Class Notes for Harmonic Analysis
MATH 693

by
S. W. Drury

Contents

1

Differentiation and the Maximal Function

We start with the remarkable Vitali Covering Lemma. The action takes place on $\mathbb{R}$, but equally well there are corresponding statement in $\mathbb{R}^d$. We denote by λ the Lebesgue measure.

LEMMA 1 (VITALI COVERING LEMMA) *Let $\mathcal{S}$ be a family of bounded open intervals in $\mathbb{R}$ and let S be a Lebesgue subset of $\mathbb{R}$ with $\lambda(S) < \infty$ and such that*

$$S \subseteq \bigcup_{I \in \mathcal{S}} I.$$

Then, there exists $N \in \mathbb{N}$ and pairwise disjoint intervals $I_1, I_2, \ldots I_N$ of $\mathcal{S}$ such that

$$\lambda(S) \leq 4 \sum_{n=1}^{N} \lambda(I_n). \tag{1.1}$$

Proof. Since $\lambda(S) < \infty$, there exists K compact, $K \subseteq S$ and $\lambda(K) > \frac{3}{4}\lambda(S)$. Now, since

$$K \subseteq \bigcup_{I \in \mathcal{S}} I$$

there are just finitely many intervals $J_1, J_2, \ldots, J_M$ with $K \subseteq \bigcup_{m=1}^{M} J_m$. Let these intervals be arranged in order of decreasing length. Thus $1 \leq m_1 < m_2 \leq M$ implies that $\lambda(J_{m_1}) \geq \lambda(J_{m_2})$. We will call this the J list. We proceed algorithmically. If the J list is empty, then $N = 0$ and we stop. Otherwise, let I_1 be the first element of the J list (in this case J_1). Now, remove from the J list, all intervals that meet I_1 (including I_1 itself). If the J list is empty, then $N = 1$ and we stop. Otherwise, let I_2 be the first remaining element of the J list. Now, remove from the J list, all intervals that meet I_2 (including I_2 itself). If the J list is empty, then $N = 2$ and we stop. Otherwise, let I_3 be the first remaining element of the J list. Now, remove from the J list, all intervals that meet I_3...

Eventually, the process must stop, because there are only finitely many elements in the J list to start with. Clearly, the I_n are pairwise disjoint, because if I_{n_1} meets I_{n_2} and $1 \leq n_1 < n_2 \leq N$, then, immediately after I_{n_1} was chosen, all those intervals of the J list which meet I_{n_1} were removed. Since I_{n_2} was eventually chosen from this list, it must be that $I_{n_1} \cap I_{n_2} = \emptyset$. Now, we claim that for every J_m is contained in an interval $I_n^\star$ which is our notation for the interval with the same centre as I_n but three times the length. To see this, suppose that J_m was removed from the J list immediately after the choice of I_n. Then J_m was in the J list immediately prior to the choice of I_n and we must have that $\mathrm{length(J_m)} \leq \mathrm{length(I_n)}$ for otherwise J_m would be strictly longer than I_n and I_n would not have been chosen as a longest interval at that stage. Also J_m must meet I_n because it was removed immediately after the choice of I_n. It therefore follows that $J_m \subseteq I_n^\star$.

So, $K \subseteq \bigcup_{m=1}^M J_m \subseteq \bigcup_{n=1}^N I_n^\star$ and $\lambda(K) \leq \sum_{n=1}^N \lambda(I_n^\star) = 3\sum_{n=1}^N \lambda(I_n)$. It follows that (1.1) holds. ■

We now get an estimate for the Hardy-Littlewood maximal function. Let us define for $f \in L^1(\mathbb{R}, \mathcal{L}, \lambda)$

$$Mf(x) = \sup_{\substack{I \text{open interval} \\ x \in I}} \frac{1}{|\lambda(I)|} \left| \int_I f(t) d\lambda(t) \right|.$$

THEOREM 2 *We have* $\lambda(\{x; |Mf(x)| > s\}) \leq 4s^{-1}\|f\|_1$.

The result says that Mf satsifies a Tchebychev type inequality for L^1. It is easy to see that we do not necessarily have $Mf \in L^1$. For example, if $f = \mathbb{1}_{[-1,1]}$, then we have

$$Mf(x) = \begin{cases} 1 & \text{if } |x| \leq 1, \\ \dfrac{2}{|x|+1} & \text{if } |x| \geq 1. \end{cases}$$

and Mf is not integrable even though f is. Note that we may also define the centred maximal function

$$M_c f(x) = \sup_{h>0} \frac{1}{2h} \left| \int_{x-h}^{x+h} f(t) dt \right|.$$

Proof. First of all, there is no loss of generality in assuming that $f \geq 0$. Let $a \in \mathbb{N}$ and let $S = [-a, a] \cap \{x; |Mf(x)| > s\}$. Let $\mathcal{S}$ be the set of intervals I such that

$$\frac{1}{\lambda(I)} \int_I f(t)dt > s. \tag{1.2}$$

Then if $x \in S$ there is some I such that (1.2) holds and so I is in $\mathcal{S}$. Hence the hypotheses of the Vitali Covering Lemma are satisfied. We can then find N disjoint intervals I_n such that $\lambda(S) \leq 4 \sum_{n=1}^{N} \lambda(I_n)$. But

$$s \sum_{n=1}^{N} \lambda(I_n) \leq \sum_{n=1}^{N} \int_{I_n} f(t)dt = \int \left(\sum_{n=1}^{N} \mathbb{1}_{I_n} \right) f d\lambda \leq \|f\|_1$$

Note that the disjointness of the intervals is key here. It is used to show that $\sum_{n=1}^{N} \mathbb{1}_{I_n} \leq \mathbb{1}$. It follows that $\lambda(S) \leq 4s^{-1}\|f\|_1$. Now it suffices to let $a \longrightarrow \infty$ to obtain the desired conclusion. ∎

The results of this section are easily generalized to $\mathbb{R}^d$ and in a variety of ways. In the Vitali Covering Lemma, we can replace intervals by either cubes with sides parallel to the coordinate axes or with balls. In fact, you can easily find in the literature many papers generalizing this lemma by replacing intervals with various families of sets satisfying various sets of conditions. Let us take balls. The key fact about balls that one is using is the following:

> If $r \geq r' > 0$ and if U and U' are open balls of radius r and r' respectively and if $U \bigcap U' \neq \emptyset$, then $U' \subseteq 3U$ where $3U$ denotes the ball with the same centre as U and radius $3r$.

Then we may define the corresponding maximal function by

$$Mf(x) = \sup_{\substack{U \text{open ball} \\ x \in U}} \frac{1}{|\lambda(I)|} \left| \int_U f(t) d\lambda(t) \right|.$$

and the estimate that will be obtained is

$$\lambda(\{x; |Mf(x)| > s\}) \leq C_d \, s^{-1} \|f\|_1.$$

where C_d is a constant depending only on the dimension. There are also corresponding estimates for the maximal function on $\mathbb{T}^d$, the d-dimensional torus.

1.1 Conditional Expectation Operators

As an example of orthogonal projections, we can look at conditional expectation operators. These arise when we have two nested σ-fields on the same set. So, let $(X, \mathcal{F}, \lambda)$ be a measure space and suppose that $\mathcal{G} \subseteq \mathcal{F}$. Then, $(X, \mathcal{G}, \lambda|_{\mathcal{G}})$ is an equally good measure space. It is easy to see that $L^2(X, \mathcal{G}, \lambda|_{\mathcal{G}})$ is a closed linear subspace of $L^2(X, \mathcal{F}, \lambda)$. It should be pointed out, that trivialities can arise even when we might not expect them. For example, let $X = \mathbb{R}^2$, $\mathcal{F}$ the Borel σ-field of $\mathbb{R}^2$ and $\mathcal{G}$ the sets which depend only on the first coordinate. Then unfortunately $L^2(X, \mathcal{G}, \lambda|_{\mathcal{G}})$ consists just of the zero vector.

The situation is very significant in probability theory, where the σ-fields $\mathcal{F}$ and $\mathcal{G}$ encode which events are available to different "observers". For example $\mathcal{G}$ might encode outcomes based on the first 2 rolls of the dice, while $\mathcal{F}$ might encode outcomes based on the first 4 rolls.

A useful example is the case where $A = [0, 1[$, $\mathcal{F}$ is the borel σ-field of X. Then partition $[0, 1[$ into n intervals and let $\mathcal{G}$ be the σ-field generated by these intervals. We take λ the linear measure on the interval. In this case $\mathbf{E}_{\mathcal{G}}$ will turn out to be the mapping which replaces a function with its average value on each of the given intervals.

Well, the orthogonal projection operator is denoted $\mathbf{E}_{\mathcal{G}}$. We view it as a map

$$\mathbf{E}_{\mathcal{G}} : L^2(X, \mathcal{F}, \lambda) \longrightarrow L^2(X, \mathcal{G}, \lambda) \subseteq L^2(X, \mathcal{F}, \lambda).$$

We usually understand $\mathbf{E}_{\mathcal{G}}$ in terms of its properties. These are

- $\mathbf{E}_{\mathcal{G}}(f) \in L^2(X, \mathcal{G}, \lambda)$.
- $\int (f - \mathbf{E}_{\mathcal{G}}(f)) g d\lambda = 0$ whenever $g \in L^2(X, \mathcal{G}, \lambda)$.

The probabilists will write this last condition as $\mathbf{E}(f - \mathbf{E}_{\mathcal{G}}(f))g = 0$ whenever $g \in L^2(X, \mathcal{G}, \lambda)$, where $\mathbf{E}$ is the scalar-valued expectation.

To get much further we will need the additional assumption that $(X, \mathcal{G}, \lambda)$ is σ-finite. So we are assuming the existence of an increasing sequence of sets $G_n \in \mathcal{G}$ with $X = \bigcup_{n=1}^{\infty} G_n$ and $\lambda(G_n) < \infty$. As an exercise, the reader should check that $\mathbf{E}_{\mathcal{G}}(\mathbb{1}_G f) = \mathbb{1}_G \mathbf{E}_{\mathcal{G}}(f)$ for $G \in \mathcal{G}$. We do this by taking the inner product against every function in $L^2(X, \mathcal{G}, \lambda)$ and using property (ii) above.

Next we claim that if $G \in \mathcal{G}$ and $\lambda(G) < \infty$ and f is a $\mathcal{F}$ measurable function carried on $\mathcal{G}$ and $|f| \leq 1$, then $|\mathbf{E}_{\mathcal{G}}(f)| \leq 1$. To see this, let $t > 1$ and let $g = \mathbb{1}_G \overline{\operatorname{sgn}(\mathbf{E}_{\mathcal{G}} f)} \mathbb{1}_{\{|\mathbf{E}_{\mathcal{G}} f| > t\}}$. Then we have

$$t\lambda\Big(\{|\mathbf{E}_{\mathcal{G}} f| > t\}\Big) \leq \int g \mathbf{E}_{\mathcal{G}} f d\lambda = \int g f d\lambda \leq \lambda\Big(\{|\mathbf{E}_{\mathcal{G}} f| > t\}\Big)$$

The only way out is that $\lambda\Big(\{|\mathbf{E}_{\mathcal{G}}f| > t\}\Big) = 0$. Since this is true for all $t > 1$ it follows that $|\mathbf{E}_{\mathcal{G}}f| \leq 1$ λ-a.e. This gives us a way of extending the definition of conditional expectation to L^∞ functions. We define for $f \in L^\infty(X, \mathcal{F}, \lambda)$,

$$\mathbf{E}_{\mathcal{G}}f(x) = \mathbf{E}_{\mathcal{G}}\mathbb{1}_{G_n}f(x) \qquad \forall x \in G_n$$

The apparent dependence of this definition on n is illusory because for $x \in G_n$

$$\mathbf{E}_{\mathcal{G}}\mathbb{1}_{G_{n+1}}f(x) = (\mathbb{1}_{G_n} \cdot \mathbf{E}_{\mathcal{G}}\mathbb{1}_{G_{n+1}}f)(x) = \mathbf{E}_{\mathcal{G}}\mathbb{1}_{G_n}\mathbb{1}_{G_{n+1}}f(x) = \mathbf{E}_{\mathcal{G}}\mathbb{1}_{G_n}f(x)$$

and indeed, as an exercise, the reader can show that the definition is independent of the choice of sequence G_n. The bottom line here is that $\mathbf{E}_{\mathcal{G}}$ is a norm decreasing map

$$\mathbf{E}_{\mathcal{G}} : L^\infty(X, \mathcal{F}, \lambda) \longrightarrow L^\infty(X, \mathcal{G}, \lambda).$$

Now let $1 \leq p < \infty$ and let $f \in V$ where V is the space of bounded $\mathcal{F}$-measurable simple functions carried by a subset $G \in \mathcal{G}$ with $\lambda(G) < \infty$. In this case we will have that $\mathbf{E}_{\mathcal{G}}f$ is a bounded $\mathcal{G}$ measurable function still carried by the subset G. We will estimate the L^p norm of $\mathbf{E}_{\mathcal{G}}f$.

$$\int |\mathbf{E}_{\mathcal{G}}f|^p d\lambda = \int (\mathbf{E}_{\mathcal{G}}f) g d\lambda$$

where $g = |\mathbf{E}_{\mathcal{G}}f|^{p-1}\overline{\mathrm{sgn}(\mathbf{E}_{\mathcal{G}}f)}$,

$$= \int f g d\lambda$$

since g is $\mathcal{G}$-measurable and all functions are in the appropriate L^2 space,

$$\leq \|f\|_p \|g\|_{p'},$$

by Hölder's Inequality. On the other hand

$$\|g\|_{p'}^{p'} = \int |\mathbf{E}_{\mathcal{G}}f|^{\frac{p}{p-1}(p-1)} d\lambda = \|\mathbf{E}_{\mathcal{G}}f\|_p^p,$$

leading to $\|g\|_{p'} \leq \|\mathbf{E}_{\mathcal{G}}f\|_p^{p-1}$. So, combining these inequalities yields

$$\|\mathbf{E}_{\mathcal{G}}f\|_p^p \leq \|f\|_p \cdot \|\mathbf{E}_{\mathcal{G}}f\|_p^{p-1} \tag{1.3}$$

We now obtain $\|\mathbf{E}_{\mathcal{G}}f\|_p \leq \|f\|_p$ because this is obvious if $\|\mathbf{E}_{\mathcal{G}}f\|_p = 0$ and if not, then we can divide out in (1.3) because we know that $\|\mathbf{E}_{\mathcal{G}}f\|_p < \infty$.

We have obtained that $\mathbf{E}_{\mathcal{G}}$ is a linear mapping from V to $L^p(X, \mathcal{F}, \lambda)$ and norm decreasing for the L^p norm. Since V is dense in $L^p(X, \mathcal{F}, \lambda)$, we can extend this mapping to a norm decreasing linear mapping $\mathbf{E}_{\mathcal{G}} : L^p(X, \mathcal{F}, \lambda) \longrightarrow L^p(X, \mathcal{F}, \lambda)$ by uniform continuity. We naturally use the same notation for this mapping, although strictly speaking it is a different mapping. This gives a nice application of "abstract nonsense" ideas to a really quite practical situation. You can check that the extended mapping satisfies the expected conditions which are valid even in the case $p = \infty$ handled earlier.

- $\mathbf{E}_{\mathcal{G}}(f) \in L^p(X, \mathcal{G}, \lambda)$ provided $f \in L^p(X, \mathcal{F}, \lambda)$ and indeed we have $\|\mathbf{E}_{\mathcal{G}} f\|_p \leq \|f\|_p$.
- $\int (f - \mathbf{E}_{\mathcal{G}}(f)) g d\lambda = 0$ whenever $g \in L^{p'}(X, \mathcal{G}, \lambda)$.

This is a typical example of the von Neumann program at work by using Hilbert space methods as a foot in the door to get results that have no obvious connection to Hilbert space.

1.2 Martingale Maximal Functions

This example develops a similar theorem with a different and instructive proof. We work on $[0, 1[$ with the Lebesgue field, we'll call it $\mathcal{F}$ and linear measure λ. A dyadic interval of length 2^{-n} is an interval $[(k-1)2^{-n}, k2^{-n}[$ for $k = 1, 2, \ldots, 2^n$ and $n = 0, 1, 2, \ldots$ The maximal function we deal with here for $f \in L^1([0, 1[, \mathcal{F}, \lambda)$ is

$$M_m f(x) = \sup \frac{1}{\lambda(I)} \int \mathbb{1}_I f d\lambda$$

where the sup is taken over all dyadic intervals that contain x. There is a more succinct way of writing this maximal function

$$M_m f(x) = \sup_{n=0}^{\infty} |\mathbf{E}_{\mathcal{F}_n} f|$$

where $\mathcal{F}_n$ is the σ-field (it's actually a field) generated by the dyadic intervals of length 2^{-n}. Note that for given $x \in [0, 1[$ and $n \in \mathbb{Z}^+$ there is a unique dyadic interval of length 2^{-n} to which x belongs.

To get further, we need to develop the probabilistic setting. A sequence of σ-fields $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots$ is called a ***stochastic base*** . The σ-field $\mathcal{F}_n$ contains those events that can be formulated at time n. The σ-field $\mathcal{F}$ is the σ-field generated by the union of all the $\mathcal{F}_n$, and contains all possible events. In our case, times n are nonnegative integers and we can imagine tossing a fair coin. So, at time 1, we toss the coin and if it ends up heads, we are in $[\frac{1}{2}, 1[$ and if it comes up tails, we are in $[0, \frac{1}{2}[$. At time 2, we toss again and we place in the upper half of the interval if we have a head and the lower half if we have a tail. The tossing is repeated indefinitely. Thus, if the result of the first 5 tosses is HTTHT, we are in the interval $[\frac{1}{2} + \frac{1}{16}, \frac{1}{2} + \frac{1}{16} + \frac{1}{32}[$.

Probabilists need to consider random times. In our case, these are mappings from the sample space $[0, 1[$ to the time space $\mathbb{Z}^+ \cup \{\infty\}$ which are $\mathcal{F}$ measurable. However, there is a very special class of random times called stopping times. We can think of a gambler who is following a fixed strategy. The quintessence of being a good gambler is knowing when to quit. But if the gambler is to quit at time n, then his decision has to be based on the information that is available to him at time n. If he were able to base his decision of whether to quit or not at time n on the information available at time $n + 1$, then he would be clairvoyant. So, a ***stopping time*** is a random time $\tau : [0, 1[\longrightarrow \mathbb{Z}^+ \cup \{\infty\}$ with the additional property

$$\{x; \tau(x) = n\} \in \mathcal{F}_n, \quad n = 0, 1, 2, \dots$$

This also implies

$$\{x; \tau(x) = \infty\} \in \mathcal{F}_\infty = \mathcal{F},$$

and in fact, you can make this part of the definition if you wish.

Let τ be a stopping time. We ask, what information is available at time τ. Well, an event $A \in \mathcal{F}$ can be formulated at time τ if and only if

$$A \cap \{x; \tau(x) = n\} \in \mathcal{F}_n, \quad n = 0, 1, 2, \dots$$

The collection of all such events A defines a σ-field $\mathcal{F}_\tau$ (prove this). This idea does not make a whole lot of sense for random times (the whole space need not be in $\mathcal{F}_\tau$), but it does make sense for stopping times. Now, since $\mathcal{F}_\tau$ is a σ-field , it has an associated conditional expectation operator $\mathbf{E}_{\mathcal{F}_\tau}$. The next thing to show is that

$$\mathbf{E}_{\mathcal{F}_\tau} f = \left(\sum_{n=0}^{\infty} \mathbb{1}_{\{x;\tau(x)=n\}} \mathbf{E}_{\mathcal{F}_n} f \right) + \mathbb{1}_{\{x;\tau(x)=\infty\}} f$$

We now have enough information to start to tackle the maximal function. Let $f \in L^1([0, 1[, \mathcal{F}, \lambda)$ and $f \geq 0$. Fix $s > 0$. We define $\tau(x)$ to be the first time n that $\mathbf{E}_{\mathcal{F}_n} f(x) > s$. If it happens that $\mathbf{E}_{\mathcal{F}_n} f(x) \leq s$ for all $n = 0, 1, 2, \dots$ then we have $\tau(x) = \infty$. This is a stopping time because $\tau(x) = n$ if and only if

$$\mathbf{E}_{\mathcal{F}_k} f(x) \leq s, \text{ for } k = 0, 1, \dots, n-1$$

and

$$\mathbf{E}_{\mathcal{F}_n} f(x) > s$$

These conditions define an event in $\mathcal{F}_n$.

The key observation is that $\mathbf{E}_{\mathcal{F}_\tau} f(x) > s$ on the set $\{x; \tau(x) < \infty\} = \{x; M_m f(x) > s\}$. So,

$$\lambda(\{x; M_m f(x) > s\}) = \lambda(\{x; \mathbf{E}_{\mathcal{F}_\tau} f(x) > s\}) \leq s^{-1} \|\mathbf{E}_{\mathcal{F}_\tau} f\|_1 \leq s^{-1} \|f\|_1.$$

This is the analogue of Theorem 2. There are interesting parallels between, for example the use of *longest* intervals in the Vitali Covering lemma and the stopping time being the *first* time that $\mathbf{E}_{\mathcal{F}_n} f(x) > s$. Note that this argument does not show that $M_m f \in L^1$ (a false statement in general) because the stopping time τ depends on s.

If we were to follow the strategy that we adopted for the Hardy–Littlewood maximal function, we would consider the set of all dyadic intervals such that

$$\frac{1}{\lambda(I)} \int_I f(t)dt > s.$$

But any two dyadic intervals are either disjoint or nested. It follows that the union of all such intervals is also a disjoint union of a certain subcollection of intervals, in fact the ones corresponding to the stopping time.

1.3 Fundamental Theorem of Calculus

THEOREM 3 *Let $f \in L^1(\mathbb{R})$ and define $F(x) = \int_0^x f(t)dt$. Then $F'(x)$ exists and equals $f(x)$ for almost all $x \in \mathbb{R}$.*

Proof. Let $\epsilon > 0$ and write $f = g + h$ where $g \in C_c(\mathbb{R})$ and $h \in L^1$ with $\|h\|_1 < \epsilon$. Let us also define $G(x) = \int_0^x g(t)dt$ and $H(x) = \int_0^x h(t)dt$. Then $F(x) = G(x) + H(x)$. Now consider

$$\limsup_{t\to 0} \left| \frac{F(x+t) - F(x)}{t} - f(x) \right|$$
$$\leq \limsup_{t\to 0} \left| \frac{G(x+t) - G(x)}{t} - g(x) \right| + \limsup_{t\to 0} \left| \frac{H(x+t) - H(x)}{t} - h(x) \right|$$

Now, the first $\limsup$ on the right is zero, by the Fundamental Theorem of Calculus, because g is continuous. So,

$$\begin{aligned}\limsup_{t\to 0}\left|\frac{F(x+t)-F(x)}{t}-f(x)\right| &\le \limsup_{t\to 0}\left|\frac{H(x+t)-H(x)}{t}-h(x)\right|\\ &\le |h(x)|+\sup_{t\neq 0}\left|\frac{H(x+t)-H(x)}{t}\right|\\ &\le |h(x)|+\sup_{t\neq 0}t^{-1}\int_x^{x+t}|h(s)|ds\\ &\le |h(x)|+\sup_{t>0}t^{-1}\int_{x-t}^{x+t}|h(s)|ds\\ &= |h(x)|+2M|h|(x)\end{aligned}$$

Let $\delta>0$ and consider the set

$$A_\delta=\{x;\limsup_{t\to 0+}\left|\frac{F(x+t)-F(x)}{t}-f(x)\right|>\delta\}.$$

Now, if $x\in A_\delta$, then either $|h(x)|>\frac{1}{3}\delta$ or $M|h|(x)>\frac{1}{3}\delta$. The first possibility occurs on a set of measure at most $3\epsilon\delta^{-1}$ (by the Tchebychev Inequality) and the second on a set of measure at most $12\epsilon\delta^{-1}$ by Theorem 2. So, the measure of A_δ is bounded by $15\epsilon\delta^{-1}$. But A_δ does not depend on ϵ, so letting $\epsilon\longrightarrow 0+$, we find that A_δ is a null set. Finally, taking a sequence of positive δs converging to zero, we find that $\{x;\limsup_{t\to 0+}\left|\frac{F(x+t)-F(x)}{t}-f(x)\right|>0\}$ is also a null set. The result follows. ■

1.4 Homogenous operators with nonnegative kernel

We will prove the following result for the homogenous integral operator on the half line.

$$Tf(x)=\int_0^\infty K(x,y)f(y)dy$$

where x runs over $]0,\infty[$. We assume that the kernel function K is nonnegative and satisfies the homogeneity condition

$$K(tx,ty)=t^{-1}K(x,y).$$

We will also need to know that K is Lebesgue measurable on the positive quadrant. We further define

$$\begin{aligned}
C_p = \int_0^\infty t^{-\frac{1}{p'}} K(t,1)dt &= \int_0^\infty t^{-\frac{1}{p'}} K(1,t^{-1})t^{-1}dt \\
&= \int_0^\infty s^{\frac{1}{p'}} K(1,s)s^{-1}ds \\
&= \int_0^\infty s^{-\frac{1}{p}} K(1,s)ds
\end{aligned}$$

and we will assume that C_p is finite.

Now comes a remarkable idea. For f and g nonnegative functions we have using Hölder's Inequality and Tonelli's Theorem

$$\begin{aligned}
&\int_0^\infty\int_0^\infty K(x,y)g(x)f(y)dxdy \\
&= \int_0^\infty\int_0^\infty \left(\frac{x}{y}\right)^{-\frac{1}{pp'}} K(x,y)^{\frac{1}{p}} f(y) \left(\frac{y}{x}\right)^{-\frac{1}{pp'}} K(x,y)^{\frac{1}{p'}} g(x)dxdy \\
&\leq \left\{ \int_0^\infty\int_0^\infty \left(\frac{x}{y}\right)^{-\frac{1}{p'}} K(x,y)f(y)^p dxdy\right\}^{\frac{1}{p}} \left\{ \int_0^\infty\int_0^\infty \left(\frac{y}{x}\right)^{-\frac{1}{p}} K(x,y)g(x)^{p'} dxdy\right\}^{\frac{1}{p'}}
\end{aligned}$$

We now have

$$\begin{aligned}
&\int_0^\infty\int_0^\infty \left(\frac{x}{y}\right)^{-\frac{1}{p'}} K(x,y)f(y)^p dxdy \\
&= \int_{y=0}^\infty \left\{ \int_{x=0}^\infty K\left(\frac{x}{y},1\right)\left(\frac{x}{y}\right)^{-\frac{1}{p'}} y^{-1}dx \right\} f(y)^p dy \\
&= C_p \|f\|_p^p
\end{aligned}$$

and similarly

$$\int_0^\infty\int_0^\infty \left(\frac{y}{x}\right)^{-\frac{1}{p}} K(x,y)g(x)^{p'} dxdy = C_p\|g\|_{p'}^{p'}$$

resulting in

$$\int_0^\infty\int_0^\infty K(x,y)g(x)f(y)dxdy \leq C_p\|f\|_p\|g\|_{p'}$$

It follows from this by duality that $\|Tf\|_p \leq C_p\|f\|_p$ for f nonnegative and the same inequality for signed or complex f then follows.

Some specific cases of interest are

- $Tf(x) = x^{-1}\int_0^x f(y)dy$, $C_p = \dfrac{p}{p-1}$.
- $Tf(x) = \int_x^\infty y^{-1}f(y)dy$, $C_p = p$.
- $Tf(x) = \int_0^\infty (x+y)^{-1}f(y)dy$, $C_p = \pi\mathrm{cosec}\left(\dfrac{\pi}{p}\right)$.

1.5 The distribution function and the decreasing rearrangement

Suppose that we have a nonnegative measurable function f defined on a measure space $(X, \mathcal{M}, \mu)$ which has the additional property that there exists $t > 0$ such that $\mu(\{x; f(x) > t\})$ is finite. If f is not nonnegative, then we replace it by $|f|$, its absolute value. The distribution function $f_\star$ is now defined on $]0, \infty[$ by

$$f_\star(t) = \mu(\{x; f(x) > t\}).$$

LEMMA 4 *The distribution function has the following properties.*

1. $f_\star$ *takes values in* $[0, \infty]$.
2. $f_\star$ *is decreasing and right continuous on* $]0, \infty[$.
3. $f_\star(t) \to 0$ *as* $t \to \infty$.
4. *If* $f \leq g$ *pointwise, then* $f_\star \leq g_\star$ *pointwise.*
5. *If* $g(x) = af(x)$ *for all* $x \in X$ *with* $a > 0$ *constant, then* $g_\star(t) = f_\star(a^{-1}t)$ *for all* $t > 0$.
6. *If* $h(x) = f(x) + g(x)$ *for all* $x \in X$, *then* $h_\star(t+s) \leq f_\star(t) + g_\star(s)$ *for all* $t, s > 0$.
7. *If* $f_n \uparrow f$ *pointwise on* X, *then* $f_{n\star} \uparrow f_\star$ *pointwise on* $]0, \infty[$.

The proofs are straightforward. The Monotone convergence theorem is used to establish the last assertion as well as the right continuity in the second assertion. Note that $f_\star(t)$ can be infinite.

The next step is to define

$$f^\star(s) = \inf\{t; f_\star(t) \leq s\}$$

In fact, the infimum is attained. To see this, let $\tau = f^\star(s)$. Then we have $t_n \downarrow \tau$ with $f_\star(t_n) \leq s$. But $f_\star$ is right continuous, so $f_\star(\tau) \leq s$. This means incidentally that we also have

$$f^\star(s) = \sup\{t; f_\star(t) > s\} \tag{1.4}$$

To see this, let $\tau = f^\star(s)$. Then $f_\star(t) > s$ if and only if $t < \tau$.

Next, we use this fact to show that $f^\star$ is right continuous. Let $s_n \downarrow s$ and again, let $\tau = f^\star(s)$. Let $\tau_n = f^\star(s_n)$. Then $\tau_n \leq \tau$ and τ_n increases with n. The danger is that $\sup_n \tau_n < \tau$. Now, let $\epsilon > 0$ then from (1.4), $\exists t$ with $f_\star(t) > s$ and $t > \tau - \epsilon$. So for n large enough, $f_\star(t) > s_n$ and it follows that $\tau_n \geq t > \tau - \epsilon$. It may also help to observe that $f^\star$ is the distribution function of $f_\star$ for Lebesgue measure on the half line.

LEMMA 5 *The function $f^\star$ has the following properties.*

1. *$f^\star$ takes values in $[0, \infty[$.*
2. *$f^\star$ is decreasing and right continuous on $]0, \infty[$.*
3. *If $f \leq g$ pointwise, then $f^\star \leq g^\star$ pointwise.*
4. *If $g(x) = af(x)$ for all $x \in X$ with $a > 0$ constant, then $g^\star(t) = af^\star(t)$ for all $t > 0$.*
5. *If $h(x) = f(x) + g(x)$ for all $x \in X$, then $h^\star(t+s) \leq f^\star(t) + g^\star(s)$ for all $t, s > 0$.*
6. *If $f_n \uparrow f$ pointwise on X, then $f_n^\star \uparrow f^\star$ pointwise on $]0, \infty[$.*
7. *$(f^\star)_\star = f_\star$*

The last assertion expresses the fact that f and $f^\star$ have the same distribution function. The function $f^\star$ is called the (equimeasurable) decreasing rearrangement of f. Although technically it may not be a rearrangement of f, it behaves as if it were. It should also be clear that if f_1 and f_2 are nonnegative decreasing right continuous functions on $]0,\infty[$, both having the same distribution function then $f_1 = f_2$. Thus if you can construct somehow a nonnegative decreasing right continuous function g on $]0,\infty[$ with the same distribution function as f, then $g = f^\star$. Note also that the intervals of constancy of $f_\star$ correspond to the jump discontinuities of $f^\star$ and vice-versa.

LEMMA 6 *Let f be a nonnegative measurable function on $(X, \mathcal{M}, \mu)$, finite almost everywhere and with the additional property. Let $M \in \mathcal{M}$ be a set of finite measure. Then*

$$\int_M f(x)d\mu(x) \le \int_{t=0}^{\mu(M)} f^\star(t)dt$$

Proof. We approximate f with an increasing sequence of measurable step functions. It is clear that the result need only be proved in case that f is a measurable step function. So, we write

$$f = \sum_{k=1}^{n} a_k \mathbb{1}_{E_k}$$

where $a_k > 0$ and the E_k are nested measurable sets $E_1 \supset E_2 \supset E_3 \ldots$ Then

$$f^\star(t) = \sum_{k=1}^{n} a_k \mathbb{1}_{[0,\mu(E_k)[}(t).$$

Now, we find

$$\int_M f(x)d\mu(x) = \sum_{k=1}^{n} a_k \mu(M \cap E_k)$$

and

$$\int_{t=0}^{\mu(M)} f^\star(t)dt = \sum_{k=1}^{n} a_k \mu([0,\mu(M)[\cap[0,\mu(E_k)[) = \sum_{k=1}^{n} a_k \min(\mu(M), \mu(E_k))$$

and the result follows. ■

LEMMA 7 *Let f, g be nonnegative measurable functions on $(X, \mathcal{M}, \mu)$, finite almost everywhere and with the additional property. Then*

$$\int f(x)g(x)d\mu(x) \leq \int_{t=0}^{\infty} f^{\star}(t)g^{\star}(t)dt$$

Proof. We apply the same approximation idea as in proof of the previous lemma. Then, we need to show that

$$\int \left(\sum_{k=1}^{n} a_k \mathbb{1}_{E_k} \right) g(x)d\mu(x) \leq \int_{t=0}^{\infty} \left(\sum_{k=1}^{n} a_k \mathbb{1}_{[0,\mu(E_k)[}(t) \right) g^{\star}(t)dt$$

or by estimating term by term

$$\int_{E_k} g(x)d\mu(x) \leq \int_{t=0}^{\mu(E_k)} g^{\star}(t)dt$$

but this is just the content of that lemma. ■

EXAMPLE Note that on infinite measure spaces, decreasing rearrangements may not behave quite the way that one would expect. For example on $[0, \infty[$ with Lebesgue measure, let $f(t) = 1 - e^{-t}$, then $f^{\star}(t) = 1$ for all $t > 0$. So, $f^{\star}$ is not a rearrangment of f and it is clear that one cannot hope to achieve this. □

LEMMA 8 *Let f be nonnegative a measurable function on $(X, \mathcal{M}, \mu)$, finite almost everywhere and with the additional property. Then*

$$\int_X f(x)d\mu(x) = \int_0^{\infty} f^{\star}(t)dt$$

Proof. We verify the lemma for step functions. Then, given an arbitrary function f approximate it by an increasing sequence of step functions. ■

LEMMA 9 *Let f be nonnegative a measurable function on $(X, \mathcal{M}, \mu)$, finite almost everywhere and with the additional property. Let $\varphi : [0, \infty[\to [0, \infty[$ be an increasing function continuous on the left. Let $g = \varphi \circ f$, then $g^{\star} = \varphi \circ f^{\star}$.*

Proof. Since $\varphi \circ f^\star$ is decreasing and right continuous, it is enough to show that it has the correct distribution function. So, we must show that the sets $\{\varphi \circ f^\star > t\}$ and $\{\varphi \circ f > t\}$ have the same measures. It is enought to show that $\varphi^{-1}(]t, \infty[$ is an interval of the form $]s, \infty[$. This is a consequence of the left continuity of φ. If for example $\varphi^{-1}(]t, \infty[= [s, \infty[$, then choose $s_n \uparrow s$ strictly increasing, then $\varphi(s_n) \leq t$ and it follows that $\varphi(s) \leq t$ contradicting $s \in \varphi^{-1}(]t, \infty[$. ∎

A typical consequence of the last two results is that

$$\int_X f(x)^p d\mu(x) = \int_0^\infty (f^\star(t))^p dt$$

for $p > 0$. In particular, $\|f\|_p = \|f^\star\|_p$ for $1 \leq p \leq \infty$.

LEMMA 10 *On a finite non-atomic measure space of total measure m, let f be a nonnegative function finite almost everywhere. Let $0 \leq t \leq m$. Then there is a measurable subset E_t such that*

$$\int_{E_t} f(x) d\mu(x) = \int_0^t f^\star(u) du.$$

Also, it may be arranged that E_t is increasing with t.

Proof. There are two cases. The easy case is where there exists $s > 0$ such that $f_\star(s) = t$. In this case, it will suffice to take $E_t = \{x; f(x) > s\}$. The reason is that

$$(\mathbb{1}_{E_t} f)^\star = \mathbb{1}_{[0,t[} f^\star$$

and this in turn follows from $[0, t[= \{f^\star > s\}$. We leave the details to the reader.

In the remaining case, t lies in an interval of constancy of $f^\star$. We have $t_0 < t \leq t_1 < \infty$ where t_0 and t_1 are the measures of the sets $T_0 = \{f > s\}$ and $T_1 = \{f \geq s\}$ respectively. The finiteness of the measure space is used to assert that t_1 is finite. Now, using the fact that the measure space is non-atomic, we may construct a measurable subset E_t with $T_0 \subset E_t \subseteq T_1$ with measure exactly t. Again, this is an exercise.

We now get

$$\int_{E_t} f(x) d\mu(x) = \int_{T_0} f(x) d\mu(x) + \int_{E_t \setminus T_0} f(x) d\mu(x)$$

$$
\begin{aligned}
&= \int_0^{t_0} f^\star(u)du + s(t-t_0) \\
&= \int_0^{t} f^\star(u)du
\end{aligned}
$$

as required. ■

THEOREM 11 *Let f, g be nonnegative measurable functions on a finite non-atomic measure space $(X, \mathcal{M}, \mu)$, finite almost everywhere. Then there is a non-negative measurable function $\tilde{g}$ equimeasurable with g such that*

$$
\int_X f(x)\tilde{g}(x)d\mu(x) = \int_0^\infty f^\star(t)g^\star(t)dt \tag{1.5}
$$

Proof. We do not need to consider g, we start from $g^\star$ which is equimeasurable with g and produce a sequence of step functions $h_n = \varphi_n \circ g^\star$ where

$$
\varphi_n(x) = \begin{cases} 2^{-n}\lfloor 2^n x \rfloor & \text{if } x < 2^n, \\ 0 & \text{if } x \geq 2^n. \end{cases}
$$

so that $h_n \uparrow g^\star$. Writing each $h = h_n$ in the form $h = \sum_{j=1}^{J_n} a_j \mathbb{1}_{[0,t_j[}$ where t_j is increasing in j and $a_j > 0$ (actually $a_j = 2^{-n}$ for all j and $J_n = 4^n$). Then, using the previous lemma, we can construct E_{t_j} measurable sets such that

$$
\tilde{h} = \sum_{j=1}^{J_n} a_j \mathbb{1}_{E_{t_j}}
$$

has the property

$$
\int_X f(x)\tilde{h}(x)d\mu(x) = \int_0^\infty f^\star(t)h(t)dt.
$$

But a moment's thought shows that $\widetilde{h_n}$ is itself an increasing sequence of functions converging to a function $\tilde{g}$ equimeasurable with $g^\star$ and hence also with g. An application of the monotone convergence theorem then yields (1.5) as required.

There is also an infinite measure version of this result.

THEOREM 12 *Let f, g be nonnegative measurable functions on a σ-finite non-atomic measure space $(X, \mathcal{M}, \mu)$, finite almost everywhere and satisfying the additional condition. Let $\epsilon > 0$. Then there is a nonnegative measurable function $\tilde{g}$ equimeasurable with g such that*

$$\int_X f(x)\tilde{g}(x)d\mu(x) > \int_0^\infty f^\star(t)g^\star(t)dt - \epsilon$$

Proof. Let X_n be increasing measurable sets of finite measure such that $X = \bigcup_{n=1}^\infty X_n$. Then, applying the previous result, there exist functions h_n carried by X_n and equimeasurable with $g\mathbb{1}_{X_n}$ such that

$$\int_{X_n} f(x)h_n(x)d\mu(x) = \int_0^{\mu(X_n)} (f\mathbb{1}_{X_n})^\star(t)(g\mathbb{1}_{X_n})^\star(t)dt.$$

Now build $\widetilde{g_n}$ by $\widetilde{g_n} = h_n + \mathbb{1}_{X\setminus X_n} g$. Then $\widetilde{g_n}$ is clearly equimeasurable with g and we have

$$\int f(x)\widetilde{g_n}(x)d\mu(x) \geq \int_0^{\mu(X_n)} (f\mathbb{1}_{X_n})^\star(t)(g\mathbb{1}_{X_n})^\star(t)dt.$$

But by monotone convergence, the right-hand side increases to $\int_0^\infty f^\star(t)g^\star(t)dt$ and the result now follows. ■

There are examples such as $g(x) = 1 - e^{-x}$, f any nice strictly positive function where equality cannot be obtained. We also have immediately

COROLLARY 13 *On a σ-finite non-atomic measure space let f be a nonnegative function finite almost everywhere satisfying the additional condition. Let $t \geq 0$ be finite. Then*

$$\sup_{\mu(E)=t} \int_E f(x)d\mu(x) = \int_0^t f^\star(u)du.$$

We can now recover a form of subadditivity for the decreasing rearrangement.

THEOREM 14 *On a σ-finite measure space let f and g be nonnegative functions finite almost everywhere satisfying the additional condition. Let $t \geq 0$ be finite and $h = f + g$. Then*

$$\int_0^t h^\star(u)du \leq \int_0^t f^\star(u)du + \int_0^t g^\star(u)du.$$

Proof. First assume that the measure space is non-atomic. Then

$$\begin{aligned}\int_0^t h^\star(u)du &= \sup_{\mu(E)=t} \int_E h(x)d\mu(x) \\ &= \sup_{\mu(E)=t} \int_E \Big(f(x)+g(x)\Big)d\mu(x) \\ &\leq \sup_{\mu(E)=t} \int_E f(x)d\mu(x) + \sup_{\mu(E)=t} \int_E g(x)d\mu(x) \\ &= \int_0^t f^\star(u)du + \int_0^t g^\star(u)du\end{aligned}$$

as required. If the measure space is atomic, we simply take the product with $[0,1]$ and Lebesgue measure. ■

Now usually, the theorm is not left in this form. We assume $t > 0$ and divide by t. This yields

$$A((f+g)^\star) \leq A(f^\star) + A(g^\star)$$

where A is Hardy's averaging operator. It is easy to see that $A(f^\star)$ is a continuous decreasing function with $A(f^\star) \geq f^\star$.

EXAMPLE Let $f = \mathbb{1}_{[0,1[}$ and $g = \mathbb{1}_{[1,2[}$. Then $h = f + g = \mathbb{1}_{[0,2[}$. We have $f^\star = g^\star = f$ and $h^\star = h$. It is false that

$$h^\star \leq f^\star + g^\star = 2f^\star$$

because the left hand side is positive on $[1, 2[$ while the right hand side is zero. However, after averaging the situation changes.

$$(Ah^\star)(t) = \begin{cases} 1 & \text{if } 0 < t \leq 2, \\ 2t^{-1} & \text{if } t \geq 2. \end{cases}$$

On the other hand

$$2(Af^\star)(t) = \begin{cases} 2 & \text{if } 0 < t \leq 1, \\ 2t^{-1} & \text{if } t \geq 1. \end{cases}$$

and indeed it is true that $(Ah^\star)(t) \leq 2(Af^\star)(t)$ for $t > 0$. □

In this situation, the Hardy operator bears a strong resemblance to the maximal operator.

We finish this chapter with a proof of a special case of the Marcinkiewicz Interpolation Theorem that allows us to conclude that the maximal operator is bounded on L^p for $1 < p \leq \infty$. We need the following lemma which allows to compute an L^p norm from the distribution function.

LEMMA 15 *Let f be a positive measurable function and $1 < p < \infty$. Then*

$$\|f\|_p^p = p \int_0^\infty t^{p-1} f_\star(t) dt.$$

Proof. We approximate f by an increasing sequence of step functions. By monotone convergence, it is enough to prove the result in the special case that f is a step function. Let $f = \sum_{j=1}^J t_j \mathbb{1}_{E_j}$ where t_j are positive and decreasing and the E_j are disjoint measurable sets. Then clearly

$$\|f\|_p^p = \sum_{j=1}^J t_j^p \mu(E_j).$$

On the other hand

$$f_\star(s) = \begin{cases} \sum_{j=1}^k \mu(E_j) & \text{if } t_{k+1} \leq s < t_k,\ 1 \leq k < J, \\ \sum_{j=1}^J \mu(E_j) & \text{if } s < t_J \text{ (case } k = J, \text{ interpret } t_{J+1} = 0), \\ 0 & \text{if } t_1 \leq s \text{ (case } k = 0, \text{ interpret } t_0 = \infty). \end{cases}$$

We get, omitting the term $k = 0$ since $f_\star(s) = 0$ in that case,

$$\begin{aligned} p \int_0^\infty t^{p-1} f_\star(t) dt &= p \sum_{k=1}^J \sum_{j=1}^k \mu(E_j) \int_{t_{k+1}}^{t_k} t^{p-1} dt \\ &= p \sum_{j=1}^J \sum_{k=j}^J \mu(E_j) \int_{t_{k+1}}^{t_k} t^{p-1} dt \\ &= p \sum_{j=1}^J \mu(E_j) \int_{t_{J+1}}^{t_j} t^{p-1} dt \\ &= \sum_{j=1}^J t_j^p \mu(E_j). \end{aligned}$$

since $t_{j+1} = 0$. The two expressions agree. ∎

THEOREM 16 *Let T be a sublinear operator (defined on $L^\infty + L^1$) that maps L^∞ to L^∞ with norm 1 and satisfies the estimate*

$$\mu(\{|Tf| > s\}) \le \frac{C\|f\|_1}{s}.$$

Then T is bounded on all L^p spaces for $1 < p < \infty$.

Proof. We take $t > 0$ and a function f and split it as $f = f_t + f^t$ where

$$f_t = \mathbb{1}_{X_t} f, \quad f^t = \mathbb{1}_{Y_t} f, \quad X_t = \{|f| \le t\}, \quad Y_t = \{|f| > t\}.$$

Then $|Tf| \le |Tf_t| + |Tf^t|$ and $|Tf_t| \le t$. So

$$\{|Tf| > 2t\} \subseteq \{|Tf^t| > t\}$$

and $|Tf|_\star(2t) \le Ct^{-1}\|f^t\|_1$. Note that

$$\|f^t\|_1 = \int_{f^\star(s)>t} f^\star(s) ds,$$

by applying Lemma 9 with φ the increasing left continuous function

$$\varphi(s) = \begin{cases} 0 & \text{if } s \le t, \\ s & \text{if } s > t. \end{cases}$$

We get

$$\begin{aligned}
\|Tf\|_p^p &= p\int_{t=0}^\infty t^{p-1}|Tf|_\star(t)dt \\
&= p2^p \int_{t=0}^\infty t^{p-1}|Tf|_\star(2t)dt \\
&\le p2^p \int_{t=0}^\infty t^{p-2}\|f^t\|_1 dt \\
&= p2^p \int_{t=0}^\infty t^{p-2} \int_{f^\star(s)>t} f^\star(s) ds dt \\
&= p2^p \int_{s=0}^\infty f^\star(s) \int_{t=0}^{f^\star(s)} t^{p-2} dt ds \\
&= \frac{p}{p-1} 2^p \int_{s=0}^\infty f^\star(s) f^\star(s)^{p-1} ds \\
&= p' 2^p \int_{s=0}^\infty f^\star(s)^p ds \\
&= p' 2^p \|f\|_p^p
\end{aligned}$$

and conclude that $\|Tf\|_p \leq 2(p')^{\frac{1}{p}}\|f\|_p$. ■

2

Harmonic Analysis — the LCA setting

2.1 Gelfand's Theory of Commutative Banach Algebra

A commutative Banach algebra A is a Banach space together with a continuous multiplication so that A becomes a linear commutative associative algebra. The continuity of the multiplication amounts to the existence of a constant C such that

$$\|xy\| \leq C\|x\|\|y\|, \quad \forall x, y \in A.$$

The algbra A is said to be unital if it has an identity element which we will denote $\mathbb{1}_A$. In a unital algebra, it may be false that $\|\mathbb{1}_A\| = 1$, but we can always renorm the algebra with an equivalent norm that has this property. For this we use the multiplier norm

$$\|x\|_M = \sup_{\|y\| \leq 1} \|xy\|.$$

While in general, this may fail to define an equivalent norm, in this case it does because

$$\|x\|_M = \sup_{\|y\| \leq 1} \|xy\| \leq \sup_{\|y\| \leq 1} C\|x\|\|y\| \leq C\|x\|$$

and

$$\|x\| = \|x\mathbb{1}_A\| \leq \|\mathbb{1}_A\|\|x\|_M.$$

Generally we will therefore work with a norm that has the property that $\|\mathbb{1}_A\| = 1$. The multiplier norm has an even more important property, namely that

$$\|xy\|_M \leq \|x\|_M\|y\|_M,$$

in other words we may always assume without loss of generality that $C = 1$ (at least if we are only interested in properties that are preserved under norm equivalence). From now on we are interested only in the case of unital commutative banach algebras over $\mathbb{C}$. The real case presents substantially more difficulties.

The spectrum of an element $x \in A$ is a subset of the complex plane defined by

$$\sigma(x) = \{\lambda; \lambda \in \mathbb{C}, (\lambda \mathbb{1}_A - x)^{-1} \text{ fails to exist in } A\}.$$

The spectrum has the following properties

1. If $\lambda \in \sigma(x)$ implies $|\lambda| \leq \|x\|$.
2. $\sigma(x)$ is closed.
3. $\sigma(x)$ is nonempty.

The proofs are easy. First if $\lambda > \|x\|$, then we can construct

$$(\mathbb{1}_A - \lambda^{-1}x)^{-1} = \sum_{n=0}^{\infty} \lambda^{-n} x^n$$

the right hand side being a norm convergent sum. It follows then that $(\lambda \mathbb{1}_A - x)^{-1}$ exists.

The second assertion is similar. If $\mu \notin \sigma(x)$, then $(\mu \mathbb{1}_A - x)^{-1}$ exists. Now we consider λ very close to μ and observe

$$(\lambda \mathbb{1}_A - x) = (\lambda - \mu)\mathbb{1}_A + (\mu \mathbb{1}_A - x) = \Big(\mathbb{1}_A + (\lambda - \mu)(\mu \mathbb{1}_A - x)^{-1}\Big)(\mu \mathbb{1}_A - x).$$

Provided $|\lambda - \mu| \|(\mu \mathbb{1}_A - x)^{-1}\| < 1$, it will be possible to construct $(\lambda \mathbb{1}_A - x))^{-1}$ with a geometric series argument. So, the complement of $\sigma(x)$ is open and therefore $\sigma(x)$ is closed.

For the third assertion, suppose the contrary. Then $(\lambda \mathbb{1}_A - x)^{-1}$ exists for all complex λ. Let u be a continuous linear functional on A and consider the complex-valued function

$$\lambda \mapsto u((\lambda \mathbb{1}_A - x)^{-1})$$

It is clear (actually by using the arguments that we have used in proving the first two assertions) that this is a holomorphic function in the whole complex plane (a so-called entire function) and also that it tends to zero at infinity since

$$\|(\lambda \mathbb{1}_A - x)^{-1}\| = \left\| \sum_{n=0}^{\infty} \lambda^{-n-1} x^n \right\| \leq |\lambda|^{-1}(1 - |\lambda|^{-1}\|x\|)^{-1} = (|\lambda| - \|x\|)^{-1}.$$

It follows from the maximum modulus principle that such a function is identically zero. It then follows from the Hahn–Banach Theorem that $(\lambda \mathbb{1}_A - x)^{-1} = 0$ for all λ which is complete nonsense since inverses can never be zero.

Having dealt with the spectrum, we now turn to the ideal structure of A. An ideal I is said to be proper if $I \subset A$. We assert that every proper ideal is contained in a maximal proper ideal. This is proved using a Zorn's Lemma argument. It is enough to show that every chain of proper ideals has an upper bound under set inclusion. Given a chain $\mathcal{C}$ of proper ideals, one simply takes

$$B = \bigcup_{I \in \mathcal{C}} I.$$

It is easy to see that B is an ideal. If it is not proper, then $\mathbb{1}_A \in B$. But then there exists $I \in \mathcal{C}$ such that $\mathbb{1}_A \in I$ contradicting the fact that I is proper. (As soon as $\mathbb{1}_A \in I$, then $x = x\mathbb{1}_A \in I$ for every $x \in A$.)

The similar argument shows that every maximal proper ideal is closed. If M is a maximal proper ideal, then it is clear that $\mathrm{cl}(M)$ is an ideal. So either $M = \mathrm{cl}(M)$ or $\mathrm{cl}(M)$ is not proper. In other words, either M is closed or M is dense. But the latter situation is not possible, since then we would be able to approximate $\mathbb{1}_A$ with elements of M. But any element of A sufficiently close to $\mathbb{1}_A$ is invertible (by the geometric series argument yet again) and so M would have to contain invertible elements and hence $\mathbb{1}_A$ itself, contradicting the fact that M is proper.

Now let M be a maximal proper ideal and consider $Q = A/M$. Then it is routine to check that Q is a unital commutative Banach algebra in the quotient norm. Also, from ring theory, it cannot contain any ideals other than the zero ideal and Q itself. (Let π be the canonical projection $\pi : A \to Q$ and let J be a nontrivial ideal of Q, then $\pi^{-1}(J)$ is a proper ideal of A strictly bigger than M — a contradiction). This implies in turn that every non-zero element of Q is invertible. We claim that $Q = \mathbb{C}\mathbb{1}_Q$. Indeed, let $x \in Q$ be arbitrary and let $\lambda \in \sigma(x)$. Then $\lambda\mathbb{1}_Q - x$ fails to be invertible and must therefore be zero. So $x = \lambda\mathbb{1}_Q$.

This means then that every maximal proper ideal has codimension 1 and is the kernel of a continuous linear form $\varphi : A \to \mathbb{C}$. We are free to normalize φ such that $\varphi(\mathbb{1}_A) = 1$. But now, let $x, y \in A$ then $x - \varphi(x)\mathbb{1}_A$ and $y - \varphi(y)\mathbb{1}_A$ are elements of M since they are clearly in the kernel of φ. But now $(x - \varphi(x)\mathbb{1}_A)y = xy - \varphi(x)y$ is in M and therefore also

$$xy - \varphi(x)\varphi(y)\mathbb{1}_A = xy - \varphi(x)y + \varphi(x)(y - \varphi(y)\mathbb{1}_A).$$

It follows that this element is in the kernel of φ and hence

$$\varphi(xy) = \varphi(x)\varphi(y).$$

Such a φ (with $\varphi(\mathbb{1}_A) = 1$) is called a multiplicative linear functional (mlf). Every maximal proper ideal is therefore the kernel of an mlf and conversely, it is obvious that the kernel of any mlf is a closed ideal of codimension one and hence a maximal proper ideal.

The next step in the saga is to define M_A the space of all mlfs and to give it a topology. This is the Gelfand topology and it is simply the relative topology inherited from the weak$\star$ ($\sigma(A', A)$ topology). It turns out that M_A is compact in this topology and it is also clearly Hausdorff.

To see this first of all we observe that any mlf has norm exactly one. Clearly $x - \varphi(x)\mathbb{1}_A$ is in the proper ideal $\ker(\varphi)$ and therefore, not invertible. So, $\varphi(x) \in \sigma(x)$ and hence $|\varphi(x)| \leq \|x\|$. On the other hand $\varphi(\mathbb{1}_A) = 1$ and $\|\mathbb{1}_A\| = 1$.

So M_A can be specified as the subset of the unit ball of A' which satisfies the following closed conditions

$$\begin{aligned} &\varphi(xy) = \varphi(x)\varphi(y) \qquad x, y \in A \\ &\varphi(\mathbb{1}_A) = 1 \end{aligned}$$

each depending on only finitely many elements from A (at most three).

Since the unit ball of A' is compact for the $\sigma(A', A)$ topology and since M_A is a $\sigma(A', A)$-closed subset of A' it follows that M_A is itself $\sigma(A', A)$ compact.

We now write $\hat{x}(\varphi) = \varphi(x)$ and observe that $\hat{x}$ is now a continuous function on M_A. The mapping $x \mapsto \hat{x}$ which maps from A to $C(M_A)$ is called the Gelfand transform of A and is an algebra homomorphism. It can happen that the Gelfand transform has a non-trivial kernel. We can even characterize the kernel of the Gelfand transform. It consists of all elements $x \in A$ such that $\sigma(x) = \{0\}$ or from power series considerations that

$$\limsup_{n \to \infty} \|x^n\|^{\frac{1}{n}} = 0.$$

It is also the Jacobson radical of A viewed as a ring.

Also it is rarely the case that the Gelfand transform is onto or that the uniform norm of $\hat{x}$ is equivalent to the norm of x. In many situations the space M_A is easy to understand, but there are also cases where its structure is totally mind boggling!

2.2 The non-unital case

We now come to the case that A is a complex commutative Banach algebra, but it does not have a unit (identity) element. In that case, we simply adjoin an identity element and use the theory in the previous section. So, the new algebra has elements

$$\tilde{A} = \{t\mathbb{1} + x; t \in \mathbb{C}, x \in A\}.$$

and we define

$$\begin{aligned}(t\mathbb{1} + x) + (s\mathbb{1} + y) &= (t+s)\mathbb{1} + (x+y)\\ (t\mathbb{1} + x)(s\mathbb{1} + y) &= ts\mathbb{1} + (ty + sx + xy)\end{aligned}$$

For the norm on $\tilde{A}$ we simply take $\|t\mathbb{1} + x\| = |t| + \|x\|$ and it is straightforward to verify that this is actually a norm. If multiplication is continuous on A, then it is also on $\tilde{A}$ and then one may replace this norm with the multiplier norm to get an equivalent submultiplicative norm. It's important to extend the norm to $\tilde{A}$ first. Taking the multiplier norm immediately does not work.

We now consider the ideals in A and we need to add an extra condition. Let I be an ideal in A. Then a modular unit (or modular identity) for I is an element $u \in A$ such that $x - ux \in I$ for all $x \in A$. When we form the quotient algebra A/I, the image of u will be an identity element. So, we say that an ideal is modular, if it possesses a modular unit and this is actually equivalent to A/I having an identity element. We now have the following lemma.

LEMMA 17 *Let I be a modular ideal in A. Then there exists an ideal J in $\tilde{A}$ such that $J \not\subseteq A$ and $I = J \bigcap A$.*

Conversely, if J is an ideal in $\tilde{A}$ such that $J \not\subseteq A$, then $I = J \bigcap A$ is a modular ideal in A.

Proof. For the first assertion, let I be a modular ideal of A with modular unit u. Define $J = \{x; x \in \tilde{A}, xu \in I\}$, clearly an ideal of $\tilde{A}$. Since $u \in A$, $u - u^2 \in I$, i.e. $(\mathbb{1} - u)u \in I$. So, $\mathbb{1} - u \in J$. But $\mathbb{1} - u \notin A$, so $J \not\subseteq A$. It remains to show that $I = J \bigcap A$ and clearly $I \subseteq J \bigcap A$. So, let $x \in J \bigcap A$. then since $x \in J$, we have $xu \in I$ and since $x \in A$, we have $x - xu \in I$. Therfore $x = (x - xu) + xu \in I$. This completes the proof of the first assertion.

For the converse, it is clear that $I = J \bigcap A$ is an ideal in A. But $J \not\subseteq A$, so there is an element of J of the form $\mathbb{1} - u$ with $u \in A$. Thus, for $x \in A$,

$x - xu = x(\mathbb{1} - u) \in J$. But also x and xu are both elements of A and hence so is $x - xu$. Thus $x - xu \in I$. We have shown that u is a modular unit for I. ■

The consequence of this correspondence is that the maximal modular ideals of A are in one-to-one correspondence with the maximal ideals of $\tilde{A}$ that are not contained in A. But since A is itself a maximal ideal in $\tilde{A}$ because it has codimension one, the maximum modular ideal space (also denoted M_A) of A is just the maximal ideal space of $M_{\tilde{A}}$ with a single point removed φ_0. We view this point as a "point at infinity", so that M_A is a locally compact Hausdorff space having $M_{\tilde{A}}$ as its one-point compactification. Of course, φ_0 is an mlf on $\tilde{A}$ vanishing on A and hence must be given by

$$\varphi_0(\lambda \mathbb{1} + x) = \lambda$$

for $x \in A$. Every other mlf on $\tilde{A}$ restricts to a (non-zero) mlf on A and conversely, every mlf on A extends to a unique mlf on $\tilde{A}$. For $x \in A$, we have $\varphi(x) \to \varphi_0(x) = 0$ as $\varphi \to \varphi_0$ in the $M_{\tilde{A}}$ topology, so its Gelfand transform $\hat{x}$ viewed as a function on M_A vanishes at infinity. We see that $x \mapsto \hat{x}$ is a continuous algebra homomorphism from A to $C_0(M_A)$.

2.3 Finding the Maximal Ideal Space

Usually this is either very easy or totally impossible.

EXAMPLE Let $A = C^1([0, 1])$ the space of continuously differentiable functions on the unit interval. It's clearly an algebra with identity and the multiplication is continuous. Clearly the point evaluations $f \mapsto f(t)$ are mlfs for $t \in [0, 1]$. It seems reasonable that these would be the only ones. How do we prove this? Let I be some maximal ideal not of this form. Then, for each $t \in [0, 1]$ there is a function $f_t \in I$ with $f_t(t) = 1$. Let $U_t = \{s; s \in [0, 1], |f_t(s)| > \frac{1}{2}\}$. This is a neighbourhood of t. Applying compactness we have $t_1, \ldots, t_N$ such that U_{t_n} cover $[0, 1]$ for $n = 1, \ldots, n$. Now, make the function f as

$$f = \sum_{n=1}^{N} \overline{f_{t_n}} f_{t_n}$$

and observe that $f > \frac{1}{4}$ everywhere on $[0, 1]$. So the reciprocal $\mathbb{1}/f$ is in C^1. But f is in I and hence so is $\mathbb{1}$. But this means that $I = A$ a contradiction. We leave the reader to chack that the Gelfand topology is just the standard topology on $[0, 1]$. □

EXAMPLE Here is another example, very different. Let K be a compact subset of $\mathbb{C}$. Consider the ring of polynomials, viewed as functions on $\mathbb{C}$, restrict them to K and let A be the uniform closure. Then A is a closed subalgebra of $C(K)$ and it has an identity. The key to this algebra is that it is singly generated. We denote by z the identity function on K. Then everything in A is limit of polynomials in z. So, if φ is an mlf, then knowledge of $\varphi(z)$ essentially determines φ everywhere on A. So $\zeta = \varphi(z) \in \mathbb{C}$ and for every polynomial p, we get $\varphi(p) = p(\zeta)$. We will have $\zeta \in M_A$ if and only if the map $p \mapsto p(\zeta)$ is continuous and indeed, in this case we will have

$$|p(\zeta)| \leq \sup_{z \in K} |p(z)| \quad \text{for all polynomials } p.$$

The ζ that satisfy this inequality form the polynomially convex hull $\hat{K}$ of K. It can be shown that $\mathbb{C} \setminus \hat{K}$ is the unbounded connected component of $\mathbb{C} \setminus K$. Again $M_A = \hat{K}$ with the usual topology. □

EXAMPLE Let $A = \ell^\infty = C(\mathbb{Z})$ the set of bounded two-sided sequences with the uniform norm. Again, A is a commutative Banach algebra with identity. The maximal ideal space is horrendous. We clearly have $\mathbb{Z} \subseteq M_A$. We claim that this inclusion is dense. Suppose not. Then there is an mlf φ which is not in the closure of $\mathbb{Z}$ interpreted as a subset of M_A via the point evaluations. Now the topology of M_A is the topology of convergence on finitely many elements of A, so there exists a neighbourhood of φ defined by finitely many functions which avoids the closure of $\mathbb{Z}$. This means (after adding a suitable constant to each function if necessary) that there exists $N \in \mathbb{N}$ and functions $f_1, f_2, \ldots, f_N \in C(\mathbb{Z})$ such that $\varphi(f_n) = 0$ for $n = 1, 2, \ldots, N$ and the origin is not in the closure of the subset

$$\{(f_1(k), f_2(k), \ldots, f_N(k)); k \in \mathbb{Z}\}$$

in $\mathbb{C}^N$. Then $g = \sum_{n=1}^N |f_n|^2 = \sum_{n=1}^N f_n \overline{f_n}$ has $\varphi(g) = 0$ and yet on $\mathbb{Z}$, g is positive and bounded away from zero. It follows that g is invertible and this is a contradiction.

Therefore M_A is a compactification of $\mathbb{Z}$ and in fact it is called the Stone–Čech compactification. This is the largest possible compactification of $\mathbb{Z}$ and enjoys the following universal property. Let K be a compact topological space into which $\mathbb{Z}$ is mapped injectively and densely (i.e. K is a compactification of $\mathbb{Z}$). Then there is a mapping $\pi : M_A \to K$ which is continuous and onto (but not in general one-to-one) such that the diagram

$$\begin{tikzcd} & M_A \arrow[d, "\pi"] \\ \mathbb{Z} \arrow[ur] \arrow[r] & K \end{tikzcd}$$

commutes. The Stone–Čech compactification is close to being incomprehensible and we should not waste too much time trying to understand it, although of course some mathematicians have spent many years trying to do so. □

2.4 The Spectral Radius Formula

THEOREM 18 *In a commutative Banach algebra we have*

$$\|\hat{x}\|_\infty = \lim_{n\to\infty} \|x^n\|^{\frac{1}{n}}$$

The quantity $\|\hat{x}\|_\infty$ is called the spectral radius of x.

Proof. Without loss of generality, we can assume that the algebra possesses an identity element. Clearly

$$\|\hat{x}\|_\infty \le \|x^n\|^{\frac{1}{n}}$$

for all $n \in \mathbb{N}$ and hence

$$\|\hat{x}\|_\infty \le \liminf_{n\to\infty} \|x^n\|^{\frac{1}{n}}.$$

It remains to show that

$$\limsup_{n\to\infty} \|x^n\|^{\frac{1}{n}} \le \|\hat{x}\|_\infty.$$

If $\zeta \in \mathbb{C}$ and $|\zeta| < \|x\|^{-1}$, then$(\mathbb{1} - \zeta x)^{-1} = \displaystyle\sum_{k=0}^{\infty} \zeta^k x^k$ and

$$x^n = \frac{1}{2\pi i} \oint_{|\zeta|=s} (\mathbb{1} - \zeta x)^{-1} \zeta^{-(n+1)} d\zeta$$

for $s < \|x\|^{-1}$. Now let $t < \|\hat{x}\|_\infty^{-1}$, then since $\zeta \mapsto (\mathbb{1} - \zeta x)^{-1}$ is analytic in $|\zeta| < \|\hat{x}\|_\infty^{-1}$, we also have

$$x^n = \frac{1}{2\pi i} \oint_{|\zeta|=t} (\mathbb{1} - \zeta x)^{-1} \zeta^{-(n+1)} d\zeta.$$

Taking norms in the integral, this yields

$$\|x^n\| \leq t^{-n} \sup_{|\zeta|=t} \|(\mathbb{1} - \zeta x)^{-1}\|$$

and, since the sup is finite,

$$\limsup_{n\to\infty} \|x^n\|^{\frac{1}{n}} \leq t^{-1}.$$

But, now letting t approach its maximum value $\|\hat{x}\|_\infty^{-1}$, we have the desired result.
■

2.5 Haar Measure

In this section we just assume the results that we need. The proofs aren't really very instructive. A locally compact abelian(LCA) group is an abelian group which is also a locally compact topological space. We demand that the multiplication map is continuous as a map from $G \times G$ to G and also that group inversion is continuous as a map from G to G. We will use additive notations. An immediate consequence of the definitions is the following proposition.

PROPOSITION 19 *Given a neighbourhood V of 0 in G, there is a symmetric neighbourhood U of 0 in G such that $U + U \subseteq V$.*

The basic fact that we need is given by the following theorem.

THEOREM 20 *On every LCA group G there is a translation invariant nonnegative regular borel measure η such that $\eta(U) > 0$ for every non-empty open subset U of G and $\eta(K) < \infty$ for every compact subset K of G. Furthermore, the measure η is unique up to a positive multiplicative constant.*

The measure η is called the Haar measure of G. Note that the image of the Haar measure under group negation is also translation invariant and hence a multiple of η. Now test both measures on a symmetric open neighbourhood of 0 to see that the constant must be equal to unity. In other words, any Haar measure is necessarily negation invariant.

2.6 Translation and Convolution

For a function f defined on G and $x \in G$ we define $f_x(y) = f(y - x)$ for all $y \in G$. We call f_x the translate of f by x.

LEMMA 21 *Let $1 \leq p < \infty$. For $f \in L^p(G)$ we have that $x \mapsto f_x$ is a uniformly continuous map from G to $L^p(G)$.*

Proof. Suppose first that $f \in C_c(G)$, the space of continuous functions of compact support on G. Then f is uniformly continuous. (Note that G has a natural uniform structure coming from group subtraction). Let K be the support of f. Let U be a compact neighbourhood of 0. Then $K + U$ is also compact. Let it have measure t. Let $\epsilon > 0$, then, since f is uniformly continuous, there exists a compact neighbourhood V of 0 such that

$$x \in V \Longrightarrow \|f - f_x\|_\infty < t^{-\frac{1}{p}}\epsilon.$$

If $x \in U \cap V$, then the support of $f - f_x$ is contained in $K + U$ and it follows that $\|f - f_x\|_p < \epsilon$.

In the general case, Let $h \in L^p$ and let $\epsilon > 0$. We first approximate h by a function $f \in C_c(G)$ so that $\|f - h\|_p < \epsilon$. It is in this last step that the fact $p < \infty$ is used. Then, since the underlying measure is translation invariant, $\|f_x - h_x\|_p = \|f - h\|_p < \epsilon$ and we have our result. if $x \in U \cap V$, then

$$\|h - h_x\|_p \leq \|f - h\|_p + \|f - f_x\|_p + \|f_x - h_x\|_p < 3\epsilon.$$

■

We now define convolution. If f and g are suitable functions, we set

$$f \star g(x) = \int f(x - y)g(y)d\eta(y).$$

If we make the substitution $y = x - z$ in this integral we get

$$f \star g(x) = \int f(z)g(x - z)d\eta(z) = g \star f(x),$$

using that η is both translation and reflection invariant.

LEMMA 22

1. *If $f \in L^1$ and $g \in L^\infty$, then $f \star g$ is bounded and uniformly continuous.*

2. *If $f, g \in C_c(G)$ then $f \star g \in C_c(G)$.*

3. *If $1, p < \infty$, $f \in L^p$, $g \in L^{p'}$, then $f \star g \in C_0$.*

4. *If $f, g \in L^1$, then $f \star g \in L^1$.*

Proof. In 1), Clearly $f \star g$ is bounded by $\|f\|_1 \|g\|_\infty$. We rewrite

$$f \star g(x) = \int f(x-y)g(y)d\eta(y) = \int h(y-x)g(y)d\eta(y) = \int h_x(y)g(y)d\eta(y)$$

where $h(x) = f(-x)$ and the uniform continuity is clear since $x \mapsto h_x$ is uniformly continuous for the L^1 norm.

For 2), clearly continuous by 1). Also $\mathrm{supp}(f \star g) \subseteq \mathrm{supp}(f) + \mathrm{supp}(g)$. Note that $\mathrm{supp}(f) + \mathrm{supp}(g)$ is the continuous image of $\mathrm{supp}(f) \times \mathrm{supp}(g)$ under the addition map $G \times G \to G$ and hence is compact.

For 3), proceed as in 1). We see that $f \star g$ is bounded by $\|f\|_p \|g\|_{p'}$ by using Hölder's inequality. By 2) $f \star g$ is a uniform limit of continuous functions of compact support. Hence $f \star g \in C_0$.

For 4), we start by oberving that if U is open in G, then $(x, y); x \in G, y \in G, x - y \in U\}$ is an open subset of $G \times G$. It follows that if B is a Borel set in G, then $(x, y); x \in G, y \in G, x - y \in B\}$ is Borel in $G \times G$. So, replacing both f and g with Borel versions, we see that $(x, y) \mapsto f(x-y)$ and $(x, y) \mapsto f(x-y)g(y)$ are Borel functions on $G \times G$. By Fubini's Theorem, this last function is absolutely integrable on the product space because

$$\iint |g(y)f(x-y)|d\eta(x)d\eta(y) = \int \|f\|_1 |g(y)| d\eta(y) = \|f\|_1 \|g\|_1 < \infty$$

It now follows that $f \star g(x) = \int f(x-y)g(y)d\eta(y)$ is a measurable function (finite almost everywhere) for the completion of the Borel σ-field with respect to η. It also follows from Fubini's Theorem that $f \star g \in L^1$ and that $\|f \star g\|_1 \leq \|f\|_1 \|g\|_1$. It is an exercise to check that different versions of f and g yield the same element of $f \star g$ viewed as an element of L^1. ∎

Note: OK, so I lied. The problem with the proof of 4) above is that one of the hypotheses of Fubini's Theorem is that the underlying measure space be σ-finite.

Unfortunately not all LCA groups are σ-finite, for example any discrete uncountable abelian group will fail to be σ-finite. We would have no difficulty handing the case of discrete groups, because L^1 functions on such groups would have to be carried by countable sets and actually by countable subgroups.

In a general LCA group G, we take an open relatively compact neighbourhood U of 0 and consider $U_n = U + U + \cdots + U$ with n summands. Note that U_n is open and relatively compact. Now consider

$$G_0 = \bigcup_{n=0}^{\infty} U_n$$

an open subgroup of G which is σ-finite. But an open subgroup of G is also closed (because it is the complement of the union of all the cosets not equal to the subgroup itself) and it follows that the quotient G/G_0 is discrete. You can now show that given $f, g \in L^1(G)$, there is a σ-finite open and closed subgroup H of G such that in fact, f and g are carried on H. Now you apply the argument in 4) above to H. We reserve the right in these notes to tell this same lie again without comment.

We now have the following theorem which is easy to check.

THEOREM 23 *For g an LCA group, $L^1(G)$ is a commutative Banach algebra with convolution multiplication.*

If G is discrete, then δ_0 is an identity element. It turns out that if $L^1(G)$ has an identity element, then G is discrete, but this is not too obvious.

A character χ on G is a continuous group homomorphism into the multiplicative group of unimodular convex numbers. We will denote the set of all characters on G by Γ. We can give Γ the structure of a group in the obvious way. We will use additive notations for consistency even though they look a trifle strange.

$$(-\chi)(x) = \overline{\chi(x)}, \quad (\chi_1 + \chi_2)(x) = \chi_1(x)\chi_2(x)$$

The Fourier transform $\hat{f}$ of $f \in L^1(G)$ is now given by

$$\hat{f}(\chi) = \int f(x)\overline{\chi(x)}d\eta(x). \tag{2.1}$$

This is a linear functional on $L^1(G)$ and furthermore multiplicative

$$\int f \star g(x)\overline{\chi(x)}d\eta(x) = \int\int f(x-y)g(y)d\eta(y)\overline{\chi(x)}d\eta(x)$$

$$= \iint f(x-y)\overline{\chi(x-y)\chi(y)}g(y)d\eta(y)d\eta(x)$$
$$= \iint f(z)\overline{\chi(z)}d\eta(z)\overline{\chi(y)}g(y)d\eta(y)$$
$$= \hat{f}(\chi)\hat{g}(\chi)$$

Note also that

$$f \star \chi = \hat{f}(\chi)\chi$$

THEOREM 24 *Every mlf on $L^1(G)$ is given by a character as in (2.1).*

Proof. Every bounded linear functional on L^1 is given by an L^∞ function. So, every non-zero mlf φ, which necessarily has norm 1 would have to be given by a function $h \in L^\infty$ with $\|h\|_\infty = 1$ by

$$\varphi(f) = \int f(x)\overline{h(x)}d\eta(x).$$

Now

$$\varphi(f)\int g(y)\overline{h(y)}d\eta(y) = \varphi(f)\varphi(g) = \varphi(f \star g) = \varphi\left(\int f_y g(y)d\eta(y)\right)$$
$$= \int (\varphi(f_y)g(y)d\eta(y)$$

and this holds for all $g \in L^1$. Therefore

$$\varphi(f)\overline{h(y)} = \varphi(f_y) \tag{2.2}$$

for almost all y. Choosing f so that $\varphi(f) \neq 0$ and since $y \mapsto f_y$ is continuous, we see that h has a continuous version. Replacing h with its continuous version, it is now clear that (2.2) holds for all $y \in G$ (a conull set must be dense). Now we have

$$\varphi(f)\overline{h(x+y)} = \varphi(f_{x+y}) = \varphi(f_x)\overline{h(y)} = \varphi(f)\overline{h(x)h(y)}$$

giving $h(x+y) = h(x)h(y)$. Put now $x = y = 0$ to get $h(0) = 0$ or 1. But $h(0) = 0$ implies that h and hence φ vanishes identically and hence we must have $h(0) = 1$. But now $h(x)h(-x) = 1$ and it comes that $|h(x)| = 1$ for all $x \in G$. ∎

So, on $L^1(G)$, the Gelfand transform and the Fourier transform are the same. We note that if G is discrete, then $L^1(G)$ has an identity and Γ is compact. If G is compact, then η has finite measure. Normally it is normalized to have total mass 1 in this case. We have

$$\int \chi(x) d\eta(x) = \begin{cases} 1 & \text{if } \chi \text{ is the zero element of } \Gamma, \\ 0 & \text{otherwise.} \end{cases}$$

since if $\chi = \mathbb{1}$, the first assertion is obvious. Otherwise there is an element $y \in G$ such that $\chi(y) \neq 1$. Then

$$\int \chi(x) d\eta(x) = \int \chi(x+y) d\eta(x) = \chi(y) \int \chi(x) d\eta(x)$$

so that

$$(1 - \chi(y)) \int \chi(x) d\eta(x) = 0.$$

Note that in this case, the characters are themselves elements of $L^1(G)$. Thus $\hat{\chi}(\psi) = 1$ if $\chi = \psi$ and $= 0$ otherwise. Since $\hat{\chi}$ is a continuous function on Γ, it follows that Γ is discrete.

THEOREM 25

1. *$(x, \chi) \mapsto \chi(x)$ is jointly continuous $G \times \Gamma \to \mathbb{T}$.*
2. *Let K and C be compact in G and Γ respectively, then for $t > 0$*

$$N(K, t) = \{\chi; |\chi(x) - 1| < t \text{ for all } x \in K\}$$
$$N(C, t) = \{x; |\chi(x) - 1| < t \text{ for all } \chi \in C\}$$

   are open in Γ and G respectively.

3. *The sets $N(K, t)$ and their translates form a base for the topology of Γ.*
4. *Γ is an LCA group.*

Proof. For 1) let $f \in L^1(G)$. We know that $x \mapsto f_x$ is continuous from G to $L^1(G)$. So, since the Gelfand transform is continuous, $x \mapsto \hat{f}_x$ is continuous from G to $C_0(\Gamma)$. But

$$\hat{f}_x(\chi) = \overline{\chi(x)} \hat{f}(\chi)$$

and it follows that $(x, \chi) \mapsto \chi(x)$ is jointly continuous on the set $\{(x, \chi); x \in G, \hat{f}(\chi) \neq 0\}$. But, for each $\chi \in \Gamma$ it is easy to construct $f \in L^1(G)$ such that $\hat{f}(\chi) \neq 0$ and we see that $(x, \chi) \mapsto \chi(x)$ is jointly continuous on $G \times \Gamma$.

Next we prove 2). Let C be compact in Γ and let $x_0 \in N(C, t)$. Then $|\chi(x_0) - 1| < t$ for all $\chi \in C$. So, for each χ there an open neighbourhood V_χ of χ in Γ and an open neighbourhood U_χ of x_0 in G such that $|\psi(x) - 1| < t$ for all $\psi \in V_\chi$ and $x \in U_\chi$. Finitely many such neighbourhoods $V\chi$ cover C. Let U be the open intersection of the corresponding U_χ. Then it is clear that $x_0 \in U \subseteq N(C, t)$. The other assertion is proved similarly.

Note that 2) states that the Gelfand topology in Γ is finer than the compact open mapping topology (i.e. the topology of uniform convergence on the compact sets). For 3), we have to show the converse and for this it is enough to show that each Gelfand transform $\hat{f}$ for $f \in L^1(G)$ is continuous for the compact open topology. If the function f has compact support, this is obvious since

$$\begin{aligned} |\hat{f}(\chi_1) - \hat{f}(\chi_2)| &\leq \int_{\mathrm{supp}(f)} |\chi_1(x) - \chi_2(x)||f(x)|d\eta(x) \\ &\leq \|f\|_1 \sup_{x \in \mathrm{supp}(f)} |\chi_1(x) - \chi_2(x)|. \end{aligned}$$

But any L^1 function can be approximated in L^1 norm by L^1 functions of compact support and the corresponding transforms converge uniformly. This completes the proof of 3).

To prove 4), we simply observe that compact open topology on Γ is clearly a group topology. This really amounts to observing that for every compact subset of G and every $t > 0$ we have

$$N(K, t/2) - N(K, t/2) \subseteq N(K, t),$$

or equivalently that the standard topology on $\mathbb{T}$ is a group topology. ■

The group Γ is called the dual group of G.

2.8 Summability Kernels

Here we give the theory of summability kernels as it applies to LCA groups. The Bernstein approximation theorem (the proof using the Bernstein polynomials gives an example of the idea in more general situations).

Let $k_n \in L^1(G)$ be indexed over $n \in \mathbb{N}$. (In general other indexing sets are used). We suppose

- $k_n \geq 0$.
- $\int_G k_n(x) d\eta(x) = 1$, for all $n \in \mathbb{N}$.
- For every measurable neighbourhood V of 0 we have

$$\lim_{n\to\infty} \int_{G\setminus V} k_n(x) d\eta(x) = 0.$$

We have the following general theorem.

THEOREM 26 *Let B be a banach space of objects on which G acts isometrically and continuously. We will denote b_x for the result of applying of the group element x to $b \in B$. Then*

$$\int b_x k_n(x) d\eta(x) \underset{n\to\infty}{\longrightarrow} b.$$

Proof. We have

$$b - \int b_x k_n(x) d\eta(x) = \int (b - b_x) k_n(x) d\eta(x)$$

and so

$$\left\| b - \int b_x k_n(x) d\eta(x) \right\| \leq \int \|b - b_x\| k_n(x) d\eta(x).$$

Now, let $\epsilon > 0$. There exists V a measurable neighbourhood of 0 such that

$$x \in V \Leftrightarrow \|b - b_x\| < \epsilon$$

and then there exists $N \in \mathbb{N}$ such that

$$n \geq N \Leftrightarrow \int_{G\setminus V} k_n(x) d\eta(x) < \epsilon.$$

We have

$$\begin{aligned}\int_G \|b - b_x\| k_n(x) d\eta(x) &\leq \int_V \|b - b_x\| k_n(x) d\eta(x) + \int_{G\setminus V} \|b - b_x\| k_n(x) d\eta(x) \\ &\leq \epsilon \int_V k_n(x) d\eta(x) + \int_{G\setminus V} (\|b\| + \|b_x\|) k_n(x) d\eta(x)\end{aligned}$$

$$\leq \epsilon \int_G k_n(x)d\eta(x) + \int_{G\setminus V} (\|b\| + \|b\|)k_n(x)d\eta(x)$$
$$\leq \epsilon + 2\|b\| \int_{G\setminus V} k_n(x)d\eta(x)$$
$$\leq \epsilon + 2\|b\|\epsilon$$

for $n \geq N$. ■

COROLLARY 27 *Let $1 \leq p < \infty$. Let $f \in L^p(G)$ and (k_n) be a summability kernel. Then $k_n \star f \to f$ in L^p norm.*

EXAMPLE Let φ be a bounded continuous function on G. Show that

$$\int \varphi(x)k_n(x)d\eta(x) \underset{n\to\infty}{\longrightarrow} \varphi(0).$$

□

2.9 Convolution of Measures

Let λ and μ be complex borel measures on G. Then we define thir convolution product $\lambda * \mu$ by

$$\lambda * \mu(B) = \lambda \otimes \mu(\alpha^{-1}(B)) \tag{2.3}$$

where α is the addition map $\alpha : G \times G \to G$ given by $\alpha(x, y) = x + y$. This extends to suitable measurable functions via

$$\int_G f d\lambda * \mu = \int_G \int_G f(x + y)d\lambda(x)d\mu(y). \tag{2.4}$$

In fact, (2.3) is just the special case $f = \mathbb{1}_B$. It's easy to check that the convolution multiplication is associative and (on an abelian group) commutative. The totality of all complex borel measures on G is denoted $M(G)$. Since all complex borel measures are necessarily bounded, we can put the total mass norm $\|\ \|_M$ on $M(G)$ and it can then be realised as the dual space of $C_0(G)$. Taking the supremum over all $f \in C_0(G)$ with norm bounded by one in (2.4), we see that $\|\lambda * \mu\|_M \leq \|\lambda\|_M\|\mu\|_M$. It follows that $M(G)$ is a commutative Banach algebra with identity δ_0. The maximal ideal space of $M(G)$ is pathological. It is true that the mappings

$$\mu \mapsto \hat{\mu}(\chi) = \int_G \overline{\chi(x)} d\mu(x)$$

for $\chi \in \Gamma$ which define the so-called Fourier-Stieltjes transform of μ are multiplicative linear functionals on $M(G)$, but there are other less obvious mlfs as well (at least when G is non-discrete).

We now have the following uniqueness theorem which is the wrong way around.

THEOREM 28 *Let $\mu \in M(\Gamma)$ be such that $\int_\Gamma \overline{\chi(x)} d\mu(\chi) = 0$ for all $x \in G$. Then $\mu = 0$ identically.*

Proof. Let f be in $L^1(G)$, then $\int_G f(x) \int_\Gamma \overline{\chi(x)} d\mu(\chi) d\eta(x) = 0$. Then we have $\int_G \int_\Gamma |f(x)| d|\mu|(\chi) d\eta(x) < \infty$ and hence by Fubini's Theorem, we have

$$\int_\Gamma \hat{f}(\chi) d\mu(\chi) = \int_\Gamma \int_G f(x) \overline{\chi(x)} d\eta(x) d\mu(\chi) = 0 \tag{2.5}$$

But the set of Fourier transforms $A(\Gamma)$ of L^1 functions on G is a self-adjoint subalgebra of $C_0(\Gamma)$ under pointwise mutiplication which separates the points of Γ (compact case) and the points of the one-point compactification of Γ in the non-compact case. To verify the self-adjointness, we check

$$\int \overline{f(-x)\chi(x)} d\eta(x) = \overline{\int f(-x)\chi(x) d\eta(x)} = \overline{\int f(x)\chi(-x) d\eta(x)}$$

so that $\hat{g}(\chi) = \overline{\hat{f}(\chi)}$, where $g(x) = \overline{f(-x)}$. Therefore, by the Stone–Weierstrass Theorem, $A(\Gamma)$ is dense in $C_0(\Gamma)$. It follows from (2.5) that $\mu = 0$. ■

2.10 Positive Definite Functions

Let φ be a complex-valued function on G, then we say that φ is positive semidefinite if and only the matrix M given by

$$m_{j,k} = \varphi(x_j - x_k)$$

is positive semidefinite for all choices of finitely many points $(x_j)_{j=1}^n$ from G. Explicitly, this means that

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\overline{c_j}c_k\varphi(x_j-x_k)\geq 0$$

for all $n\in\mathbb{N}$, $c_j\in\mathbb{C}$ and $x_j\in G$. Let φ be a positive semidefinite function. Then clearly $\varphi(0)\geq 0$ (take $n=1$ and $c_1\neq 0$). Also, a positive semidefinite matrix has to be hermitian, so $\varphi(-x)=\overline{\varphi(x)}$. Now the matrix

$$\begin{pmatrix}\varphi(0) & \varphi(x)\\ \varphi(-x) & \varphi(0)\end{pmatrix}=\begin{pmatrix}\varphi(0) & \varphi(x)\\ \overline{\varphi(x)} & \varphi(0)\end{pmatrix}$$

is positive semidefinite and has a nonnegative determinant, so $|\varphi(x)|\leq\varphi(0)$ for all $x\in G$. Similarly, the matrix

$$\begin{pmatrix}\varphi(0) & \varphi(x) & \varphi(y)\\ \varphi(-x) & \varphi(0) & \varphi(y-x)\\ \varphi(-y) & \varphi(x-y) & \varphi(0)\end{pmatrix}=\begin{pmatrix}\varphi(0) & \varphi(x) & \varphi(y)\\ \overline{\varphi(x)} & \varphi(0) & \overline{\varphi(x-y)}\\ \overline{\varphi(y)} & \varphi(x-y) & \varphi(0)\end{pmatrix}$$

is positive semidefinite and hence, using simulaneous row and column reduction, so is

$$\begin{pmatrix}\varphi(0) & \varphi(x)-\varphi(y)\\ \overline{\varphi(x)-\varphi(y)} & 2\Re(\varphi(0)-\varphi(x-y))\end{pmatrix}.$$

It follows that

$$|\varphi(x)-\varphi(y)|^2\leq 2\varphi(0)\Big(\varphi(0)-\Re\varphi(x-y)\Big).$$

It is easy to check that if $f\in L^2(G)$, then $\tilde{f}\star f$ is a continuous positive semidefinite function on G tending to zero at infinity. However, there is a complete characterization of the continuous positive semidefinite functions on G.

THEOREM 29 (BOCHNER'S THEOREM) *Every continuous positive semidefinite function φ on G has the form*

$$\varphi(x)=\int_{\Gamma}\overline{\chi(x)}d\mu(\chi) \tag{2.6}$$

where μ is a nonnegative Borel measure (of finite total mass) on Γ and conversely.

Proof. It is routine to check that if φ is defined by (2.6) then φ is continuous and positive semidefinite on G. For the converse, it is an exercise to check that (since φ is bounded and continuous), we have

$$\iint_{G\times G}\varphi(x-y)f(x)\overline{f(y)}d\eta(x)d\eta(y)\geq 0,$$

for $f\in L^1(G)$. We use the formula to define a quasi inner product on $L^1(G)$ by

$$<f,g>=\iint_{G\times G}\varphi(x-y)g(x)\overline{f(y)}d\eta(x)d\eta(y)=\int_G(\tilde{f}*g)\varphi d\eta.$$

This is in all respects like an inner product, except that the implication $<f,f>=0$ does not necessarily imply that f is the zero element of $L^1(G)$. Nevertheless, the proof of the corresponding Cauchy–Schwarz–Bunyakowski inequality goes thru, giving

$$|<f,g>|^2\leq <f,f><g,g>$$

Now pass to the limit as f runs over a summability kernel on G. We get

$$\left|\int_G\varphi(x)g(x)d\eta(x)\right|^2\leq\varphi(0)<g,g>$$

for all $g\in L^1(G)$. Let $g_1=\tilde{g}\star g$ and $g_{n+1}=\tilde{g_n}\star g_n$ for $n=1,2,\ldots$ Actually, $\tilde{g_n}=g_n$ and it follows that $g_{n+1}=\star^{2^n}g_1$, the 2^n-fold convolution product of g_1 with itself. The point is that

$$\left|\int_G\varphi(x)g_n(x)d\eta(x)\right|^2\leq\varphi(0)<g_n,g_n>=\varphi(0)\int_G\varphi(x)g_{n+1}(x)d\eta(x)$$

It follows from this and a simple induction that

$$\left|\int_G\varphi(x)g(x)d\eta(x)\right|^{2^n}\leq\varphi(0)^{2^n}\|\star^{2^{n-1}}g_1\|_1$$

and, after taking the root of order 2^{n-1} and passing to the limit with the spectral radius formula, we get

$$\left|\int_G\varphi(x)g(x)d\eta(x)\right|^2\leq\varphi(0)^2\|\hat{g}_1\|_\infty\leq\varphi(0)^2\|\hat{g}\|_\infty^2$$

or

$$\left|\int_G \varphi(x)g(x)d\eta(x)\right| \le \varphi(0)\|\hat{g}\|_\infty$$

This tell us that $\displaystyle\int_G \varphi(x)g(x)d\eta(x)$ depends only on the value of $\hat{g}$ and (since $A(\Gamma)$ is dense in $C_0(\Gamma)$) that there is a measure μ on Γ of total mass at most $\varphi(0)$ such that

$$\int_G \varphi(x)g(x)d\eta(x) = \int_\Gamma \hat{g}(\chi)d\mu(\chi)$$

But now,

$$\begin{aligned}\int_G \varphi(x)g(x)d\eta(x) &= \int_\Gamma\int_G g(x)\overline{\chi(x)}d\eta(x)d\mu(\chi)\\ &= \int_G g(x)\int_\Gamma \overline{\chi(x)}d\mu(\chi)d\eta(x) \qquad (2.7)\end{aligned}$$

The functions φ and $x \mapsto \int_\Gamma \overline{\chi(x)}d\mu(\chi)$ are both continuous on G and since 2.7 holds for all $g \in L^1(G)$, we have (2.6) holding for all $x \in G$ as required. ■

Something special happened in the proof above. Before this theorem, we didn't know that the points of G could be separated by its characters, but now we do. Given $x \neq 0$ in G, find a symmetric neighbourhood V of 0 such that $x \notin V + V$. Then apply Bochner's Theorem to $\mathbb{1}_V \star \mathbb{1}_V$.

We are now ready to prove a preliminary form of the inversion theorem.

THEOREM 30 *Let $f \in L^1(G)$ be also given by $f(x) = \int_\Gamma \chi(x)d\mu_f(\chi)$ where μ_f is a complex measure. (Note that complex measures have finite total mass). Then $\mu_f = \hat{f}\nu$ where ν is a suitably normalized Haar measure on Γ.*

Proof. Let f and g be two such functions with associated measures μ_f and μ_g. Let $h \in L^1(G)$. Then

$$\begin{aligned}\iint h(-x-y)f(x)g(y)d\eta(x)d\eta(y) =&\iint h(-x-y)\chi(x)g(y)d\mu_f(\chi)d\eta(x)d\eta(y)\\ =&\iint h(-x-y)\chi(x)g(y)d\eta(x)d\eta(y)d\mu_f(\chi)\\ =&\iint h(-x-y)\chi(x)g(y)d\eta(x)d\eta(y)d\mu_f(\chi)\end{aligned}$$

$$
\begin{aligned}
&= \iint h(x-y)\overline{\chi(x)}g(y)d\eta(x)d\eta(y)d\mu_f(\chi)\\
&= \iint h(x-y)\overline{\chi(x)}g(y)d\eta(y)d\eta(x)d\mu_f(\chi)\\
&= \iint g\star h(x)\overline{\chi(x)}d\eta(x)d\mu_f(\chi)\\
&= \iint \widehat{g\star h}(\chi)d\mu_f(\chi)\\
&= \iint \hat{h}(\chi)\hat{g}(\chi)d\mu_f(\chi)
\end{aligned}
$$

and also by the symmetry of the initial expression in f and g

$$
= \iint \hat{h}(\chi)\hat{f}(\chi)d\mu_g(\chi)
$$

Again, since $A(\Gamma)$ is dense in $C_0(\Gamma)$, we find $\hat{g}d\mu_f = \hat{f}d\mu_g$.

Now we can construct functions like f and g easily. Let V be a measurable neighbourhood of 0 and let $h = \mathbb{1}_V$, then by Bochners Theorem, there is a measure $\mu_{h\star\tilde{h}}$ such that

$$
\begin{aligned}
h\star\tilde{h}(x) &= \int_\Gamma \chi(x)d\mu_{h\star\tilde{h}}(\chi)\\
\widehat{h\star\tilde{h}}(\chi) &= |\hat{h}(\chi)|^2
\end{aligned}
$$

Furthermore, if $\psi\in\Gamma$

$$
\begin{aligned}
(\psi h\star\widetilde{\psi h})(x) = \psi(x)\, h\star\tilde{h}(x) &= \int_\Gamma \chi(x)d\mu_{h\star\tilde{h}}(\chi-\psi)\\
&= \int_\Gamma \chi(x)d\mu_{(\psi h\star\widetilde{\psi h})}(\chi)\\
\widehat{\psi h\star\widetilde{\psi h}}(\chi) &= |\hat{h}(\chi-\psi)|^2.
\end{aligned}
$$

Note also that $\hat{h}$ is continuous and $\hat{h}(0)$ is nonzero. This leads to

$$
|\hat{h}(\chi-\psi)|^2 d\mu_f(\chi) = \hat{f}(\chi)d\mu_{(\psi h\star\widetilde{\psi h})}(\chi)
$$

showing that μ_f is uniquely determined near ψ and hence everywhere on Γ.

Also we may infer the existence (the details are an exercise) of a positive measure ν such that

$$\mu_f = \hat{f}\nu.$$

A straightforward compactness argument shows that ν is finite on the compact sets and charges every nonempty open set.

Now let us abbreviate $h \star \tilde{h}$ to g. Then

$$\widehat{\psi g}(\chi) d\mu_g(\chi) = \hat{g}(\chi) d\mu_{\psi g}(\chi) = \hat{g}(\chi) d\mu_g(\chi - \psi)$$

leading to

$$\hat{g}(\chi - \psi)\hat{g}(\chi) d\nu(\chi) = \hat{g}(\chi)\hat{g}(\chi - \psi) d\nu(\chi - \psi).$$

Now suppose that ψ is given, then, choosing ϵ suitably small, choose V such that $V \subseteq \{x; |\psi(x) - 1| < \epsilon\}$. Then for χ in a neighbourhood of 0_Γ, $\hat{g}(\chi - \psi)\hat{g}(\chi) \neq 0$, showing that $d\nu(\chi) = d\nu(\chi - \psi)$ at least for values of χ in a neighbourhood of 0_Γ. It follows that ν is translation invariant and hence a multiple of Haar measure on Γ. Again, the details are an exercise. ■

Now let V be a symmetric neighbourhood of 0 in G. Let $g = (\eta(V))^{-1} \mathbb{1}_V \star \mathbb{1}_V$. Then $g(0) = 1$ and g is positive definite. It follows that $g(x) = \int_\Gamma \hat{g}(\chi)\chi(x) d\nu(\chi)$. Now $\hat{g}$ is in $L^1(\Gamma)$ with norm 1 and there is a compact subset C of Γ such that

$$\int_{\Gamma \setminus C} g(\chi) d\nu(\chi) < \frac{1}{5}$$

Suppose that $x \in N(C, \frac{1}{5})$. Then

$$\begin{aligned}
\left| 1 - \int_\Gamma \hat{g}(\chi)\chi(x) d\nu(\chi) \right| &\leq \int_\Gamma \hat{g}(\chi) |1 - \chi(x)| d\nu(\chi) \\
&\leq \int_{\Gamma \setminus C} \hat{g}(\chi) |1 - \chi(x)| d\nu(\chi) + \int_C \hat{g}(\chi) |1 - \chi(x)| d\nu(\chi) \\
&\leq \frac{2}{5} + \frac{1}{5} = \frac{3}{5}
\end{aligned}$$

So, $g(x) \geq \frac{2}{5}$ and $x \in V + V$. It follows from this that the compact open topology defined on G by means of the duality with Γ is finer than and therefore equivalent to the original topology on G.

2.11 The Plancherel Theorem

This is an immediate consequence of the inversion theorm.

THEOREM 31 (PLANCHEREL THEOREM) *Let $f \in L^1(G) \cap L^2(G)$. Then*

$$\int_G |f(y)|^2 d\eta(y) = \int_\Gamma |\hat{f}(\chi)|^2 d\nu(\chi) \tag{2.8}$$

so that

$$\begin{aligned} f \quad &\mapsto \hat{f} \\ L^1(G) \cap L^2(G) &\longrightarrow L^2(\Gamma) \end{aligned}$$

extends by continuity to a surjective isometry

$$L^2(G) \longrightarrow L^2(\Gamma)$$

Proof. Let $h = f \star \tilde{f}$, then $h(x) = \int_G f(x-y)\overline{f(-y)} d\eta(y)$ and $h(0) = \|f\|_2^2$. Since h is both in L^1 and is positive definite, it can be represented by a measure μ_h of total mass $h(0)$. Also, $\mu_h = \hat{h}\nu$. Since $\hat{h}(\chi) = \widehat{f \star \tilde{f}}(\chi) = |\hat{f}(\chi)|^2$, we have (2.8). The remainder of the result is obvious, except for the fact that the isometry is surjective. To see this, suppose not. Then there is a nonzero function $\phi \in L^2(\Gamma)$ such that

$$\int_\Gamma \hat{f}(\chi)\phi(\chi) d\nu(\chi) = 0$$

for all $f \in L^1(G) \cap L^2(G)$. Fix such an f and consider its translation f_x. We get

$$\int_\Gamma \overline{\chi(x)}\hat{f}(\chi)\phi(\chi) d\nu(\chi) = \int_\Gamma \hat{f}_x(\chi)\phi(\chi) d\nu(\chi) = 0$$

for all $x \in G$. But now by Theorem 28 and since $\hat{f}\phi\nu$ is a measure ($\hat{f}\phi \in L^1$), we have that $\hat{f}\phi$ vanishes ν almost everywhere. But we know how to choose f such that $\hat{f}$ is non-vanishing in a neighbourhood of any given point of Γ. Hence $\phi = 0$ almost everywhere (and as an element of L^2). ■

COROLLARY 32 *Let $f, g \in L^2(G)$, then $\widehat{fg} = \hat{f} \star \hat{g}$.*

Proof. Polarizing the Plancherel identity leads to

$$\int_G f(y)\overline{g(y)}d\eta(y) = \int_\Gamma \hat{f}(\chi)\overline{\hat{g}(\chi)}d\nu(\chi)$$

for $f, g \in L^2(G)$. The notations $\hat{f}, \hat{g}$ now stand for the abstract nonsense fourier transforms of f and g respectively. Replace g by $\overline{g}$ and f by $\overline{\psi}f$ where $\psi \in \Gamma$. We get

$$\int_G \overline{\psi}f(y)g(y)d\eta(y) = \int_\Gamma \hat{f}(\chi+\psi)\hat{g}(-\chi)d\nu(\chi)$$

which after a change of variables gives exactly $\widehat{fg} = \hat{f} \star \hat{g}$. ■

COROLLARY 33 *Let Ω be a nonempty open subset of Γ. Then there is a function $f \in L^1(G)$ such that $\hat{f}$ is not identically zero and $\hat{f}(\chi) = 0$ for all $\chi \in \Gamma \setminus \Omega$.*

Proof. First, find V_1 and V_2 open nonempty and relatively compact with $V_1 + V_2 \subset \Omega$. Then let $f_j \in L^2(G)$ be the elements such that $\hat{f}_j = \mathbb{1}_{V_j}$ for $j = 1, 2$. Then $f = f_1 f_2$ does the trick, since $\hat{f} = \mathbb{1}_{V_1} \star \mathbb{1}_{V_2}$. ■

2.12 The Pontryagin Duality Theorem

Let H be the dual group of Γ. Every element of G defines a continuous character on Γ, so there is a map $\alpha : G \to H$ which is clearly one-to-one (different elements of G define different characters since we know that the characters of G separate the points of G).

THEOREM 34 (PONTRYAGIN DUALITY THEOREM) *The mapping α is an isomorphism of topological groups.*

Proof. It is clear that α is an injective group homomorphism. We also know that the topologies of G and H can be identified to the compact open topology when these spaces are viewed as function spaces on Γ. Therefore the topology of G is the subspace topology coming from H. Now the uniform structure of an abelian topological group is given from the topology by means of translation. Therefore, the uniform structure on G is just the restriction of the uniform structure on

H. But G is locally compact and hence as a uniform space, it is complete. But when complete spaces occur as subsets of other spaces, they are necessarily closed. Hence $\alpha(G)$ is a closed subset of H[1].

It remains only to show that $\alpha(G)$ is dense in H. But, if not, then by one of the corollaries of the Plancherel Theorem, we can find $f \in L^1(\Gamma)$ nonzero, with $\hat{f}(x) = 0$ for all $x \in \alpha(G)$. But then Theorem 28 implies that f is almost everywhere zero on Γ a contradiction. ■

Some of the consequences of the Pontryagin Duality Theorem are as follows:

- Every compact abelian group is the dual of a discrete abelian group.
- Every discrete abelian group is the dual of a compact abelian group.
- If $\mu \in M(G)$ and $\hat{\mu}(\chi) = 0$ for all $\chi \in \hat{G}$, then $\mu = 0$. In particular, both $L^1(G)$ and $M(G)$ are semisimple Banach algebras.
- If G is not discrete, then $\hat{G}$ is not compact and hence $L^1(G)$ does not have an identity element.
- We can restate the inversion theorem the correct way around. If $\mu \in M(G)$ and $\hat{\mu} \in L^1(\hat{G})$, then there exists $f \in L^1(G)$ such that $\mu = f\eta$ and the inversion formula
$$f(x) = \int_{\hat{G}} \hat{\mu}(\chi)\chi(x)d\eta(\chi)$$
holds.

[1] If you are reading along in Rudin's book, please note that it is in general false that a locally compact subspace of a locally compact topological space is necessarily closed. Whatever Rudin intended in §1.7 is by no means clear.