolution.

Assume first that $T : \mathcal{D}(\mathbf{R^d}) \to \mathbf{C}(\mathbf{R^d})$ is translation invariant. Then $\varphi \to T(\varphi)(0)$ is a distribution, and using translation invariance we find that $T(\varphi) = \varphi * u$ for $u \in \mathcal{D}'(\mathbf{R^d})$, this is the general form of T. That is almost a tautology.

It will take values in $\mathcal{D}(\mathbf{R^d})$ if and only if $u \in \mathcal{E}'(\mathbf{R^d})$. Thus, convolution by $u \in \mathcal{E}'(\mathbf{R^d})$ is the general form of a continuous t.i.p. from $\mathcal{D}(\mathbf{R^d})$ to itself.

In an analogous way, convolution by $u \in \mathcal{E}'(\mathbf{R^d})$ is the general form of a continuous translation invariant operator from $C^\infty(\mathbf{R^d})$ to itself, and convolution by $u \in \mathcal{S}'(\mathbf{R^d})$ is the general form of one from $\mathcal{S}(\mathbf{R^d})$ to $C^\infty(\mathbf{R^d})$. It will take $\mathcal{S}(\mathbf{R^d})$ to itself iff $\hat{u} \in \mathcal{B}(\mathbf{R^d})$.

Now, a basic fact of the spaces $C^\infty(\mathbf{R^d}), \mathcal{D}(\mathbf{R^d})$ and $\mathcal{S}(\mathbf{R^d})$ is that they are *reflexive*. This means that they equal their second dual. In other words, if for instance $\omega : \mathcal{D}'(\mathbf{R^d}) :\to \mathbf{C}$ is a continuous linear functional, then there exists $\varphi \in \mathcal{D}(\mathbf{R^d})$ such that $\omega(u) = \langle u, \varphi \rangle$. Analogously, every continuous linear functional on $\mathcal{E}'(\mathbf{R^d})$ is given by testing on some $\varphi \in C^\infty(\mathbf{R^d})$ and every continuous linear functional on $\mathcal{S}'(\mathbf{R^d})$ is given by testing on some $\varphi \in \mathcal{S}(\mathbf{R^d})$.

Using this fact it is easy to prove that the general form of a continuous translation invariant operator $T : \mathcal{D}'(\mathbf{R^d}) \to \mathcal{D}'(\mathbf{R^d})$ is convolution against some $u \in \mathcal{E}'(\mathbf{R^d})$. Indeed, for all $\varphi \in \mathcal{D}(\mathbf{R^d})$, the composition $u \to \langle Tu, \varphi \rangle$ is a continuous linear functional, and so it is given by testing against some other function $S(\varphi) \in \mathcal{D}(\mathbf{R^d})$. The operator S thus defined is continuous and translation invariant and so it consists in convolving with some distribution with compact support. From this one we get the required u.

In a similar way, the general form of a continuous translation invariant $T : \mathcal{E}'(\mathbf{R^d}) \to \mathcal{E}'(\mathbf{R^d})$ is convolution by some $u\mathcal{E}'(\mathbf{R^d})$, and convolution with $u \in \mathcal{S}'(\mathbf{R^d})$, such that $\hat{u} \in \mathcal{B}(\mathbf{R^d})$, is the general form of a continuous translation invariant operator from $\mathcal{S}'(\mathbf{R^d})$ to $\mathcal{S}'(\mathbf{R^d})$.

In general, if X, Y are some spaces of tempered distributions, every continuous t.i.p T from X to Y is given by convolution with some distribution u. For instance,

Theorem 21. If $T : L^p(\mathbf{R^d}) \to \mathbf{L^q}(\mathbf{R^d})$ is bounded and translation invariant, there exists $u \in \mathcal{S}'(\mathbf{R^d})$ such that $T(f) = f*, f \in \mathcal{S}(\mathbf{R^d})$.

Proof. Indeed T it will commute with convolution with L^1 functions so

$$T(\varphi * \psi) = \varphi * T(\psi) = T(\varphi) * \psi,$$

hence

$$\hat{\varphi} * \widehat{T\psi} = \hat{\psi}\widehat{T\varphi}, \varphi \in L^1(\mathbf{R^d}) \cap \mathbf{L^p}(\mathbf{R^d}), \psi \in \mathbf{L^p}(\mathbf{R^d}).$$

Thus $\widehat{T\psi} = m\hat{\psi}$, with $m = \frac{\widehat{T\varphi}}{\hat{\varphi}}$. Choosing $\varphi(x) = e^{-|x|}$, $\hat{\varphi}(\xi)$ is the Poisson kernel whose inverse has polynomial growth. Then m is the product of an L^q function with a function of polynomial growth, and so is a tempered distribution. It now suffices to choose u so that $\hat{u} = m$. □

5.9 Fundamental solutions

Definition 8. If $T : \mathcal{D}'(\mathbf{R^d}) \to \mathcal{D}'(\mathbf{R^d})$ is continuous and translation invariant, we say that $E \in \mathcal{D}'(\mathbf{R^d})$ is a *fundamental solution* if $T(E) = \delta_0$.

In this case $T(E * v) = T(E) * v = \delta * v = v$ for $v \in \mathcal{E}'(\mathbf{R^d})$ and generally whenever $E * f$ makes sense. Thus convolution with E is an inverse of T.

We state without proof the following theorem:

Theorem 22. (Malgrange-Ehrenpreis theorem): every linear constant coefficient operator $T = P(D)$ has a fundamental solution.

Note that if T is a translation invariant operator in $\mathcal{S}'(\mathbf{R^d})$, then it is convolution with $u \in \mathcal{S}'(\mathbf{R^d})$ with $\hat{u} = m \in \mathcal{B}(\mathbf{R^d})$, so $TE = \delta$ is equivalent to $1 = m\hat{E}$. If it happens that $\frac{1}{m} \in \mathcal{S}'(\mathbf{R^d})$, then the tempered distribution with $\hat{E} = \frac{1}{m}$ is a fundamental solution.

For example, for the laplacian Δ, $m(\xi) = -4\pi^2|\xi|^2$. Here, $\frac{1}{m}$ is locally integrable if $d > 2$ and so we can find a fundamental solution $E = \widehat{1/m}$. It must be radial, and by homogeneity it must be of the form $c_d|x|^{d-2}$. For $d = 2$, a direct proof shows that $E = c_2 \log|x|$ is a fundamental solution.

The fundamental solution is of course not unique since we cal add to it a solution of $Tu = 0$. For the Laplacian, the tempered distributions u such that $\Delta u = 0$ must satisfy $|\xi|^2\hat{u} = 0$, whence $\hat{u}$ is supported in zero, whence $\hat{u}$ is a finite linear combination of derivatives of the delta function, and therefore u is a (harmonic) polynomial. The fundamental solution E found above is the only one vanishing at ∞.

In connection to this we mention the Weyl's lemma that holds true more generally for elliptic operators.

Theorem 23. Weyl's lemma: if $f \in C^\infty(\mathbf{R^d})$ and $\Delta u = f$ in the sense of distributions, then $u \in C^\infty(\mathbf{R^d})$ and $\Delta u = f$ in the classical sense.

5.10 The Poisson summation formula.

Let us go back now to what we found before, formula

$$\frac{1}{a}\widehat{\Delta_{\frac{1}{a}}} = \Delta_a$$

as tempered distributions. This means exactly that

$$\sum_n \varphi(na) = \frac{1}{a}\sum_n \widehat{\varphi}(\frac{n}{a}), \varphi \in \mathcal{S}(\mathbf{R^d}).$$

Replacing φ by $\tau_{-x}\varphi$ we get

$$\sum_n \varphi(x-na) = \frac{1}{a}\sum_n \widehat{\varphi}(\frac{n}{a})e^{2\pi i x\frac{n}{a}}, \tag{5.2}$$

or interchanging φ and $\hat{\varphi}$,

$$\sum_n \hat{\varphi}(x-na) = \frac{1}{a}\sum_n \varphi(na)e^{-2\pi i x\frac{n}{a}}.$$

This is known as *Poisson's summation formula* and if fact holds for a much larger class of functions, as the following argument shows, that constitutes as well a direct proof, that works as well in $\mathbf{R^d}$

For $f \in L^1(\mathbf{R})$ we consider its a-periodized function

$$F(x) = \sum_n f(x-na).$$

Indeed, this series converges a.e. and defines an a-periodic function in $L^1(0,a)$. This is so because

$$\int_0^a \sum_{N<|n|<M} |f(x-na)|dx \leq \int_{|x|\geq Na} |f(x)|\,dx.$$

The Fourier coeficients of F are

$$c_k(F) = \frac{1}{a}\int_0^a \sum_n f(x-na)e^{-2\pi i k\frac{x}{a}}\,dx = \sum_n \int_n^{n+1} f(x)e^{-2\pi i k\frac{x}{a}} =$$
$$= \frac{1}{a}\int_{\mathbf{R}} f(x)e^{-2\pi i k\frac{x}{a}} = \frac{1}{a}\hat{f}(\frac{k}{a}).$$

Hence the right hand side of (5.2) is the formal power series of F. If F satisfies some of the criteria for pointwise convergence, then (5.2) will hold pointwise. One such criteria is that F is absolutely continuous. It can be seen that this is the case if f is differentiable with derivative in $L^1(\mathbf{R})$.

5.11 Tempered distributions, a common frame for the Fourier transform in $\mathbf{Z}, \mathbf{Z_N}, \mathbf{T}, \mathbf{R}$

Assume that f is a a-periodic function integrable in one period. It has a sequence of Fourier coefficients

$$c_n(f) = \frac{1}{a}\int_0^a f(x)e^{-2\pi i \frac{n}{a}x}\,dx$$

and a formal series

$$\sum_n c_n(f)e^{2\pi i \frac{n}{a}x}.$$

This series does not converge in general to f. Let us look at f as a tempered distribution and let us compute its Fourier transform.

$$\langle \hat{f}, \varphi\rangle = \langle f, \hat{\varphi}\rangle = \int_{\mathbf{R}} f(x)\hat{\varphi}(x)\,dx =$$

$$= \int_0^a f(x)\sum_n \hat{\varphi}(x-na)dx =$$

$$= \frac{1}{a}\int_0^a f(x)\sum_n \varphi(\frac{n}{a})e^{-2\pi i x\frac{n}{a}} = \sum_n c_n(f)\varphi(\frac{n}{a})$$

This means that as a tempered distribution

$$\hat{f} = \sum_n c_n\delta_{\frac{n}{a}}$$

the sum being convergent in $\mathcal{S}'(\mathbf{R^d})$. By applying the inverse Fourier transform we find that

$$f = \sum_n c_n(f)e^{2\pi i\frac{n}{a}x},$$

this development being in $\mathcal{S}'(\mathbf{R^d})$.

Something similar happens with an infinite digital sequence $x = (x_n)$. This is identified with $x = \sum_n x_n\delta_n$, which certainly is a tempered distribution if $x \in l^p(\mathbf{Z})$ or if x_n has slow growth. Its Fourier transform as a tempered distribution is

$$\sum_n x_n e^{-2\pi i n\xi}$$

which again is its definition as a Fourier transform in $\mathbf{Z}$.

Now suppose that this signal is N-periodic, that is, $x_{j+N} = x_j$. This means that the tempered distribution x can be written

$$x = \sum_{k=0}^{N-1} x_k\tau_k\Delta_N,$$

and hence

$$\hat{x}(\xi) = \sum_{k=0}^{N-1} x_k \widehat{\tau_k \Delta_N}(\xi) = \sum_{k=0}^{N-1} x_k e^{-2\pi i \xi k} \frac{1}{N} \Delta_{\frac{1}{N}} =$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} x_k \sum_{j \in \mathbf{N}} e^{-2\pi i \frac{j}{N} k} \delta_{\frac{j}{N}} =$$

$$= \sum_{j \in \mathbf{N}} (\frac{1}{N} \sum_{k=0}^{N-1} e^{-2\pi i \frac{j}{N} k}) \delta_{\frac{j}{N}},$$

This is another N- periodic signal whose N defining numbers are the Fourier transform defined in section 2.4.

We thus see that all four definitions of the Fourier transform for each of the different groups fit together in the common frame of tempered distributions. It can be proved that any a- periodic tempered distribution T has a Fourier series convergent in $\mathcal{S}(\mathbf{R}^{\mathbf{d}})'$, and that $\hat{T}$ is supported in $\frac{\mathbf{N}}{a}$.