# AN INTRODUCTION TO SOME MODERN ASPECTS OF HARMONIC ANALYSIS

by A.H. Dooley
School of Mathematics
University of New South Wales
Sydney, 2052, Australia

## FOREWORD.

The following notes were prepared as the text of a set of lectures given at the Workshop in Geometry and Analysis, held at the Australian National University in January 1995, and aimed at a mixed audience of postgraduate students and academics.

The aim of the lecture series was ambitious: in five hours to start from the level of undergraduate Fourier analysis and bring the participants to the point where they could take subsequent advanced lecture courses on Calderón-Zygmund theory, semigroup theory, distributions etc.

I have sought to give an overview of what I see as the important parts of the theory. There are no proofs or examples to speak of, indications of some of these being provided in lectures. An exception is in Chapter 5, where I felt that the proof of Malgrange-Ehrenpreis nicely drew together the rest of the material.

The last chapter provides an introduction to noncommutative harmonic analysis.

## CHAPTER 1: THE FOURIER TRANSFORM

### 1.1. Introduction.

In most undergraduate degrees, one studies the theory of Fourier series and Fourier transforms. In this lecture, we begin with a brief resumé of this material and then show how it is part of a more general picture.

### 1.2. The Fourier transform on the circle.

Given a $2\pi$-periodic function $f$ on $\mathbb{R}$, we define its Fourier transform by

$$\hat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}dx, \text{ for } n \in \mathbb{Z}.$$

For this to make sense, $f$ should be (Lebesgue) integrable over the interval $(-\pi, \pi]$ and complex-valued. Of course, any function on $(-\pi, \pi]$ defines a $2\pi$-periodic function on $\mathbb{R}$ and one often considers instead, functions on $(-\pi, \pi]$. In fact, it is most convenient to think of the periodic functions on $\mathbb{R}$ as functions on the unit circle $\mathbb{T}$ in $\mathbb{C}$, via

$$f(x) = g(e^{ix}).$$

Here $g$ is a function on $\mathbb{T}$ and $f$ a $2\pi$-periodic function on $\mathbb{R}$.

If we adopt this identification, the formula for the Fourier series is given by

$$\hat{g}(n) = \int_{\mathbb{T}} g(z)\bar{z}^n dz.$$

where $\int_{\mathbb{T}} dz$ is the usual contour integral around the unit circle.

From an engineer's point of view, the Fourier coefficients are the frequencies that "make up" the signal $f$. The basic problem of Fourier series is, to study the relationship between a function and its Fourier series, with regard to integrability, etc.

I should note that the average engineer does not use complex notation for the Fourier series. For her, the function $f$ is real-valued and our $\hat{f}(n)$ is written as $\frac{1}{2}(a_n - ib_n)$, $a_n$, $b_n$ being real numbers, given by $a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\cos nx\, dx$ and $\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\sin nx\, dx$ respectively. However, the notation above is more compact and convenient.

### 1.3. Some basic facts about the Fourier transform.

Let $L^1(\mathbb{T})$ (or $L^1$) denote the set of complex-valued function on $\mathbb{T}$ so that

$$\|f\|_1 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(e^{ix})|dx = \int_{\mathbb{T}} |f(z)|dz < \infty.$$

**Theorem.** (i) *Let $f$ and $g$ belong to $L^1, \alpha \in \mathbb{C}$. Then $(f + \alpha g)\hat{}(n) = \hat{f}(n) + \alpha\hat{g}(n)$. ($\hat{}$ is linear)*

(ii) *If $\hat{f}(n) = 0$ for all $n$ then $f = 0$ a.e. ($\hat{}$ is injective).*

(iii) *If $f$ is differentiable and $f' \in L^1$ then $(f')\hat{}(n) = -in\hat{f}(n)$.*

(iv) $\sup_n |\hat{f}(n)| \leq \|f\|_1$.

(v) *If $f \in L^1$, $\hat{f}(n) \to 0$ as $n \to \infty$ (Riemann-Lebesgue Lemma) ($\hat{}$ maps $L^1(\mathbb{T})$ into $c_0(\mathbb{Z})$).*

(vi) *If* $f$ *and* $g \in L^1$, *their* **convolution** *is defined by* $f * g(z) = \int_{\mathbb{T}} f(w^{-1}z)g(w)dw$ *and one has*

$$(f * g)\hat{}(n) = \hat{f}(n)\hat{g}(n)$$

*(*$\hat{}$ *converts convolution to multiplication).*

These facts will be proved later in a more general context.

While the Fourier transform maps $L^1(\mathbb{T})$ into $c_0(\mathbb{Z})$, it is certainly not an onto map, and this is one of the more annoying features of the whole subject – or to put it another way, this fact is the basis for many fascinating research directions!

In fact, the image of $L^1$ can be easily seen to be dense in $c_0(\mathbb{Z})$ (Stone-Weistrass) but attempts to characterize it are not satisfactory. For example, given any positive sequence $\xi_n \to 0$, there is an $L^1$ function $f$ such that $\hat{f}(n) \geq \xi_n$. There are many other results of this kind.

### 1.4. The Fourier series.

Of course, the engineer's ultimate dream is to reconstruct the signal from the Fourier coefficients.

Thus, we define the Fourier series of a function $f$ by

$$f(x) \sim \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{inx}.$$

Just what "$\sim$" means is left to be described hereunder. At the moment, it represents a hope that we can recover $f$ from its Fourier transform. The problem, intimately connected to the problem of the previous section, is to say under what conditions on $f$ the series on the right hand side converges, and in what sense it converges.

There are two circumstances when things work which are traditionally taught to engineers, when $f$ is differentiable and when $f$ is square integrable. Let us discuss these in turn.

The $L^2$ theory is the most complete and satisfying, so we start with that. By $L^2(\mathbb{T})$ we mean the set of all functions on $\mathbb{T}$ so that $(\int_{\mathbb{T}} |f(z)|^2 dz)^{1/2} = \|f\|_2 < \infty$.

Indeed, we define $L^p(\mathbb{T})$ to be the set of functions so that $(\int_{\mathbb{T}} |f(z)|^p dz)^{1/p} = \|f\|_p < \infty$ and $L^\infty(\mathbb{T})$ to be the set of essentially bounded function on $\mathbb{T}$; those for which $\|f\|_\infty =$ ess sup $|f(x)| < \infty$.

**Beware:** We treat the elements of $L^p$ as functions; they are in fact equivalence classes of functions under the equivalence $f = g$ a.e.

By standard Lebesgue theory (dominated convergence theorem) one shows that for each $p$, $1 \leq p \leq \infty$, $L^p$ is a complete space and indeed a Banach space under the norm $\|\cdot\|_p$. (Minkowski's inequality)

$L^2$ is distinguished amongst the $L^p$-spaces by the fact that it is a Hilbert space with inner product

$$\langle f, g\rangle = \int_{\pi} f(z)\overline{g(z)}dz.$$

[**Exercise:** show that if $f, g \in L^2$ then the integral converges.]

Furthermore, the functions $\{z^n : n \in \mathbb{Z}\}$ (or if you prefer, $\{e^{inx} : n \in \mathbb{Z}\}$) form a complete orthonormal set in $L^2$.

The Fourier coefficient then can be seen as the coefficient of $f$ in this expansion

$$\hat{f}(n) = \langle f, z^n\rangle$$

Let $\ell^2(\mathbb{Z})$ denote the Hilbert space of square-summable sequences $a = \{a_n\}_{n\in\mathbb{Z}}$ (i.e. $(\Sigma|a_n|^2)^{1/2} = \|a\|_2 < \infty$) with inner product $\langle a, b\rangle = \Sigma a_n\overline{b_n}$. We have, by functional analytic techniques,

**Theorem** (Plancherel) The Fourier transform is an injective isometry from $L^2(\mathbb{T})$ to $\ell^2(\mathbb{Z})$.

This means that $f \in L^2$ if and only if $\hat{f} \in \ell^2$, that $\|f\|_2 = (\Sigma|\hat{f}(n)|^2)^{1/2}$, and that furthermore, the Fourier series

$$(S_N f)(x) = \sum_{|n|\leq N} \hat{f}(n)e^{inx}$$

converges to $f$ in the $L^2$ sense, i.e.

$$\|S_N f - f\|_2 \to 0 \text{ as } N \to \infty.$$

The engineer's dream has been realized, in some senses. From an engineer's point of view, the $L^2$ norm represents the energy in the signal and the number $|\hat{f}(n)|$ the energy in the $n$th component. Unfortunately, this theorem is not totally satisfactory from the engineering perspective. Even if $f$ is continuous, there can be points $x_0$ where $S_N f(x_0)$ diverges. This can even happen on a dense set of points (example of Kolmogorov).

In 1965, Carleson showed that the Fourier series of an $L^2$ function converges almost everywhere. This was a major breakthrough, and the proof is still very complicated.

The following result goes right back to the beginning of the subject and ensures that if $f$ is continuously differentiable then the Fourier series converges uniformly to the function.

**Theorem** (Fourier) Suppose that $f$ is piecewise continuous on $\mathbb{T}$ and has left and right hand derivatives at $x = a$. Then the Fourier series for $f$ converges at $x = a$ to

$$\frac{1}{2}(f(a^+) + f(a^-)).$$

In particular, if $f$ is continuous at $a$, $S_N f(a) \to f(a)$.

Further, if $f$ is piecewise differentiable then $S_N f \to f$ uniformly on any interval on which $f$ is continuous.

**Hint of proof:**

Write

$$\begin{aligned} S_N f(x) &= \sum_{n=-N}^{N} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(u) e^{-inu} du e^{inx} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(u) \sum_{n=-N}^{N} e^{-in(u-x)} du \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(u) \frac{\sin(N+\frac{1}{2})(u-x)}{\sin\frac{1}{2}(u-x)} du \end{aligned}$$

(You will recognise this as a convolution of $f$ with a function known as the Dirichlet kernel.)

$$\begin{aligned} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(u) \frac{\frac{1}{2}(u-x)}{\sin\frac{1}{2}(u-x)} \frac{\sin(N-\frac{1}{2})(u-x)}{\frac{1}{2}(u-x)} du \\ &\to \frac{f(u^+) + f(u^-)}{2} \quad \text{as } N \to \infty. \end{aligned}$$

The reason for this is that $\frac{\sin u}{u}$ has integral $\frac{\pi}{2}$ on $(0, \infty)$. By the Riemann-Lebesgue lemma, $\lim_{n\to\infty} \int_0^a \psi(x) \frac{\sin nx}{x} dx = \frac{\pi}{2}\psi(0^+)$ whenever $\psi$ is continuously differentiable on $(0, a)$ and $\psi'(0^+)$ exists. $\square$

These results notwithstanding, Kolmogorov has given an example of an $L^1$ function whose Fourier series diverges **everywhere!** Notice that $\|D_N\|_1 = O(\log N)$.

## 1.5. Fourier transforms on $\mathbb{R}$ and $\mathbb{R}^n$.

We now pass to consideration of Fourier transforms on $\mathbb{R}$. Actually, it is almost equivalent to write it all down for $\mathbb{R}^n$, so I will do that.

If $f$ is a complex-valued Lebesgue integrable function on $\mathbb{R}^n$, we define its Fourier transform by

$$\hat{f}(\xi) = \frac{1}{(2\pi)^{n/2}} \int f(x) e^{-i\xi . x} dx.$$

Here, $\xi \in \mathbb{R}^n$ and $\xi . x$ denotes the usual dot product. The integral is with respect to Lebesgue measure on $\mathbb{R}^n$. The factor outside the integral is chosen differently by different authors. I have followed the convention of Stein. In fact, it is necessary to choose some factor for the "$dx$" and some other factor for "$d\xi$". One of these factors having been chosen, the other is prescribed by the desire to recover $f$ from its Fourier transform. In our case, we have the formula

$$f(x) \sim \frac{1}{(2\pi)^{n/2}} \int \hat{f}(\xi) e^{ix.\xi} d\xi.$$

for "suitable" functions $f$.

We ask to what extent the results of the previous section hold. In fact, there is a very precise analogue of the first theorem.

**Theorem.** *The Fourier transform is a linear injection from $L^1(\mathbb{R}^n)$ into the space $C_0^b(\mathbb{R}^n)$ of uniformly continuous functions which vanish at infinity.*

*Furthermore, if $f, g \in L^1$ then*

$$(f * g)\hat{}(\xi) = \hat{f}(\xi)\hat{g}(\xi)$$

*and*

$$(\frac{1}{i}\frac{\partial f}{\partial \xi_k})\hat{}(\xi) = \xi_k \hat{f}(\xi)$$

*(provided the partial derivative also belongs to $L^1$).*

*Finally, if $f \in L^1, \lambda \in \mathbb{R}_+$ $\hat{f}_\lambda(\xi) = \lambda^{-n}\hat{f}(\xi/\lambda)$, where $f_\lambda(x) = f(\lambda x)$.*

**Remarks:** Convolution is defined as you would expect:

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

Notice that $\mathbb{Z}$ is replaced by $\mathbb{R}^n$ here. We think of this as a different copy of $\mathbb{R}^n$. Engineers call it phase space. Sometimes we write $\widehat{\mathbb{R}}^n$ to distinguish it from the original copy of $\mathbb{R}^n$.

The $L^2$-theory also has a very precise analogue, after some initial technical difficulties. Let me remind you of Hölder's inequality. If $f \in L^p$ and $g \in L^q$ and $\frac{1}{p} + \frac{1}{q} = 1$, then

$$|\int f(x)g(x)dx| \leq \|f\|_p \|g\|_q$$

(when $p = q = 2$ this reduces to the familiar Cauchy-Schwarz inequality which guarantees that $L^2$ is a Hilbert space.) For other values of $p$, this inequality can be used to show that the dual of $L^p$ is $L^q$ (for $1 \leq p < \infty$: it is not true that the dual of $L^\infty$ is $L^1$).

In the case of the circle, Hölder's inequality implies that $L^2 \subseteq L^1$, and hence for $f \in L^2$, $\hat{f}(n)$ is defined. (Take $g(x) = sgn f(x)$).

On $\mathbb{R}^n$, however, which is an infinite measure space, $L^2$ is no longer contained in $L^1$.

Furthermore, the functions $x \mapsto e^{i\xi . x}$ no longer belong to $L^1$ or to $L^2$.

Nevertheless, one may consider $L^1 \cap L^2(\mathbb{R}^n)$, the functions which belong to both $L^1$ and $L^2$. This set of functions is dense in $L^2$, and one shows that if $f \in L^1 \cap L^2$ then $\hat{f} \in L^1 \cap L^2$ and $\|f\|_2 = \|\hat{f}\|_2$.

By functional analysis, one may extend the Fourier transform to a linear isometry (also called the Fourier transform)

$$\hat{} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n).$$

This is our version of Plancherel's theorem for $\mathbb{R}^n$. In fact, it is not hard to show that

$$\hat{\hat{f}}(x) = f(-x)$$

and so the fourth power of the Fourier transform is the identity.

In fact, the same technique can be used to extend the Fourier transform to $L^p(\mathbb{R}^n)$ for $1 < p < 2$. One considers $f \in L^1 \cap L^p$ and shows that $\hat{f} \in L^q$ and $\|\hat{f}\|_q \leq \|f\|_p$. Thus $\hat{}$

can be extended to a bounded linear map $L^p \to L^q$, where $\frac{1}{p} + \frac{1}{q} = 1$. We shall study this more in Lecture 2.

There are various analogues of Fourier's theorem on $\mathbb{R}^n$. Probably the most primitive have the form

$$f(x) = \int \hat{f}(\xi) e^{i\xi \cdot x} d\xi$$

for $f \in C_c^\infty(\mathbb{R}^n)$. Notice that a continuous function is not necessarily integrable on $\mathbb{R}^n$. We shall return to this topic in Lecture 3.

Indeed, a preoccupation of the rest of the lectures is to what extent one can make sense of the Fourier transform in larger and larger settings.

**1.6. LCA groups.**

The theory just advanced is all part of a much more general theory called "abstract harmonic analysis" as distinct from "Fourier analysis". Although most of the remainder of these lectures will be about $\mathbb{R}^n$, I would like to give you some indications of what happens in general.

In a sense, the previous theory is about breaking down a symmetry, either periodic (circular) or linear into its fundamental building blocks. To generalize the theory we need to have symmetries; a group, and things work most nicely when you have a locally compact group $G$.

That is $G$ is simultaneously a group and a locally compact topological space, and the operations

$$(x, y) \mapsto xy \text{ from } G \times G \text{ to } G$$

$$\text{and} \qquad x \mapsto x^{-1} \text{ from } G \text{ to } G$$

are both assumed continuous.

The open sets of the topology on $G$ then generate a $\sigma$-algebra $\mathcal{B}$, called the Borel $\sigma$-algebra.

**Theorem.** *There exists a measure $\lambda$ on $\mathcal{B}$ such that*

(i) *For all $E \in \mathcal{B}$ $\lambda(g.E) = \lambda(E)$ ($\lambda$ is left invariant).*

(ii) $\lambda(E) = $ *inf* $\{\lambda(U) : U$ *open* $U \supseteq E\} = \sup\{\lambda(C) : C$ *compact,* $C \subseteq E\}$. *($\lambda$ is regular).*

(iii) *There is an open set $U$ with $0 < \lambda(U) < \infty$.* *($\lambda$ is nontrivial)*

*$\lambda$ is unique up to scalar multiplication.*

Such a measure is called **Haar measure** on $G$. In fact, instead of writing $\int f(x)d\lambda(x)$, one usually writes $\int_G f(x)dx$ to indicate the analogy with Lebesgue measure on $\mathbb{R}^n$ on $\mathbb{T}$.

Thus, in the setting of locally compact groups, one has analogues of $L^1(G)$, $L^2(G)$ and $L^p(G)$, $C(G)$ etc.

In my last lecture I will consider what happens when $G$ is nonabelian. For the moment, let's just consider $G$ to be locally compact abelian (LCA).

The analogue of $\mathbb{Z}$ (in the case of $\mathbb{T}$) and $\mathbb{R}^n$ (phase space) in the case of $\mathbb{R}^n$ is well understood in this setting.

Let $\widehat{G}$ denote the set of **characters** of $G$. These are the continuous homomorphisms $G \to \mathbb{T}$.

(Exercise: Check that $\widehat{\mathbb{T}} = \mathbb{Z}$ and $\widehat{\mathbb{R}} = \mathbb{R}$).

Obviously, the product of two characters is again a character. In fact, we can endow $\widehat{G}$ with a topology that makes this product into a continuous one, the hull-kernel topology.

[For each compact set $K \subseteq G$, and for each $\epsilon > 0$, and for each $\chi_0 \in \widehat{G}$, let

$$B(K, \epsilon, \chi_0) = \{\chi \in \hat{G} : |\chi(x) - \chi_0(x)| < \epsilon \ \forall x \in K\}$$

The $B(K, \epsilon, \chi_0)$'s are a neighbourhood base at $\chi_0$.]

In fact

**Theorem.** (Pontrjagin duality) *If $G$ is a locally compact group, then $\widehat{G}$ is also locally compact. Furthermore $\widehat{\widehat{G}} = G$.*

*Thus, $\widehat{G}$ gives us our analogue of $\mathbb{Z}$ (resp. $\mathbb{R}^n$) for any LCA group.*

(Here is an example of an LCA group which is not $\mathbb{R}^n$, $\mathbb{T}^n$ or $\mathbb{Z}^n$: $G = \prod_{n=1}^{\infty}\{0,1\}$).

We can thus ask about the Fourier transform, defined by $\hat{f}(\chi) = \int_{\widehat{G}} f(x)\overline{\chi(x)}dx$.

**Theorem.** *Let $G$ be an LCA group*

a) *$\hat{}$ is a linear injection on $L^1$.*

b) *Let $f \in L^1(G)$. Then $\hat{f}$ is a continuous bounded function on $\widehat{G}$ which approaches zero at infinity.*

c) *Define for $f, g \in L^1(G)$ $f * g = \int_G f(y^{-1}x)g(y)dy$.*

*Then $(f * g)\hat{}(\chi) = \hat{f}(\chi)\hat{g}(\chi)$.*

**Note:** We do not say anything about derivatives here. You can't differentiate on an arbitrary LCA group unless it is also a Lie group. In this case it must have the form $\mathbb{T}^n \times \mathbb{R}^n$, or an extension of such a group by a discrete group.

**Proof.** Let's finally prove some parts of this theorem. For example, c)

$$(f * g)\hat{}(\chi) = \int \int_G f(y^{-1}x)g(y)dy\overline{\chi(x)}dx.$$

Replace $x$ with $yx$ since $dx$ is invariant.

We get

$$\begin{aligned}(f * g)\hat{}(x) &= \int_G \int_G f(x)g(y)\overline{\chi}(yx)dy\ dx \\ &= \int_G \int_G f(x)g(y)\overline{\chi}(x)\overline{\chi}(y)dy\ dx\end{aligned}$$

since $\chi$ is a character

$$= \hat{f}(x)\hat{g}(\chi).$$

a) is an exercise in the Stone-Weierstrass theorem and b) is a nice exercise in the definition of the topology on $\widehat{G}$. □

Actually, it turns out that a version of the Plancherel theorem also holds in this setting. One needs to define, as one did for $\mathbb{R}^n$, $\hat{f}(\chi)$ for $f \in L^2$.

**Theorem.** *For a suitable choice of constant in the Haar measure $d\chi$ on $\widehat{G}$, the Fourier transform is an isometry*

$$L^2(G) \to L^2(\widehat{G}).$$

This theorem generalizes the above theorems for $\mathbb{T}$ and $\mathbb{R}$, and enables one for example, to expand $L^2$ functions on $\prod\{0,1\}$ in orthogonal expansions known as Walsh series. But that's another story.

## CHAPTER 2: SPACES OF FUNCTIONS AND MEASURES

**2.1. Introduction.** In the case of the circle $\mathbb{T}$, the simplest kinds of functions are the trigonometric polynomials $T(\mathbb{T})$. These are finite linear combinations of exponentials of the form

$$f(x) = \sum_{n \in F} a_n e^{inx}, \quad \text{where } F \text{ is a finite set.}$$

The Fourier transform of such a function is easily seen to be $\hat{f}(n) = \begin{cases} a_n & n \in F \\ 0 & n \notin F. \end{cases}$

In fact, for any $f \in L^1(\mathbb{T})$, $S_N f$ belongs to $T(\mathbb{T})$; one of the themes of the last lecture was to see how to approximate $f$ by the trigonometric polynomial $S_N f$. Our first aim in this chapter will be to replace $S_N f$ by other trigonometric polynomials, obtaining better convergence. This leads naturally to the ideas of a summability kernel and an approximate identity.

The natural analogue of $T(\mathbb{T})$ in the case of $\mathbb{R}$ is the set $C_c(\mathbb{R})\hat{}$ of functions whose Fourier transforms have compact support. We study this space and approximate identities on $\mathbb{R}$.

We then give a description of how to extend the Fourier transform to other values of $p$, followed by a treatment of Fourier-Stieljes transforms and Bochner's theorem.

**2.2. Summability kernels on $\mathbb{T}$**

In our "proof" of Fourier's theorem, we noticed that on $\mathbb{T}$ $S_N f(x) = D_N * f(x)$, where $D_N(x) = \sum_{n=-N}^{N} e^{inx} = \frac{\sin(N+\frac{1}{2})x}{\sin\frac{1}{2}x}$. $D_N$ is called the Dirichlet kernel.

Probably the next best known kernel is the Fejér kernel, given by

$$\sigma_N f(x) = \sum_{n=-N}^{N} \left(1 - \frac{|n|}{N}\right) \hat{f}(n) e^{inx} = F_N * f(x)$$

where
$$F_N(x) = \sum_{n=-N}^{N} \left(1 - \frac{|n|}{N}\right) e^{inx} = \frac{1}{N+1} \left\{ \frac{\sin\left(\frac{N+1}{2}\right) x}{\sin\frac{1}{2}x} \right\}^2$$

Another kernel which we shall use is the Poisson kernel, defined by

$$P_r * f(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n) r^{|n|} e^{inx}$$

Here, $P_r(x) = \dfrac{1-r^2}{1-2r\cos x + r^2}$.

In fact, we define a **summability kernel** (also called an **approximate identity**) to be a sequence $\{k_n\}$ of $2\pi$-periodic functions satisfying

(1) $\displaystyle\int_{\mathbb{T}} k_n(x)dx = 1 \quad \forall n$

(2) $\displaystyle\int_{\mathbb{T}} |k_n(x)|dx \leq const$

(3) $\forall 0 < \delta < \pi$, $\lim_{n\to\infty} \int_{|x|>\delta} |k_n(x)|dx = 0$.

Given such a sequence of functions, one enquires whether $k_n * f \to f$ in $L^1, L^p$, etc. There is a standard answer to this. A Banach space $B$ contained in $L^1(\mathbb{T})$ is called **homogeneous** if its norm is translation-invariant (i.e. $f \in B \Rightarrow f_y : x \mapsto f(y^{-1}x) \in B\ \forall y \in \mathbb{T}$ and $\|f_y\|_B = \|f\|_B$) and translation is continuous (i.e. $\forall f \in B \lim_{y\to 0} \|f_y - f\|_B = 0$.)

The continuous functions and all $L^p$ spaces $1 \leq p < \infty$ have this property. The space $L^\infty$ is an exception as its norm is not translation continuous.

**Theorem.** *Let $\{k_n\}$ be a summability kernel and $B$ a homogeneous Banach space. Then $\forall f \in B\ \|k_n * f - f\|_B \to 0$ as $n \to \infty$.*

**Proof.**

$$\begin{aligned}
\|k_n * f - f\|_B &= \| \int k_n(y) f_y(\cdot) dy - \int k_n(y) f(\cdot) dy \|_B \\
&\leq \| \int_{-\delta}^{\delta} k_n(y)(f_y(\cdot) - f(\cdot)) dy \|_B + \| \int_{|y|>\delta} k_n(y)(f_y(\cdot) - f(\cdot) dy \|_B \\
&\leq \int_{-\delta}^{\delta} |k_n(y)|\ \|f_y(\cdot) - f(\cdot)\|_B dy + 2\|f\|_B \int_{|y|>\delta} |k_n(y)| dy
\end{aligned}$$

If $\epsilon > 0$, choose $\delta$ so small that $\|f_y - f\|_B < \frac{\epsilon}{2}$ for $|y| < \delta$, and then $n$ so large that $2\|f\|_B \int_{|y|>\delta} |k_n(y)| dy < \frac{\epsilon}{2}$.

We have $\|k_n * f - f\|_B < \epsilon$. □

It follows that convolution with a summability kernel yields $L^p$ convergence for $1 \leq p < \infty$.

The Dirichlet kernel is, unfortunately not a summability kernel, as $\|D_N\|_1 \sim \log N$. The Fejér and Poisson kernels are. In some senses, this is why $S_N f \to f$ fails in $L^1$. We do, however, have $\sigma_N f \to f$ in $L^1$. From this, the injectivity of the Fourier transform follows immediately as does the Riemann-Lebesgue lemma. The above theorem also implies that if $f$ is a continuous function, then $\sigma_N f \to f$ uniformly. We cannot in general deduce pointwise convergence from convergence in norm. However Fejér proved that if $f \in L^1$ and $\lim_{h\to 0} \left( \frac{f(x_0 + h) + f(x_0 - h)}{2} \right)$ exists (or is $\pm\infty$) then $\sigma_n f(x_0)$ does approach this limit. In particular, if $f$ is continuous at $x_0, \sigma_n f(x_0) \to f(x_0)$. Lebesgue showed that $\sigma_n f(x) \to f(x)$ almost everywhere.

## 2.3. Approximate identities on $\mathbb{R}^n$.

An approximate identity on $\mathbb{R}$ (or $\mathbb{R}^n$) is a family of continuous functions $\{k_\lambda\}$ with either continuous or discrete parameter $\lambda$ satisfying

(i) $\int k_\lambda(x)dx = 1$.

(ii) $\|k_\lambda\| = 0(1)$ as $\lambda \to \infty$.

(iii) $\lim\limits_{\lambda\to\infty} \int_{|x|>\delta} |k_\lambda(x)|dx = 0 \quad \forall\delta > 0$.

The most common way of producing approximate identies is to take a function $f \in L^1(\mathbb{R}^n)$ such that $\int f(x)dx = 1$ and set

$$k_\lambda(x) = \lambda^n f(\lambda x) \text{ for } \lambda > 0.$$

It is easy to see that conditions (i) - (ii) are satisfied, by making the change of variables $u = \lambda x$ in the integrals involved.

**Examples on $\mathbb{R}$:** The Fejér kernel $F_\lambda(x) = \lambda F(\lambda x)$ where

$$F(x) = \frac{1}{2\pi}\left(\frac{\sin x/2}{x/2}\right)^2 = \frac{1}{2\pi}\int_{-1}^{1}(1-|\xi|)e^{i\xi x}dx.$$

- The de la Vallée Poussin kernel $V_\lambda(x) = 2K_{2\lambda}(x) - K_\lambda(x)$
- The Poisson kernel $P_\lambda(x) = \lambda P(\lambda x)$ where $P(x) = \dfrac{1}{\pi(1+x^2)}$ $(\hat{P}(\xi) = e^{-|\xi|})$.
- The Gaussian kernel $G_\lambda(x) = \lambda G(\lambda x)$ where $G(x) = \dfrac{1}{\sqrt{\pi}}e^{-x^2}$ and $\widehat{G}(\xi) = e^{-\xi^2/4}$.

A re-write of the proof given in the previous section shows that $\|k_\lambda * f - f\|_1 \to 0$ as $\lambda \to \infty$. Hence, we can deduce injectivity of the Fourier transform, the Riemann-Lebesgue lemma and the interesting fact that the functions with compactly supported Fourier transforms are dense in $L^1$.

It is not entirely satisfactory in the noncompact case to define a homogeneous Banach space to be a subspace of $L^1$-as this excludes $L^p$ for $p \neq 1$, $C(\mathbb{R})$ etc. Actually, the convergence theorem holds for homogeneous Banach spaces of locally integrable functions, i.e. functions integrable on every compact subset of $\mathbb{R}^n$. $L^1_{loc}(\mathbb{R}^n) = \{f : f \text{ is measurable}$ and $\forall K \subseteq \mathbb{R}^n$ compact, $\int_K |f(x)|dx < \infty\}$.

**Theorem:** *If $B$ is a homogeneous Banach space of locally intergable functions, and if convergence in $B$ implies convergence in measure, then for any summability kernel $\{K_\lambda\}$*

$$K_\lambda * f \to f \quad \forall f \in B.$$

## 2.4. The Fourier transform on $L^p(\mathbb{R}^n)$.

In the previous chapter, we briefly alluded to the problems of defining the Fourier transform on $L^p$ where $p > 1$. Here, we give some of the details and crucial inequalities.

**Lemma.** *Let $f$ belong to $C_c(\mathbb{R}^n)$. Then*

$$\int |\hat{f}(\xi)|^2 d\xi = \frac{1}{(2\pi)} \int_n |f(x)|^2 dx$$

**Proof.** We'll just prove this for $n = 1$. Suppose first that the support of $f$ is contained in $(-\pi, \pi)$.

Then by the result for $\mathbb{T}$

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2 dx = \sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2.$$

Replace $f$ by $e^{-i\alpha x} f(x)$. We get

$$\int |f(x)|^2 dx = \int_{-\pi}^{\pi} |f(x)|^2 dx = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} |\hat{f}(n+\alpha)|^2.$$

Now integrate over $\alpha$, $0 \leq \alpha \leq 1$ to obtain

$$\int |f(x)|^2 dx = \int |\hat{f}(\xi)|^2 d\xi.$$

If the support of $f$ is not included in $(-\pi, \pi)$, we replace $f$ by $g(x) = \lambda^{1/2} f(\lambda x)$. If $\lambda$ is sufficiently large, the support of $g$ is included in $(-\pi, \pi)$ and $\hat{g}(\xi) = \lambda^{-1/2} \hat{f}(\xi/\lambda)$. Substituting this in, we obtain the desired result. □

Since the functions of compact support are dense in $L^1 \cap L^2$, this lemma shows that we have

$$\|f\|_2 = \|\hat{f}\|_2 \text{ for all } f \in L^1 \cap L^2.$$

Since $L^1 \cap L^2$ is dense in $L^2$, we can now extend $\hat{}$ to an isometry of $L^2(\mathbb{R}^n)$ to $L^2(\widehat{\mathbb{R}}^n)$.

If $1 < p < 2$, we require the Hausdorff-Young inequality:

$$\left( \int |\hat{f}(\xi)|^q d\xi \right)^{1/q} \leq \left( \int |f(x)|^p \right)^{1/p}$$

for $f \in L^1 \cap L^2$. (Here $\frac{1}{p} + \frac{1}{q} = 1$).

Given this inequality, we can define the Fourier transform of $f \in L^p(\mathbb{R}^n)$ as an element of $L^{p'}(\mathbb{R}^n)$.

How does one prove the Hausdorff-Young inequality? By interpolation! We know the two inequalities for $f \in L^1 \cap L^2$ :

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

$$\text{and } \sup|\hat{f}(\xi)| = \|\hat{f}\|_\infty \leq \|f\|_1.$$

The Riesz-Thorin convexity theorem may be stated as follows:

**Theorem.** Let $(X, d\mu)$ and $(Y, d\nu)$ be two measure spaces. Let

$$B = L^{p_0} \cap L^{p_1}(d\mu) \text{ and } B' = L^{p'_0} \cap L^{p'_1}(d\nu)$$

and let $S$ be a linear transformation from $B$ to $B'$, continuous from $L^{p_0}$ to $L^{p'_0}$ and from $L^{p_1}$ to $L^{p'_1}$.

For $0 < \alpha < 1$, let

$$p_\alpha = \frac{p_0 p_1}{p_0\alpha + p_1(1-\alpha)} \quad \text{(similarly for } p'_\alpha).$$

Then $S$ is continuous from $L^{p_\alpha}$ to $L^{p'_\alpha}$ and $\|S\|_{p_\alpha, p_{\alpha'}} \leq \|S\|^{1-\alpha}_{p_0, p'_0} \|S\|^{\alpha}_{p_1, p'_1}$.

The proof of this theorem would take us rather far abroad and I have decided to omit it.

Note that the Fourier transform is not onto for $p < 2$. Indeed, the problem of deciding which elements of $L^q (q > 2)$ are Fourier transforms of something in $L^p$ is extremely difficult.

Unfortunately, the Hausdorff-Young theorem fails for $p > 2$, and indeed one can show that there is no homogeneous Banach space $B$ on $\mathbb{R}$ so that for some $p > 2$ and for some constant $C$, $\|\hat{f}\|_B \leq C\|f\|_p$.

We return to the problem of $L^p, p > 2$ next chapter.

The entire content of this section holds for LCA groups.

**2.5. Fourier-Stieltjes transforms.**

Another extension of the Fourier transform is to spaces of Borel measures instead of functions.

In the case of the circle group (or indeed, any compact abelian group), the space of finite Borel measures may be defined as the dual space of $C(\mathbb{T})$, the continuous functions on $G$, i.e. the continuous linear functionals on $C(\mathbb{T})$

$$M(\mathbb{T}) = C(\mathbb{T})^*.$$

You might be more familiar with the notion of a measure as a complex-valued functions on the Borel sets. Of course, given any such function, one can define

$$\langle \mu, f\rangle = \int_{\mathbb{T}} f(x)d\mu(x)$$

for every continuous function $f$. Conversely, given an element of $C(\mathbb{T})^*$, the set function $\mu(E)$ may be constructed by approximating the characteristic function of $E$ by continuous functions and taking limits.

In the case of $\mathbb{R}$, $\mathbb{R}^n$ or any LCA group the finite Borel measures are defined to be the elements of the dual of $C_0(\widehat{G})$ – the continuous functions which vanish at infinity. Again, these are in one-one correspondence with finite-valued set functions on the Borel $\sigma$-algebra in the usual way.

$L^1(G)$ is contained in $M(G)$, because for each integrable function $f$, and for each $g \in C_0(G)$ $\langle \mu_f, g\rangle = \int_G f(x)g(x)dx < \infty$.

$M(G)$ is naturally equipped with a norm, called the total variation norm, defined by

$$\|\mu\| = \int d|\mu| = |\mu|(G)$$

and on $L^1$ this coincides with the $L^1$-norm, i.e. $\|\mu_f\| = \|f\|_1$.

The convolution of a measure and a function in $C_0$ may be defined by

$$\mu * \varphi(x) = \int \varphi(y^{-1}x)d\mu(y).$$

Hence, we may define the convolution $\mu * \nu$ of two measures by $\langle \mu * \nu,\ \varphi\rangle = \langle \nu, \mu * \varphi\rangle$, or directly by $(\mu * \nu)(E) = \int \mu(y^{-1}E)d\nu(y)$. One shows that this convolution corresponds to $L^1$-convolution and that $\|\mu * \nu\| \leq \|\mu\|\ \|\nu\|$.

We define the Fourier-Stieljes transform of a measure $\mu \in M(\mathbb{R}^n)$ by

$$\hat{\mu}(\xi) = \int e^{-i\xi x}d\mu(x) \quad \xi \in \mathbb{R}^n.$$

Clearly if $\mu = \mu_f$ then $\hat{\mu}_f(\xi) = \hat{f}(\xi)$. Many of the properties of Fourier transforms are shared by Fourier-Stieljes transforms e.g. $(\mu * \nu)\hat{} = \hat{\mu}\hat{\nu}$ and $|\hat{\mu}(\xi)| \leq \|\mu\|$. Furthermore, $\hat{\mu}$ is uniformly continuous. One important difference is that the Riemann-Lebesgue lemma fails: $\hat{\mu}(\xi)$ will not necessarily vanish at $\infty$. The easiest example to see this from is $\delta_0$, defined by $\delta_0(E) = \begin{cases} 1 & \text{if } 0 \in E \\ 0 & \text{otherwise} \end{cases}$. Then $\hat{\delta}_0(\xi) = 1$ for all $\xi$.

There is a version of Plancherel's theorem.

**Theorem.** *Let $\mu \in M(\mathbb{R})$ and $f$ be a continuous function in $L^1(\mathbb{R})$ so that $\hat{f} \in L^1(\widehat{\mathbb{R}})$. Then*

$$\int f(x)d\mu(x) = \frac{1}{2\pi}\int \hat{f}(\xi)\hat{\mu}(-\xi)d\xi.$$

**Proof.**

$$f(x) = \frac{1}{2\pi}\int \hat{f}(\xi)e^{i\xi x}d\xi$$

so

$$\int f(x)d\mu(x) = \frac{1}{2\pi}\int\int \hat{f}(\xi)e^{i\xi x}d\xi d\mu(x)$$
$$= \frac{1}{2\pi}\int \hat{f}(\xi)\hat{\mu}(-\xi)d\xi.$$

The use of Fubini's theorem is justified by the fact that $\hat{f}(\xi) \in L^1(\widehat{\mathbb{R}})$. □

An application of this Plancherel theorem which will be useful later is:

**Corollary.** *A function $\varphi$, defined and continuous on $\widehat{\mathbb{R}}$, is a Fourier-Stieljes transform iff $\exists C$ such that*

$$\left|\int \hat{f}(\xi)\varphi(-\xi)d\xi\right| \leq C\|f\|_\infty$$

*for all $f$ continuous, belonging to $L^1(\mathbb{R}^n)$ and such that $\hat{f}$ has compact support.*

There is a characterization of which functions are Fourier-Stieljes transforms of positive measures, due to Bochner.

A function $\varphi$ on $\widehat{\mathbb{R}}$ is positive definite if for every choice of $\xi_1, \ldots, \xi_N \in \widehat{\mathbb{R}}$ and complex numbers $z_1, \ldots z_N$

$$\sum_{j,k=1}^{N} \varphi(\xi_j - \xi_k)z_j\bar{z}_k \geq 0.$$

[Each such function satisfies $\varphi(-\xi) = \overline{\varphi(\xi)}$ and $|\varphi(\xi)| \leq \varphi(0)$.]

**Theorem** (Bochner): *A function $\varphi$ defined on $\widehat{\mathbb{R}}$ is a Fourier-Stieljes transform of a positive measure iff it is positive definite and continuous.*

**Proof.** If $\varphi = \hat{\mu}$ with $\mu \geq 0$

$$\sum_{j,k} \varphi(\xi_j - \xi_k) z_j \bar{z}_k = \int \sum e^{-i\xi_j x} z_j e^{i\xi_k x} \bar{z}_k d\mu(x)$$

$$= \int \left| \sum_{j=1}^{N} z_j e^{-i\xi_j x} \right|^2 d\mu(x) \geq 0.$$

The other direction involves a reduction to $\mathbb{T}$ and I will omit it. □

This theorem also holds for any LCA group $G$. In this setting it is a key step in the proof of the Pontrjagin duality theorem.

**2.6. The Poisson summation formula.**

This formula gives a very pretty relationship between the Fourier transforms on $\mathbb{T}$ and on $\mathbb{R}$. While it doesn't really "fit" here as regards didactic progression, I cannot resist mentioning it, as it has some profound consquences.

Let $f \in L^1(\mathbb{R})$ and define $\varphi$ by $\varphi(t) = 2\pi \sum_{j=-\infty}^{\infty} f(t + 2\pi j)$.

Clearly $\varphi$ is $2\pi$-periodic so we may consider $\varphi$ as defined on $\mathbb{T}$. In fact $\varphi \in L^1(\mathbb{T})$ and $\|\varphi\|_1 \leq \|f\|_1$.

If $n \in \mathbb{Z}$,

$$\hat{\varphi}(n) = \frac{1}{2\pi} \int_0^{2\pi} \varphi(t) e^{-int} dt = \sum_{j=-\infty}^{\infty} \int_0^{2\pi} f(t + 2\pi j) e^{-int} dt$$

$$= \int_{-\infty}^{\infty} f(x) e^{-inx} dx = \hat{f}(n).$$

Thus $\hat{\varphi}$ is the restriction of $\hat{f}$ to $\mathbb{Z}$.

Since $\varphi(0) = \sum_{n \in \mathbb{Z}} \hat{\varphi}(n)$, we obtain the famous Poisson summation formula

$$2\pi \sum_{n=-\infty}^{\infty} f(2\pi n) = \sum_{n=-\infty}^{\infty} \hat{f}(n).$$

(In fact, replacing $f$ by $\lambda f(\lambda x)$, we get

$$2\pi \sum_{n=-\infty}^{\infty} f(2\pi \lambda n) = \sum_{n=-\infty}^{\infty} \hat{f}\left(\frac{n}{\lambda}\right).)$$

# CHAPTER 3: DISTRIBUTIONS

## 3.1. Introduction.

The theory of this chapter sits squarely in $\mathbb{R}^n$. It is also valid on certain Lie groups ... see Ch. 6.

Here we set up the theory necessary to "solve" arbitrary constant coefficient differential operators. (The actual solution will be in Ch. 5).

Distributions were initially introduced by physicists in some sense. The famous Dirac $\delta$ function which is zero everywhere except at $x = 0$ where its value is infinite – and which has integral 1 is the first example. How can we make sense of this and of its Fourier transform? As a result we will be able to find Fourier transforms of $L^p$ functions for $p > 2$.

The idea is to follow the same idea as for measures, but to replace the continuous functions with the $C_c^\infty$ functions. But these do not have a norm.

## 3.2. The space $\mathbf{C}^\infty(\Omega)$.

If $\alpha = (\alpha_1, \ldots, \alpha_n)$ is an $n$-tuple of nonnegative integers and $x \in \mathbb{R}^n$, let $x^\alpha$ denote $(x_1^{\alpha_1}, \ldots, x_n^{\alpha_n})$. Similarly associate to $\alpha$ the differential operator

$$D^\alpha = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n} \quad \text{of order } |\alpha| = \alpha_1 + \ldots + \alpha_n.$$

To $f$ defined on an open set $\Omega \subseteq \mathbb{R}^n$, we associate $D^\alpha f$. We say $f \in C^\infty(\Omega)$ if $D^\alpha f$ is continuous for all $\alpha$.

We define a topology on $C^\infty(\Omega)$ as follows: For each compact $K \subseteq \Omega$, let

$$P_{K,N}(f) = \sup\{|D^\alpha f(x)| : x \in K \text{ and } |\alpha| \leq N\}.$$

Then $P_{K,N}$ is a semi-norm (a metric where the property "$P_{K,N}(f) = 0 \Rightarrow f = 0$" fails).

Let $V(K, N, \epsilon) = \{f : P_{K,N}(f) < \epsilon\}$ and take the open neighbourhoods of 0 in $C^\infty(\Omega)$ to be the sets $V(K, N, \epsilon)$ as $K$ ranges over compact subsets of $\Omega, N$ ranges over $\mathbb{N}$, and $\epsilon > 0$.

**Theorem.** (1) *The topology is countably generated.*

(2) $C^\infty(\Omega)$ *is a locally convex topological vector space.*

(3) *$C^\infty(\Omega)$ has a metric under which it is complete (so it is a Fréchet space).*

*For each $K \subseteq \Omega$, let*

$$\mathcal{D}_K = \{f \in C^\infty(\Omega) : f \text{ is supported on } K\}$$

*and $\mathcal{D}(\Omega) = \cup_{K \text{ compact } \subseteq \Omega} \mathcal{D}_K$. Let $\|\phi\|_N = \sup\{|D^\alpha \phi(x)| : x \in \Omega,\ |\alpha| \leq N\}$.*

**Definition.** A **distribution** is a linear map $\Lambda : \mathcal{D} \to \mathbb{C}$ s.t. $\forall$ compact $K \subseteq \Omega$ $\exists$ an integer $N$ and a constant $C$

$$\text{s.t. } |\Lambda\phi| \leq C\|\phi\|_N \quad \forall \phi \in \mathcal{D}_K.$$

The space of distributions is called $\mathcal{D}'(\Omega)$.

If the same $N$ will do for all $K$ and is the smallest with this property, we say that $\Lambda$ has order $N$.

(If no such $N$ exists, $\Lambda$ has infinite order.)

**Examples** (1) Let $x \in \Omega$ and set $\delta_x \phi = \phi(x)$. Then $\delta_x$ is a distribution of order zero.

(2) A distribution of order zero is a Radon measure.

(3) Let $f$ be a complex-valued locally integrable function on $\Omega$ (i.e. $\forall$ $K$ compact $\subseteq \Omega$, $\int_K |f(x)|dx < \infty$).

Let $\Lambda_f \phi = \int_K f(x)\phi(x)dx \quad \forall \phi \in \mathcal{D}_K$.

Then $\Lambda_f$ is a distribution of order 0.

(4) The derivative of the delta function is a distribution of order 1.

**Theorem.** There is a topology on $\mathcal{D}(\Omega)$ which makes it into a LCTVS and such that $\mathcal{D}'(\Omega)$ is precisely the set of continuous linear functionals on $\mathcal{D}(\Omega)$.

$\mathcal{D}(\Omega)$ is complete in this topology although not metrizable.

[The definition is rather grotty and you don't need to see it for this course. Look it up in Rudin's Functional Analysis if you wish to.]

### 3.3. Properties of distributions.

- One can differentiate distributions: if $\Lambda \in \mathcal{D}'$ and $\alpha$ is a multi-index, let

$$(D^\alpha \Lambda)\phi = (-1)^{|\alpha|}\Lambda(D^\alpha \phi).$$

(Note that if $f$ is locally integrable and differentiable of order $|\alpha|$ then

$$D^\alpha \Lambda_f = \Lambda_{D^\alpha f}.)$$

- One can multiply distributions by functions: Let $\Lambda \in \mathcal{D}'(\Omega)$ and $f \in C^\infty(\Omega)$. Define $f\Lambda$ by $(f\Lambda)(\phi) = \Lambda(f\phi)$.

  Then $f\Lambda \in \mathcal{D}'$.
- One can take limits of distributions. The **weak* topology** on $\mathcal{D}'$ is defined as follows: $\Lambda_i \to \Lambda$ if $\forall \phi \in \mathcal{D},\ \Lambda_i\phi \to \Lambda\phi$.

For example, if $f_i$ is locally integrable then $f_i \to \Lambda$ in the distribution sense means $\int f_i\phi \to \Lambda\phi \quad \forall \phi \in \mathcal{D}$.

Thus for example, if $f_i$ is an approximate identity, we have $f_i \to \delta_0$.

Limits behave rather pleasantly. For example, suppose $\Lambda_i \in \mathcal{D}'(\Omega)$ and $\forall \phi \in \mathcal{D}(\Omega)$, $\lim_{i\to\infty} \Lambda_i\phi$ exists as a complex number. Call the limit $\Lambda\phi$. Then $\Lambda \in \mathcal{D}'(\Omega)$ and for all multi-indexes $\alpha$, $D^\alpha\Lambda_i \to D^\alpha\Lambda$.

- The support of a distribution is defined as follows: Let $\Omega_1 \subseteq \Omega$ be an open set. Say $\Lambda$ **vanishes on** $\Omega_1$ if $\Lambda\phi = 0$ whenever $\phi \in \mathcal{D}(\Omega_1)$. The support of $\Lambda$ is the complement of the largest $\Omega_1$ such that $\Lambda$ vanishes on $\Omega_1$. It is denoted $\mathrm{supp}\Lambda$.

The kinds of things you would expect hold. A distribution of support 0 is zero. If $\mathrm{supp}\ \phi \cap \mathrm{supp}\Lambda = \emptyset$ then $\Lambda\phi = 0$, etc.

Of particular interest are distributions of compact support. These are denoted $\mathcal{E}(\Omega)$. It can be seen that a distribution of compact support has finite order. In fact, it satisfies

$$|\Lambda\phi| \leq C\|\phi\|_N \text{ for some } C, N,\ \forall \phi \in \mathcal{D}(\Omega).$$

Thus, it extends in a unique way to a linear functional on $C^\infty(\Omega)$.

Every distribution is a combination of derivatives of continuous functions. In fact, we have the following:

**Structure Theorem.** *Let $\Lambda \in \mathcal{D}'$. For each $\alpha$, there is a continuous function $g_\alpha$ on $\Omega$ such that*

(a) *Each compact subset $K$ of $\Omega$ intersects the support of only finitely many $g_\alpha$.*

(b) $\Lambda = \sum_\alpha D^\alpha g_\alpha$.

*If $\Lambda$ has finite order then the $g_\alpha$ can be chosen so that only finitely many are nonzero.*

- The formula

$$\Lambda * \phi(x) = \Lambda((\phi^\nu)_x)$$

$\Lambda \in \mathcal{D}$, $\phi \in \mathcal{D}$ and $\phi^\nu(y) = \phi(-y)$, defines convolution of a distribution and a function in $\mathcal{D}$.

One shows easily that $\mathcal{D}^\alpha(\Lambda * \phi) = (D^\alpha \Lambda) * \phi = \Lambda * D^\alpha \phi$.

In fact, one can show that if $\Lambda \in \mathcal{D}'$, $\phi \in \mathcal{D}$ then $\Lambda * \phi \in C^\infty$ (though **not** of compact support).

One can convolve a distribution of compact support with an element of $C^\infty$, but **not** an arbitrary distribution. Thus if either $\Lambda_1$ or $\Lambda_2$ has compact support we can define $\Lambda_3$ by

$$\Lambda_3 \phi = \Lambda_1 * (\Lambda_2 * \phi)(0), \quad \phi \in \mathcal{D}$$

and it can be shown that $\Lambda_3 \in \mathcal{D}'$, and we set $\Lambda_3 = \Lambda_1 * \Lambda_2$.

There is however no notion of convolution for arbitrary distributions (and it can be shown that this cannot exist in some senses).

### 3.4. Tempered distributions.

To get Fourier transforms to work, we need to decrease somewhat our space of distributions (or equivalently increase our space of test functions.)

**Definitions.** We say that $f \in C^\infty(\mathbb{R}^n)$ is **rapidly decreasing** if $\forall N \in \mathbb{N}$, $\forall \alpha, \exists C_N^\alpha$ s.t.

$$\sup_{x \in \mathbb{R}^n} |D^\alpha f(x)| \leq C_N^\alpha (1 + |x|^2)^{-N}.$$

The rapidly decreasing functions are denoted by $\mathcal{S}(\mathbb{R}^n)$. Clearly $\mathcal{D}(\mathbb{R}^n) \subseteq \mathcal{S}(\mathbb{R}^n)$ but they are not equal – for example $e^{-x^2} \in \mathcal{S}(\mathbb{R})$. If $f \in \mathcal{S}(\mathbb{R}^n)$ then for any polynomial $P(x)$, $P(x)f(x) \in L^1$, and $D^\alpha f \in \mathcal{S}$ for all $\alpha$.

The smallest possible value of $C_N^\alpha$ in the above inequality gives a seminorm. These seminorms define a topology on $\mathcal{S}$ under which it becomes a Fréchet space.

$\mathcal{D}$ is dense in $\mathcal{S}$ and $i : \mathcal{D} \to \mathcal{S}$ is continuous.

**Theorem.** $\hat{}: \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ *is an isomporhism.*

The proof is based upon the fact that the Fourier transform changes differentiation into multiplication. It also uses some rather neat approximate identity ideas.

Since $i : \mathcal{D} \to \mathcal{S}$ is continuous, we get $i' : \mathcal{S}' \to \mathcal{D}'$ is continuous.

Thus $\mathcal{S}' \subseteq \mathcal{D}'$ is a certain set of distributions called **tempered distributions.**

Any distribution of compact support is tempered. If $1 \leq p < \infty$, $N > 0$ and $g$ is measurable on $\mathbb{R}^n$ such that

$$\int_{\mathbb{R}^n} |(1+|x|^2)^{-N} g(x)|^p dx = C < \infty$$

then $\Lambda_g$ is a tempered distribution. Thus, the spaces $L^p$, $p > 2$ are tempered distributions.

If $\Lambda$ is tempered, so is $\phi\Lambda$, $D^\alpha \Lambda$ and $P\Lambda$.

The Hilbert transform $H$ on $\mathbb{R}$ defined by

$$Hf(x) = p.v. \int_{-\infty}^{\infty} \frac{f(x)}{x-t} dt$$

is tempered.

We define the Fourier transform of $\Lambda \in \mathcal{S}'$ by $\Lambda\hat{}(\phi) = \Lambda(\hat{\phi})$.

**Theorem** (i)$\hat{}$ is a continuous linear isomorphism of $\mathcal{S}'$ to $\mathcal{S}'$ whose inverse is also continuous

(ii) If $u \in \mathcal{S}'$ and $P$ is a polynomial, then

$$(P(D)u)\hat{} = P\hat{u} \text{ and } (Pu)\hat{} = P(-D)\hat{u}$$

### 3.5. Paley-Wiener Theorems.

The idea of this is that the Fourier transform of a distribution of compact support is actually an entire function of $z$, where we embed $\mathbb{R}^n$ in $\mathbb{C}^n$.

First, the classical Paley-Wiener theorem.

**Theorem A** (Paley Weiner) (a) *Suppose $\phi \in \mathcal{D}(\mathbb{R}^n)$ is supported in $rB$, $B$ being the unit ball of $\mathbb{R}^n$.*

*Let $f(z) = \int_{\mathbb{R}^n} \phi(t) e^{-iz.t} dt \quad (z \in \mathbb{C}^n)$.* $\quad (*)$

*Then $f$ is entire and there exist constants $\gamma_N$ such that*

$$|f(z)| \leq \gamma_N (1+|z|)^{-N} e^{r|Im\ z|} (N \in \{0,1,2,\ldots\}) \qquad (**)$$

(b) *If $f$ satisfies $(**)$ there exists $\phi \in \mathcal{D}(\mathbb{R}^n)$ supported in $rB$ such that $(*)$ holds.*

**Proof.** (a) is reasonably straightforward.

(b) Define $\phi(t) = \int_{\mathbb{R}^n} f(x)e^{it.x}dx \quad (t \in \mathbb{R}^n)$ and use (**) to see that $\phi \in C^\infty(\mathbb{R}^n)$. Then work hard to show that $\phi$ is actually supported in $rB$. $\square$

This theorem also extends to distributions of compact support:

**Theorem B.** (Paley Wiener)

(a) *Suppose $\Lambda \in \mathcal{D}'$ is supported in $rB$ and has order $N$. Set $f(z) = \Lambda(e^{iz})$, $z \in C^n$. Then $f(z)$ is an entire function and $f\big|_{\mathbb{R}^n}$ is the Fourier transform of $\Lambda$. Further, there exists $C$ such that*

$$|f(z)| \leq C(1 + |z|^N)e^{r|Im\ z|} \tag{$*$}$$

*(b) Conversely, if $f(z)$ is an entire function satisfying $(*)$ then $\exists\Lambda \in \mathcal{D}'$, supported in $rB$ and of order $N$ such that $f(z) = \Lambda(e^{iz})$.*

## CHAPTER 4: THE HILBERT TRANSFORM, $H^p$ AND BMO

### 4.1. Introduction.

We first introduce a particular distribution known as the Hilbert transform on $\mathbb{T}$. This is intimately connected with the process of extending functions on $\mathbb{T}$ to functions analytic on the unit disc via the Poisson kernel. The M. Riesz theorem says that the Hilbert transform is bounded on $L^p$ for $1 < p < \infty$ (though not on $L^1$). This motivates the introduction of the Hardy spaces $H^p$ for $0 < p < \infty$, which coincide with $L^p$ for $1 < p < \infty$. The space $H^1$ is interesting as it provides an alternative to $L^1$ which has nice properties. Fefferman has identified the dual of $H^1$ as the functions of bounded mean oscillation (BMO). All this extends to $\mathbb{R}^n$ by replacing the disc with an upper half space, and is connected to the theory of singular integrals.

### 4.2. The Hilbert transform and the Poisson integral.

Recall that convolution with an $L^1$ function maps $L^p$ to $L^p(1 \leq p \leq \infty)$ since $||f*g||_p \leq ||f||_1||g||_p$. Furthermore, on the Fourier transform side, the operation of convolving with an $L^1$ function is simply a multiplication

$$(f * g)\hat{}(x) = \hat{f}(\chi)\hat{g}(\chi) \qquad (\text{on } \mathbb{R}^n \text{ or } \mathbb{T})$$

This leads to the concept of a multiplier – an operator that maps $L^p$ to $L^p$ (or $L^p$ to $L^q$) by a multiplication operator on the Fourier transform side.

Let $\phi$ be a complex-valued function on $\widehat{\mathbb{R}}^n$ (or $\mathbb{Z}$).

We say that $\phi$ is an **$\mathbf{L^p}$-multiplier** if $f \in L^p \Rightarrow \exists g \in L^p$ with

$$\hat{g}(\xi) = \phi(\xi)\hat{f}(\xi) \qquad (\forall \xi \in \widehat{\mathbb{R}}^n \text{ or } \forall \xi \in \mathbb{Z}).$$

The multiplier norm is given by $|||\phi|||_p = \sup_{\|f\|_p=1} \|\phi f\|_p$.

We can think of $\phi$ as a kind of generalized convolution operator. Many distributions given by singular integrals are $L^p$-multipliers. Still others are $L^p$-$L^q$-multipliers.

An example that we shall study in this chapter is the Hilbert transform, defined on $\mathbb{T}$ or $\mathbb{R}$ by

$$(Hf)\hat{}(\xi) = -isgn\xi . \hat{f}(\xi),$$

where $sgn\xi$ is the sign of $\xi$ ($sgn0 = 0$).

**Remark:** $Hf$ is an $L^p$ multiplier if and only if $S_N f \to f$ in $L^p$.

There are at least two distinct ways of studying $H$, via complex analysis, or via real-variable methods, and we shall discuss each in turn.

Let us now fix our attention on $\mathbb{T}$, which we consider as the boundary of the unit disc $D = \{z : |z| \leq 1\}$.

Let $P_r = \frac{1-r^2}{1-2r\cos t + r^2} (0 \leq r < 1)$ be the Poisson kernel, and for $f \in L^1(\mathbb{T})$, define $F$ on $D$ by $F(re^{it}) = \Sigma \hat{f}(n) r^{|n|} e^{int} = P_r * f(t)$. One shows relatively easily that $F$ is harmonic on $D$ and that for almost every $t \in \mathbb{T}$, $F(re^{it}) \to f(t)$ as $r \nearrow 1$. In fact, every bounded harmonic function on $D$ is the Poisson integral of a bounded function on $\mathbb{T}$.

Given a function $F$, let $\widetilde{F}$ be its conjugate harmonic function – i.e. $F$ and $\widetilde{F}$ satisfy the Cauchy-Riemann equations.

In fact, given $f \in L^1(\mathbb{T})$, one passes to $F$ on $D$, takes the conjugate $\widetilde{F}$, and shows that $\tilde{f}(t) = \lim_{r \to 1} \widetilde{F}(re^{it})$ is the Hilbert transform $Hf$ of $f$.

### 4.3. Functions of weak type $\mathbf{L^p}$. The M. Riesz theorem.

Let me recall the definition of weak $L^p$. To a real-valued measurable function $f$ on $\mathbb{T}$ we associate the distribution function, defined for $-\infty < x < \infty$ by

$$m_f(x) = \lambda\{t : f(t) \leq x\}$$

where $\lambda$ denotes Haar measure. Then $m_f : \mathbb{R} \to \mathbb{R}$ is a right continuous monotone function, which defines a Stieljes measure $dm_f$ on $\mathbb{R}$ by $dm_f[a, b) = m_f(b) - m_f(a)$ on intervals. (This extends to a measure.)

The basic property of $m_f$ is that for any continuous function $h$ on $\mathbb{R}$, $\int_{\mathbb{T}} h(f(t))dt = \int_{-\infty}^{\infty} h(x)dm_f(x)$.

A measurable function $f$ on $\mathbb{T}$ is said to be of weak type $L^p$ for some $0 < p < \infty$ if $\exists C$ such that $\forall \mu \in \mathbb{R}^+$

$$\lambda\{t : |f(t)| \geq \mu\} \leq C\mu^{-p}.$$

A simple calculation involving $m_{|f|}$ enables one to see that if $f \in L^p$, $f$ is in weak $L^p$. The converse is not true, as may be seen by examining the function $f(t) = |\sin t|^{-1/p}$.

In fact, it is not hard to show that if $f$ belongs to weak $L^p$, then $f \in \cap_{r<p} L^r$.

**Theorem.** *If $f \in L^1$ then $Hf$ is in weak $L^1$.*

**Theorem.** (M. Riesz) *$H$ is an $L^p$ multiplier for $1 < p < \infty$.*

The proof of both theorems may be done by complex analytic techniques.

### 4.4. Hardy space.

A corollary of M. Riesz' theorem is that if $f \in L^p$, $1 < p < \infty$, and if

$$H(z) = F(z) + i\widetilde{F}(z)$$

then for all $r \in [0, 1)$

$$\left(\int_0^1 |H(re^{it})|^p dt\right)^{1/p} \leq C\|f\|_p.$$

We define, for $H$ analytic on $D$, $0 \leq r < 1$, and $0 < p < \infty$,

$$\mu_p(H, r) = \left(\int |H(re^{it})|^p dt\right)^{1/p}.$$

We say that $H$ belongs to the Hardy space $H^p$ if

$$\sup_{0 \leq r < 1} \mu_p(H, r) = \|H\|_{H^p} \leq \infty.$$

Then $\|\cdot\|_{H^p}$ is a metric on $H^p$ under which $H^p$ is complete. It is a norm for $1 \leq p \leq \infty$. Thus $H^p$ is a Banach space for $1 \leq p < \infty$.

It is a very interesting space for $0 < p < 1$, when $||.||_{H^p}^p$ satisfies the triangle inequality, and is homogeneous of degree $p$, but this space of distributions is not a Banach space.

In fact by M. Riesz, for $p > 1$, the mapping $f \mapsto H$ is a bounded invertible mapping from $L^p$ into $H^p$ and so $H^p$ is identifiable with the closed subspace of $L^p$ consisting of all functions whose negative Fourier coefficients vanish. In fact $L^p = H^p \oplus \overline{H^p}$.

The space $H^1$ is of particular interest, as it provides a substitute for $L^1$ which shares many of the pleasant properties of $L^p$ for $p > 1$. In fact, if $f \in L^1(\mathbb{T})$ and $\tilde{f} \in L^1(\mathbb{T})$ then $F + i\widetilde{F} \in H^1$ and $\widetilde{F}$ is the Poisson integral of $\tilde{f}$. Each element of $H^1$ is the Poisson integral of its boundary value. From these considerations it can be shown that a function $F$ belongs to $H^p$ if and only if it is the Poisson integral of some $f \in L^p(\mathbb{T})$ with $\hat{f}(n) = 0 \ \forall n < 0$ (for $p \geq 1$).

These are the complex Hardy spaces. The real Hardy spaces, studied by Stein, Fefferman et al consist of $Re(H^p)$. It is not hard to show that any real-valued element of $L^p(\mathbb{T})$ is the boundary value of $Re(H)$ for some $H \in H^p$.

**4.5. Hardy spaces on $\mathbb{R}^n$.$(\mathbf{H}^1)^* = \mathbf{BMO}$.**

The preceding discussion of Hardy spaces in the disc may be generalized to the upper half plane,

$$\mathbb{R}_+^{n+1} = \{(x,t) : x \in \mathbb{R}^n, t > 0\}.$$

In fact, given a measurable function $f$ in $L^p(\mathbb{R}^n)$, let $u = f * P_t$, where $P_t$ is the Poisson kernel

$$P_t(x) = \frac{c_n}{(1 + |x|^2)^{-(n+1)/2}}, \quad c_n = \frac{\Gamma(\frac{n+1}{2})}{\pi^{(n+1)/2}}.$$

and define $u^*(x) = \sup_{|x-y|<t} |u(y,t)|$.

Define $H^p(\mathbb{R}^n)$ to be the set of measurable functions $f$ so that

$$(f * P_t)^* \in L^p.$$

Then one shows that if $u$ is harmonic in $\mathbb{R}_+^{n+1}$ then $u^* \in L^p(\mathbb{R}^n)$ iff $u$ is the Poisson integral of an $f$ in $H^p$ and $||u^*||_{L^p} \simeq ||f||_{H^p}$.

Again, these spaces are complete with the metric $d(f,g) = ||f - g||_{H^p}^p$ for $p \leq 1$, and they are isometric to $L^p$ for $p > 1$. If $f \in H^p$ and $\phi \in \mathcal{S}$, with $\int \phi dx = 1$ then

$\|f * \phi_\lambda\|_{H^p} \leq c\|f\|_{H^p}$ and $f * \phi_\lambda \longrightarrow f$ in $H^p$ as $\lambda \longrightarrow \infty$. For $p \leq 1$, the Fourier transform of $H^p$ functions are continuous on $\mathbb{R}^n$.

Fefferman has identified the dual of $H^1(\mathbb{R}^n)$. It is the space BMO. A locally integrable function $f$ belongs to BMO if there is a constant $A$ so that for all balls $B$,

$$\frac{1}{|B|}\int_B |f(x) - f_B| dx \leq A$$

Here, $f_B = \frac{1}{|B|}\int_B f\ dx$ is the average value of $f$ over $B$.

All bounded functions belong to BMO. The converse is not true, though one can show for example that if $f \in BMO$, $|f(x)|(1+|x|)^{-n-1} \in L^1(\mathbb{R}^n)$.

Fefferman's result is that every continuous linear functional $\ell$ on $H^1$ can be written as

$$\ell(g) = \int_{\mathbb{R}^n} f(x)g(x)dx \qquad (g \in H^1)$$

where $f$ is in BMO.

### 4.6. Singular integrals.

I mentioned in section 2 that there was another approach to the Hilbert transform and associated distributions – that is the method of singular integrals.

In fact, taking the inverse Fourier transform of $-isgn$ we obtain $\frac{1}{x}$ (ignoring niceties of convergence!) Thus we should expect $Hf = \frac{1}{x}$ "$*$" $f$ whatever that means. Here $\frac{1}{x}$ is a distribution.

There are two approaches to the problem that $\frac{1}{x}$ is infinite at the origin. The first is to "regularize" it, replacing $\frac{1}{x}$ by $\frac{1}{x}\chi_{\{|x|>\epsilon\}}$, estimate the $L^p$ (or weak $L^1$) norms of $\frac{1}{x}\chi_{|x|>\epsilon}$ as a convolution operator, show they are uniformly bounded as $\epsilon \longrightarrow 0$ and conclude that $H$ is bounded. This amounts to realizing $Hf$ as $\lim\limits_{\epsilon\longrightarrow 0}\int_{|x-y|<\epsilon} f(y)dy$.

The other approach, which I wish to discuss is a little more detail, is to realize that $Hf = \frac{1}{x} * f$ is actually valid on a suitable space of test functions $f$ and then to attempt a density argument in $f$. It is this direction that leads most simply to the Calderón-Zygmund theory.

Indeed, given any multiplier $T$, one assumes that $T$ is bounded on $L^q$ ($q = 2$ for the

Hilbert transform) and that it has the form

$$Tf(x) = \int_{\mathbb{R}^n} K(x,y)f(y)dy \tag{$*$}$$

where the kernel $K$ is singular near $x = y$.

Make the following assumptions on $K$

(1) There is a constant $A$ so that

$$\|Tf\|_q \leq A\|f\|_q \qquad \forall f \in L^q.$$

(2) For the same constant $A$ and for some constant $c > 1$,

$$\int_{x \notin B(y,c\delta)} |K(x,y) - K(x,\bar{y})|dx \leq A \text{ whenever } \bar{y} \in B(y,\delta).$$

(3) For all $f$ in $L^q$ with compact support , $(*)$ converges absolutely for almost all $x$ in the complement of the support of $f$ and $(*)$ holds for all these elements.

**Theorem.** *Under these assumptions, the operator $T$ is bounded in the $L^p$ norm on $L^p \cap L^1$ whenever $1 < p < q$.*

*More precisely*

$$\|Tf\|_p \leq A_p\|f\|_p \qquad \forall 1 < p < q.$$

*The bound $A_p$ depends only upon $A$.*

*Furthermore, $T$ is weak type $(1,1)$.*

**Proof.** See Stein pp20-22. □

The analogue in $\mathbb{R}^n$ of the Hilbert transform $sgn\xi = \frac{\xi}{|\xi|}$ on $\mathbb{T}$ or $\mathbb{R}$ are the operators $R_j$ given by multiplying the Fourier transform by $\frac{\xi_j}{|\xi|}$, for $j = 1, \ldots, n$. These are called Riesz transforms and are canonical examples of Calderón-Zygmund operators. One can in fact show that if $1 - \frac{1}{n} < p < \infty$, $f \in H^p(\mathbb{R}^n)$ if and only if $f$ and the $R_j(f)$ belong to $L^p$.

The study of Calderón-Zygmund theory will be taken up by Duong in Week 2.

# CHAPTER 5: FUNDAMENTAL SOLUTIONS OF PARTIAL DIFFERENTIAL OPERATORS

## 5.1. Introduction.

We apply the theory of distributions to PDO's, giving in particular the Malgrange-Ehrenpreis theorem which states that every linear PDO with constant coefficients has a fundamental solution.

We give the distributional solution of the standard fundamental equations of mathematical physics, preceded by a brief discussion of hypoellipticity.

## 5.2. Local solvability.

Let $D_\alpha = (i)^{-|\alpha|}\frac{\partial^{|\alpha|}}{\partial x^\alpha}$, so that $(D_\alpha f)\hat{}\,(\xi) = \xi^\alpha \hat{f}(\xi)$.

A PDO with constant coefficients is an expression of the form

$$L = \sum_{|\alpha|\le k} a_\alpha D_\alpha,$$

where $a_\alpha \in \mathbb{C}$. It has order $k$ if $\sum_{|\alpha|=k}|a_\alpha| \ne 0$. If $P(\xi) = \sum a_\alpha \xi^\alpha$, then $L = P(D)$ and we have $(P(D)u)\hat{}\,(\xi) = P(\xi)\hat{u}(\xi)$.

**Problem.** Given $f \in C^\infty(\Omega)$, find a distribution $\Lambda \in \mathcal{D}'(\Omega)$ such that

$$P(D)\Lambda = f. \tag{$*$}$$

We say that $L$ is locally solvable at $x_0 \in \mathbb{R}^n$ if $\exists$ a neighbourhood of $(x_0)$ so that $(*)$ holds for all points in that neighbourhood.

We may as well assume that $f$ has compact support.

**Theorem.** *Let $L$ be a PDO with constant coefficients. If $f \in \mathcal{D}$, there is a $C^\infty$ function $u$ satisfying $Lu = f$ on $\mathbb{R}^n$.*

**Proof.** Take the Fourier transform of the equation $P(D)u = f$. We get $P(\xi)\hat{u}(\xi) = \hat{f}(\xi)$ and it is natural to try to define $u$ by

$$u(x) = \int_{\mathbb{R}^n} e^{ix.\xi}\frac{\hat{f}(\xi)}{P(\xi)}d\xi.$$

In general, $P(\xi)$ will have many zeroes, so there will be a problem with taking the inverse Fourier transform of $\hat{f}/P$. But, since $f \in \mathcal{D}$, $\hat{f}$ is entire, so we may deform the contour of integration to avoid the zeroes of $P(\xi)$.

Let's make this precise. Choose $|\eta| = 1$ so that $\sum_{|\alpha|=k} a_\alpha \eta^\alpha \neq 0$. By rotating the coordinates if necessary, we can assume that $\eta = (0, 0, \ldots, 1)$. Assume also that $a_{\alpha_0} = 1$, where $\alpha_0 = (0, 0, \ldots, k)$. Then $P(\xi) = \xi_n^k +$ lower order terms in $\xi_n$.

Let $\xi = (\xi', \xi_n)$ where $\xi' \in \mathbb{R}^{n-1}$, and consider $P(\xi) = P(\xi', \xi_n)$ as a polynomial in $\xi_n (\in C)$ for $\xi' \in \mathbb{R}^{n-1}$ fixed.

Let $\lambda_1(\xi'), \ldots, \lambda_k(\xi')$ be the roots of $P(\xi', \cdot)$ arranged so that if $i \leq j$, $Im(\lambda_i(\xi')) \leq Im(\lambda_j(\xi'))$, and $Re(\lambda_i(\xi')) \leq Re(\lambda_j \xi')$ if the imaginary parts are equal.

Since the roots of a polynomial depend continously on the coefficients, the $Im(\lambda_j(\xi'))$ depend continuously on $\xi'$.

We need two lemmas.

**Lemma A.** There is a measurable function $\phi : \mathbb{R}^{n-1} \to [-k-1, k+1]$ such that for all $\xi' \in \mathbb{R}^{n-1}$,

$$\min_{1 \leq j \leq k} |\phi(\xi') - Im\lambda_j(\xi')| \geq 1.$$

**"Proof"** At least one of the intervals $[-k-1, -k+1]$, $[-k+1, -k+3], \ldots, [k-1, k+1]$ contains none of the numbers $Im(\lambda_j(\xi'))$, so we take $\phi(\xi')$ to be the mid point of that interval. Now show that $\phi$ is measurable. $\square$

**Lemma B** Let $P(\xi) = \xi_n^k +$ lower terms in $\xi_n$. Let $N(P) = \{\xi \in \mathbb{C}^n : P(\xi) = 0\}$ and let $d(\xi) = d(\xi, N(P))$. Then $|P(\xi)| \geq \left(\frac{d(\xi)}{2}\right)^k$.

**Proof.** Take $\xi \in \mathbb{R}^n$ such that $P(\xi) \neq 0$. Let $\eta = (0, 0, \ldots, 1)$ and define $g(z) = P(\xi + z\eta)$ $\forall z \in \mathbb{C}$. Then $g$ is a polynomial in one complex variable $z$.

Let $\lambda_1, \ldots \lambda_k$ be the zeroes of $g$. Then

$$g(z) = C(z - \lambda_1) \ldots (z - \lambda_k), \text{ so } \left|\frac{g(z)}{g(0)}\right| = \prod_{j=1}^{k} \left|1 - \frac{z}{\lambda_j}\right|.$$

Since $\xi + \lambda_j \eta \in N(P)$, $|\lambda_j| \geq d(\xi)$, so $|1 - \frac{z}{\lambda_j}| \leq 2$. Thus, for $|z| \leq d(\xi)$, $\left|\frac{g(z)}{g(0)}\right| \leq 2^k$. Also,

$$\begin{aligned} |g^{(k)}(0)| &= |\frac{k!}{2\pi i} \int_{|\xi|=d(\xi)} \frac{g(\zeta)}{\zeta^{k+1}} d\zeta| \leq \frac{k!}{2\pi} \frac{|g(0)|}{(d(\xi))^{k+1}} 2^k . 2\pi d(\xi) \\ &= k! \frac{|P(\xi)|}{(d(x))^k} 2^k. \end{aligned}$$

But $g^{(k)}(0) = \frac{\partial^k}{\partial \xi_n^k} P(\xi)\big|_{\xi=0} = k!$

Thus we obtain $k! \le k!|P(\xi)|\left(\frac{2}{d(\xi)}\right)^k$, or $|P(\xi)| \ge \left(\frac{d(\xi)}{2}\right)^k$. $\square$

Now let's finish the proof of the theorem. Let $u(x) = \int_{\mathbb{R}^{n-1}} \int_{Im\xi_n=\phi(\xi')} e^{2\pi i x.\xi} \left(\frac{\hat{f}(\xi)}{P(\xi)}\right) d\xi_n d\xi'$. By Lemmas $A$ and $B$,

$$|P(\xi)| \ge \left(\frac{d(\xi)}{2}\right)^k \ge \frac{1}{2^k} \text{ along } Im\xi_n = \phi(\xi').$$

Since $f \in \mathcal{D}$, $\hat{f}(\xi)$ is rapidly decreasing as $|Re\xi| \to \infty$, provided $|Im\xi|$ stays bounded (and it is less than $k+1$), so the integral converges absolutely and uniformly, together with all its derivatives, defining a $C^\infty$ function $u$.

Finally, we must show that $u$ is a solution. By the Cauchy integral formula

$$\begin{aligned}(P(D)u)(x) &= \int_{\mathbb{R}^{n-1}} \int_{Im\xi_k=\phi(\xi')} P(\xi) e^{2\pi i x.\xi} \frac{\hat{f}(\xi)}{P(\xi)} d\xi_n d\xi' \\ &= \int_{\mathbb{R}^n} e^{2\pi i x.\xi} \hat{f}(\xi) d\xi = f(x).\end{aligned}$$ $\square$

**5.3. Fundamental solutions.**

**Problem.** Can you (locally) solve $P(D)\Lambda = f$, where $f$ is a distribution?

As before, we may assume that $f$ has compact support.

A distribution $\Lambda$ is called a fundamental solution for $L = P(D)$ if it satisfies $P(D)\Lambda = \delta_0$.

By the theory of Ch.3, if $f \in \mathcal{D}'$ has compact support and $\Lambda$ is a fundamental solution for $L$, then $u = \Lambda * f$ is a distributional solution for $Pu = f$.

**Theorem.** (Malgrange-Ehrenpreis) *Every PDO with constant coefficients has a fundamental solution.*

**Proof.** We proceed as in the previous proof. Try to define

$$\Lambda(x) = \int_{\mathbb{R}^{n-1}} \int_{Im\xi_n=\phi(\xi')} e^{2\pi i x.\xi} \frac{1}{P(\xi)} d\xi_n d\xi'.$$

Unfortunately, without the $\hat{f}(\xi)$ on top, this may diverge at infinity. We proceed to regularize. For a positive integer $N$, set

$$P_N(\xi) = P(\xi)(1 + 4\pi^2 \sum_{j=1}^{n} \xi_j^2)^N, \text{ and let}$$

$$\Lambda_N(\xi) = \int_{\mathbb{R}^{n-1}} \int_{Im\xi_n=\phi_N(\xi')} \frac{e^{2\pi i x.\xi}}{P_N(\xi)} d\xi_n d\xi',$$

where $\phi_N$ is the function chosen for $P_N$ as in the previous proof.

On the region of integration, we have $P_N(\xi) \geq c(1+|\xi|^2)^N$, so the integral converges provided that $N \geq \frac{n}{2}$.

**Claim.** $P_N(D)\Lambda_N = \delta_0$. To see this, test the left hand side against $\phi \in \mathcal{D}$.

$$\begin{aligned}
(P_N(D)\Lambda_N)\phi &= \Lambda_N(P_N(-D)\phi) \\
&= \int_{\mathbb{R}^n}\int_{\mathbb{R}^{n-1}}\int_{Im\xi_n=\phi(\xi')} \frac{e^{2\pi i x.\xi}(P_N(-D)\phi)(x)}{P_N(\xi)}\, d\xi_n d\xi' dx \\
&= \int_{\mathbb{R}^{n-1}}\int_{Im\xi_n=\phi(\xi')} \frac{1}{P_N(\xi)}\int_{\mathbb{R}^n} e^{2\pi i x.\xi} P_N(-D)\phi(x)dx d\xi_n d\xi' \\
&= \int_{\mathbb{R}^{n-1}}\int_{Im\xi_n=\phi(\xi')} \hat{\phi}(-\xi)d\xi_n d\xi' = \int_{\mathbb{R}^n} \hat{\phi}(-\xi)d\xi \\
&= \phi(0) = \delta(\phi).
\end{aligned}$$

Thus, $\delta = P_N(D)\Lambda_N = P(D)((1-\Delta)^N\Lambda_n)$ where $\Delta$ is the Laplacian, $\sum_{i=1}^n \frac{\partial^2}{\partial x_i^2}$. Thus, if we put $\Lambda = (1-\Delta)^N\Lambda_N$, we have $P(D)\Lambda = \delta$ and so $\Lambda$ is a fundamental solution for $P(D)$. □

This elegant proof is due to Folland.

Hörmander has improved this result to get the best possible conditions on $\Lambda$ given the degree of $P$.

It doesn't take much to write down differential equations which do not have fundamental solutions. H. Lewy (Ann. Math. 1957) gave the following fascinating example. Let $L$ be defined on $\mathbb{R}^3$, $L\phi = -\phi_x - i\phi_y + 2i(x+iy)\phi_z$.

There exists a function $F(x,y,z) \in C^\infty(\mathbb{R}^3)$ such that the equation $Lu = F$ has no local solution anywhere.

### 5.4. Hypoellipticity.

The singular support of a distribution $\Lambda \in \mathcal{D}'$ is the complement of the largest open set on which $\Lambda$ is a $C^\infty$ function. This is denoted sing supp $\Lambda$.

Let $L = \sum a_\alpha(x)D^\alpha$ where $a_\alpha(x) \in C^\infty(\Omega)$ be a differential operator. $L$ is said to be **hypoelliptic** if for all $\Lambda \in \mathcal{D}'$, sing supp $\Lambda \subseteq$ sing supp $L\Lambda$, i.e. if for every open set $\Omega \subseteq \mathbb{R}^n$, $\Lambda \in \mathcal{D}'(\Omega)$, $L\Lambda \in C^\infty(\Omega) \Rightarrow \Lambda \in C^\infty(\Omega)$. (Recall that $L$ is **elliptic** if for all $x \in \mathbb{R}$ $\sum_{|\alpha|=k} a_\alpha(x)\xi^\alpha \neq 0$). Hörmander's theorem (below) says that every elliptic operator is hypoelliptic. Observe that if $L$ is hypoelliptic then every fundamental solution of $L$ is $C^\infty$ in $\mathbb{R}^n \backslash \{0\}$.

Before stating Hörmander's theorem, let me remind you of the definition of the Sobolev space $H_s$, for $s \in \mathbb{R}$. Define the operator $\Lambda^s : \mathcal{S} \to \mathcal{S}$ by $(\Lambda^s f)\hat{}(\xi) = (1 + |\xi|^2)^{s/2} \hat{f}(\xi)$, i.e. $\Lambda^s = (1 - \frac{\Delta}{4\pi^2})^{s/2}$. Clearly, $\Lambda^s$ is a continuous map from $\mathcal{S}$ to itself and hence by duality it defines a continuous map $\mathcal{S}' \to \mathcal{S}'$.

The Sobolev space $H_s = \{f \in \mathcal{S}' : \Lambda^s f \in L^2\}$, is endowed with the norm $\|f\|_{(s)} = \|\Lambda^s f\|_2$. Note that $C = \cap_{s \in \mathbb{R}} H_s^{\text{loc}}(\Omega)$.

Consider the following conditions on the polynomial $P$

(H) There exists $\delta > 0$ such that for all $\alpha$

$$\left| \frac{P^{(\alpha)}(\xi)}{P(\xi)} \right| = 0(|\xi|^{-\delta|\alpha|}) \text{ as } |\xi| \to \infty.$$

(H') $\left| \frac{P^{(\alpha)}(\xi)}{P(\xi)} \right| \to 0$ as $|\xi| \to \infty$ $\forall \alpha \neq 0$.

(H") $|Im \xi| \to \infty$ as $|\xi| \to \infty$ in the set $\{\zeta \in \mathbb{C}^n : P(\zeta) = 0\}$.

**Hörmander's theorem.** The following are equivalent

(1) $L$ is hypoelliptic.

(2) $P$ satisfies condition (H).

(3) $P$ satisfies condition (H').

(4) $P$ satisfies condition (H").

**Remarks:**

- Hörmander actually showed that if $P$ satisfies condition $(H)$, that if $f \in \mathcal{D}'(\Omega)$ and $P(D)f \in H_s^{\text{loc}}(\Omega)$ then $f \in H_{s+k\delta}^{\text{loc}}$, where $k$ is the degree of $P$. This shows that (2) $\Rightarrow$ (1).

  The remainder of the proof requires some sophisticated algebraic geometry.
- Recall that $P$ is elliptic if $\sum_{\alpha=|k|} a_\alpha \xi^\alpha \neq 0$ whenever $\xi \neq 0$. One can show that $P$ is elliptic if and only if $P$ satisfies condition $(H)$ with $\delta = 1$.

**5.5. Some fundamental solutions.**

- The **Laplace operator** $\Delta = \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}$ on $\mathbb{R}^n$ has fundamental solution $\frac{1}{(2-n)w_n}|x|^{2-n}$ where $w_n$ is the area of the unit sphere in $\mathbb{R}^n$, for $n \neq 2$, and $\frac{1}{2\pi} \log |x|$ for $n = 2$.

- The **Heat equation** on $\mathbb{R}^n$ is $\frac{\partial}{\partial t} - \Delta$ and we look for a distribution $K$ on $\mathbb{R}^n \times \mathbb{R}$ such that $\left(\frac{\partial}{\partial t} - \Delta\right) K = \delta_0(x)\delta_0(t)$.

  One can show that $K(x,t) = \begin{cases} (4\pi t)^{-n/2}e^{-|x|^2/4t} & t > 0 \\ 0 & t < 0 \end{cases}$ is a fundamental solution.
- The **Wave equation** $\left(\frac{\partial^2}{\partial t^2} - \Delta\right)$ has fundamental solutions $K(\xi, t) = \frac{H(t)}{4\pi^2(|\xi|^2 - t^2)}$, where $H(t)$ is the characteristic function of $[0, \infty)$.

  Actually, there are two ways of making this into a distribution, setting

$$K_+ = \lim_{\epsilon \to 0} \frac{1}{4\pi^2(|\xi|^2 - (t - i\epsilon)^2)}$$

and

$$K_- = \lim_{\epsilon \to 0} \frac{1}{4\pi^2(|\xi|^2 - (t + i\epsilon)^2)}.$$

The difference $K_+ - K_-$ is a distribution supported on the cone $|\xi| = |t|$.

**5.6. Postscript: $\Psi$ DO's**

Note that if $L = \sum_{|\alpha| \leq k} a_\alpha(x) D^\alpha$ is a PDO, then $Lu(x) = \int e^{2\pi i x.\xi} p(x,\xi)\hat{u}(\xi)d\xi$, where $p(x,\xi) = \sum_{|\alpha| \leq k} a_\alpha(x)\xi^\alpha$. The notion of a **pseudo-differential operator** ($\Psi$ DO) is obtained by replacing the function $p(x,\xi)$ in the above formula by symbols, i.e. functions $p$ in

$$S^m(\Omega) = \{p \in C^\infty(\Omega \times \mathbb{R}^n) : \forall \alpha, \beta,\ \forall \Omega' \subseteq C\Omega \exists C = C_{\alpha,\beta,\Omega'}$$

$$\text{such that } \sup_{x \in \Omega'} \left|D_x^\beta D_\xi^\alpha p(x,\xi)\right| \leq C(1 + |\xi|)^{m - |\alpha|}\}$$

This opens up a whole new chapter of analysis.

# CHAPTER 6: NONCOMMUTATIVE HARMONIC ANALYSIS

## 6.1. Introduction.

In this chapter, I will discuss some of the basic features of noncommutative harmonic analysis. Time does not permit a detailed description of how all aspects of the first five lectures go through in the noncommutative case – though many of them do. In fact, I will go little further than setting up the Fourier transform – that is, the analogue of Ch.1.

## 6.2. Locally compact groups. Unitary representations.

As we mentioned before, every locally compact group $G$ is equipped with a left invariant Haar measure, that is a positive regular Borel measure $d\lambda$ so that $\forall y \in G$ $\int_G f(y^{-1}x)d\lambda(x) = \int_G f(x)d\lambda(x)$.

Furthermore, $\lambda$ is unique up to scalar multiplication. ($G$ is likewise equipped with a right invariant Haar measure $d\rho$, which is not necessarily the same as the left one. In fact, the mapping $x \mapsto x^{-1}$ interchanges $\lambda$ and $\rho$ and the modular function of $G$ is $\triangle(x) = \frac{d\lambda(x^{-1})}{d\lambda(x)}$. A group is unimodular if $\triangle(x) \equiv 1$.)

We can thus form the usual function spaces, $L^p(G)$, $C(G)$, $C_0(G)$, $M(G)$, etc. and ask how much of classical Fourier analysis goes through in this setting.

The first, major problem is how to define the Fourier transform. Unfortunately, there are not sufficiently many characters of a typical locally compact group to separate points. One replaces characters by irreducible unitary representations. The Gelfand-Raikov theorem tells us that we may then separate points.

We say that $\sigma : G \to U(\mathcal{H})$, $\mathcal{H}$ a Hilbert space, is a **continuous unitary representation** if $\sigma(xy) = \sigma(x) \circ \sigma(y)$ for all $x, y \in G$ and $\sigma(e) = I$. A subspace $\mathcal{H}_1$ of $\mathcal{H}$ is called **invariant** if $\pi(x)\mathcal{H}_1 \subseteq \mathcal{H}_1$ for all $x \in G$, and $\sigma$ is **irreducible** if $\mathcal{H}$ has no nontrivial invariant subspaces.

If $G$ is commutative, every continuous unitary irreducible representation (CUIR) is one-dimensional, i.e. is a character, and so the set of CUIR's reduce to $\widehat{G}$.

Actually, to avoid redundancy, we wish to identify equivalent representations of $G$. Two representations $\sigma, \sigma_1$ on $\mathcal{H}, \mathcal{H}_1$ are **equivalent** if there is an isometry $i : \mathcal{H} \to \mathcal{H}_1$ so that for all $x \in G$ $i \circ \sigma(x) = \sigma_1(x) \circ i$. Thus, for a noncommutative group, we define $\widehat{G}$ as the set of all CUIR's modulo this equivalence relation (OK, there are some set-theoretic problems with the set of all Hilbert spaces, but they can be safely dealt with!) Actually, it's more convenient to think of $\widehat{G}$ as a maximal set of pairwise inequivalent CUIR's of $G$.

Given $\sigma \in \widehat{G}$ and $f \in L^1$, we may define $\hat{f}(\sigma) \in \mathcal{B}(\mathcal{H}_\sigma)$ by

$$\hat{f}(\sigma) = \int_G f(x)\sigma(x^{-1})dx.$$

This is a vector valued integral. In fact, many of the elementary properties of the

Fourier transform continue to hold.

**Theorem.** *$\hat{}$ is a linear mapping from $L^1(G)$ to elements of a "fibre bundle" over $\widehat{G}$, where the fibre over $\sigma$ is $\mathcal{B}(\mathcal{H}_\sigma)$.*

*One has $\|\hat{f}(\sigma)\|_{\mathcal{B}(\mathcal{H}_\sigma)} \leq \|f\|_1$, and $(f * g)\hat{}(\sigma) = \hat{f}(\sigma)\hat{g}(\sigma)$.*

**Proof.** The proof is just a re-write of the proof from Ch.1. □

In the commutative case, $\widehat{G}$ became a topological group. Here, neither the topology nor the group structure is evident!

There is a version of the hull-kernel topology, defined by Fell. We say that a representation $\sigma$ is in the closure of a set of representations $\mathcal{F}$ if every function of the form $x \mapsto \langle\sigma(x)\xi,\xi\rangle, \xi \in \mathcal{H}_\sigma$ can be uniformly approximated on compact sets of $G$ by functions of the form $x \mapsto \langle\eta(x)\xi,\xi\rangle,\ \xi \in \mathcal{H}_\eta,\ \eta \in \mathcal{F}$.

Although this topology generalizes the classical hull-kernel topology, it is not nice. It is not even $T_0$ for a general group! One can, however, show that if $f \in L^1$, $\hat{f}$ approaches zero at infinity, in some sense.

There is also a kind of arithmetic operation on representations. Let $\sigma, \eta \in \widehat{G}$, form the tensor product $\sigma \otimes \eta$ and decompose into a direct sum, or direct integral, of irreducibles. This gives a kind of set-valued multiplication on $\widehat{G}$ which makes it in to a "hypergroup" in some circumstances. Unfortunately, for many groups, there is no guarantee of uniqueness of decomposition of tensor products, so this operation is not well-defined.

Fortunately, for many groups, those of type I, all this can be overcome and everything works, more or less.

### 6.3. The Group C*-algebra. Type I groups.

In the case of $\mathbb{T}$, the set of characters is the Gelfand space of the commutative Banach algebra $c_0(\mathbb{Z})$, and an analogue of this statement persists in the general case. In fact, the algebra $L^1(G)$ is a Banach algebra has an enveloping $C^*$-algebra, defined as follows. Let

$$\|f\|_{C^*(G)} = sup_{\sigma\in\widehat{G}}\|\hat{f}(\sigma)\|_{\mathcal{B}(\mathcal{H}_\sigma)}, \quad \text{for } f \in L^1(G).$$

Then $\|f\|_{C^*(G)} \leq \|f\|_1$ and we may complete $L^1$ in the $C^*$-norm so that it becomes a $C^*$-algebra. Furthermore, $\widehat{G}$ is in one-one correspondence with the set of equivalence classes of non-degenerate $*$-representations of $C^*(G)$.

This process allows one to use a massive amount of the machinery of $C^*$-algebras to study the duals of groups. However, the subtleties of characterizing the Fourier transforms of functions in $L^1$ are lost in the embedding of $L^1$ in $C^*(G)$!

In particular, there is a class of $C^*$-algebras called type I $C^*$-algebras for which everything works out nicely. (Every nontrivial quotient algebra of $A$ has a closed nontrivial two-sided ideal with the property that every irreducible representation of this ideal is in the compact operators.)

In fact, for type I groups, for every irreducible representation $\sigma(C^*(G))$ contains the compact operators of $\mathcal{H}_\sigma$, two representations are equivalent iff their kernels agree, the Fell topology is $T_0$ (though still not $T_2$) and defines a reasonable Borel sigma algebra on $\widehat{G}$.

For groups of type I, there is a nice Plancherel theory.

**Theorem.** *Let $G$ be locally compact and type I. There exists a unique measure $\mu$ on $\widehat{G}$ so that for all $f \in L^1 \cap L^2(G)$, $\hat{f}(\sigma)\hat{f}(\sigma)^*$ is a.e. trace class and*

$$\int_G |f(x)|^2 dx = \int_{\widehat{G}} Tr(\hat{f}(\sigma)\hat{f}(\sigma)^*) d\mu(\sigma).$$

An accessible proof of this theorem is in the Appendix to Dixmier's book, $C^*$-algebras.

This theorem is a complete analogue of Plancherel's theorem. For type I groups, any representation decomposes into an essentially unique direct integral of irreducible representations. This enables us to define an arithmetic structure on $\widehat{G}$ as explained above and one can ask to recover $G$ from $\widehat{G}$. Various versions of this theory exist – Tannaka duality for compact groups, and a reasonable duality for all type I groups due to Tatsuma. Recently, Doplicher and Roberts have provided a rather satisfying kind of duality for all locally compact groups, involving knowing all the intertwining operators for all tensor decompositions. However, this is calculationally complex even for relatively simple groups. The problem of duality arises even in the case of finite groups, where $A_3$ and $D_2$ have the same character tables.

Which groups are of type I? Fortunately for suffering humanity, many of the ones we want to study are! The list includes all compact groups, nilpotent groups and semisimple Lie groups.

Groups which are not of type I are relatively easy to find. Free groups cannot be of type I, and many solvable Lie groups also fail to be. Here is a simple example, due to Mautner, which has inspired a number of people. It is a five dimensional solvable Lie group, the semidirect product of $\mathbb{R}$ with $\mathbb{C}\times\mathbb{C}$, where $\mathbb{R}$ acts on $\mathbb{C}\times\mathbb{C}$ by $r(z_1, z_2) = (e^{ir}z_1, e^{i\alpha r}z_2)$ where $\alpha$ is an irrational number. If a group is not of type I, its representation theory is necessarily extremely complicated.

### 6.3. Compact groups.

Everything is nice for compact groups. Each irreducible representation $\sigma$ has finite degree $d_\sigma$. The dual $\widehat{G}$ consists of a discrete set of points. The left regular representation of $G$ on $L^2(G)$ (defined by $L_gF(x) = f(g^{-1}x)$) may be decomposed into a direct sum of irreducible representations, each representation occurring with multiplicity equal to its degree

$$L^2(G) = \oplus_{\sigma\in\widehat{G}} d_\sigma\sigma \quad \text{(Peter – Weyl theorem)}.$$

The irreducible characters $\chi_\sigma(x) = tr(\sigma(x))$ on $G$ play a special role; convolution by $d_\sigma\chi_\sigma$ is the projection in $L^2(G)$ onto the invariant subspace associated with $\sigma$.

### 6.4. Lie groups.

The theory takes on some interesting new dimensions when one can differentiate – when $G$ is a Lie group. That is to say, $G$ is a manifold and also a group such that the multiplication and inverse are differentiable. The tangent space $T_e(G)$ at the identity may be endowed with the structure of a Lie algebra $\mathbf{g}$ as follows. Conjugation in $G$ (the map $A_g : x \mapsto g^{-1}xg$) being differentiable, its derivative gives a mapping $Ad(g) : \mathbf{g} \to \mathbf{g}$. Indeed, $Ad$ is a representation of $G$ and may be thought of as a homomorphism $G \to GL(\mathbf{g})$. Equipping $GL$ with the obvious structure of Lie group and differentiating we get a mapping $\mathbf{g} \to M_n(\mathbf{g})$ denoted by ad. The Lie bracket is defined by $[X, Y] = ad(X)Y$.

We have the exponential map exp : $\mathbf{g} \to G$ which is a diffemorphism near $0 \in \mathbf{g}$, and for $X \in \mathbf{g}$, $f \in C^\infty(G)$ we take the Lie derivative $Xf \in C^\infty(G)$ defined by

$$(Xf)(x) = \frac{d}{dt}\Big|_{t=0} f(x \exp tX).$$

We can now ask how much of Chapters 2-5 generalize. Much has been done for particular classes of groups.

Given a particular Lie group, or class of Lie groups, the program is more or less as follows:

(1) If it is not of type I, panic!

If it is of type I.

(2) Describe $\widehat{G}$ as explicitly as possible, or at least that part $\widehat{G}_r$ of $\widehat{G}$ which is needed to decompose the regular representation.

(3) Describe the topology Borel structure and Plancherel measure on $\widehat{G}$ (or at least on $\widehat{G}_r$).

(4) Set up the basic theorems of Fourier analysis.

(5) Theory of distributions, differential operators etc.

In the case of compact Lie groups, the Borel-Weil theorem provides a very elegant answer to (2), (3) is trivial, and much of (4) and (5) has been done.

In the case of real semisimple Lie groups, Harish-Chandra's ambitious program started in the 50's, and continuing through the works of Langlands, Vogan, Helgason etc. has solved much of (2)-(5).

In the case of nilpotent groups, Kirillov's orbit method provided a simple description of $\widehat{G}$, which has also been used for much of (3)-(5). Some of this work has also been carried through for solvable groups.

To exemplify some of this work, I shall state a rather elegant version of one of Harish-Chandra's theorems, valid for compact, semisimple and nilpotent groups.

**Theorem.** *Let $f \in C_c^\infty(G)$, $d\mu$ denote Plancherel measure on $\widehat{G}$.*

*Then*

$$f(x) = \int_{\widehat{G}} Tr(\hat{f}(\sigma)\sigma(x))d\mu(\sigma).$$

Actually, Bruhat set up the theory of distributions on a Lie group, and in the above statement, it is sometimes convenient to think of the character $Tr(\sigma(\cdot)) = \Theta(\cdot)$ as a distribution on $G$. We give some formulae for this in the next section.

**6.5. The orbit method.**

It turns out that both the Borel-Weil theorem and the Kirillov theorem are parts of the same machine – and many (but not all) of Harish-Chandra's representations also fall under this umbrella. In this section, I will give a brief sketch of how this works – it really deserves a lecture, or a course, of its own.

The idea is that one can construct representations of a Lie group $G$ be looking at its coadjoint orbits, that the set of these coadjoint orbits should parametrize a large set in $\widehat{G}$, and that the Euclidean Fourier transforms of orbits (functions on $\mathbf{g}$) composed with the exponential map and multiplied by a certain function, should give characters of representations.

If only life were so simple! However, the method works well for many Lie groups, so let me describe it briefly.

Let $\mathbf{g}^*$ be the dual of $\mathbf{g}$, and $Ad^*$ the coadjoint representation, defined by $(Ad^*(g)\beta, X) = (\beta, Ad(g^{-1})X)$.

Fix $\beta \in \mathbf{g}^*$, and let $\mathcal{O}_\beta$ be its orbit. Let $G_\beta$ be the stabilizer of $\beta$. Then for all $X, Y \in \mathbf{g}_\beta$, $\beta([X,Y]) = 0$. We say that $\beta$ is **integral** if there is a character $\lambda$ of $G_\beta$ so that $d\chi = \beta$.

Form the induced representation $\rho = \chi \uparrow_{G_\beta}^{G}$ acting by the left regular representation in $\{f \in C^\infty(G) : f(xg) = \overline{\chi(x)} f(g),\ g \in G,\ x \in G_\beta\}$. Now enlarge $g_\beta^{\mathbf{C}}$ to a $G_\beta$-stable maximal isotropic subspace $\mathbf{h}$ of $\mathbf{g}^{\mathbf{C}}$ for the symplectic form $\omega(X,Y) = \beta([X,Y])$, and restrict $\rho$ to the subspace of functions $f$ so that $Xf = 0\ \forall X \in \mathbf{h}$. Call it $\rho_\beta$.

In many cases, the resulting space of $C^\infty$ functions can be given an $L^2$-norm for which $\rho_\beta$ is irreducible and unitary. In many cases, this geometric procedure gives almost all of $\widehat{G}$.

In many cases, we have the character formula for $X \in \mathbf{g}$, $\Theta_\beta(\exp X) = j(X)\widehat{\mu}_\beta(X)$, where $\mu_\beta$ is $G$-invariant measure on $\mathcal{O}_\beta$ and $j$ is the square root of the determinant of the exponential map.

For groups for which the above theory works, the fact that everything is done using essentially Euclidean tools gives us readily accessible methods for extending Fourier analysis.

## References


(1) J. Dixmier, $C^*$-algebras and their representations, Gauthier-Villars 1969.

(2) E. Hewitt and K.A. Ross, Abstract Harmonic Analysis, Vol.I, II Springer 1968, 1974.

(3) L. Hörmander, The analysis of linear partial differential operators, Vol.2 Springer, 1984.

(4) Y. Katznelson, An introduction to harmonic anlaysis, Dover 1976.

(5) A. Kirillov, Elements of the theory of representations, Springer 1976.

(6) W. Rudin, Functional analysis, McGraw Hill 1972.

(7) W. Rudin, Fourier analysis on groups, Wiley 1960.

(8) E.M. Stein, Harmonic analysis, Princeton 1993.