A note on Fourier Fourier transform of measures and almost periodic functions

Son, Nguyen Thai Tu

March 26, 2018

Abstract

This is an expository note based on the materials from the book "Introduction to Harmonic Analysis" by Katznelson.

Contents

1 Fourier series

- We denote by $\mathbb{R}$ the additive group of real numbers and by $\mathbb{Z}$ the subgroup consisting of the integers. The group $\mathbb{T}$ is quotient group $\mathbb{R}/2\pi\mathbb{Z}$.
- There is an obvious identification between functions on $\mathbb{T}$ and 2π-periodic functions on $\mathbb{R}$, which allows the notions of continuity, differentiability, etc. for functions on $\mathbb{T}$.
- The Lebesgue measure on $\mathbb{T}$ is the defined in the same manner, which is roughly understood as the restriction of the Lebesgue measure to $[0,2\pi)$, and a function f is integrable on $\mathbb{T}$ if the corresponding 2π-periodic function, which we denote again by f, is integrable on $[0,2\pi)$ and we set
$$\int_{\mathbb{T}} f(t)\,dt = \int_0^{2\pi} f(x)\,dx.$$

- An important property of dt on $\mathbb{T}$ is its translation invariance, that is, for all $t_0 \in \mathbb{T}$ and f defined on $\mathbb{T}$ we have
$$\int_{\mathbb{T}} f(t-t_0)\,dt = \int_{\mathbb{T}} f(t)\,dt.$$

1.1 Fourier coefficients

1. We denote bt $L^1(\mathbb{T})$ the space of all (equivalent classes of) complex-valued, Lebesgue integrable functions on $\mathbb{T}$. For $f \in L^1(\mathbb{T})$, we put
$$\|f\|_{L^1} = \frac{1}{2\pi}\int_{\mathbb{T}} f(t)\,dt.$$
The total mass of dt on $\mathbb{T}$ equal to 2π and thus (many of) our formula(s) would be simpler if we normalized dt to have total mass 1, but we don't do that in order to avoid confusion, so we have pay by having the factor $1/2\pi$ in front of every integral. It is well known that $L^1(\mathbb{T})$ with $\|\cdot\|_{L^1(\mathbb{T})}$ is a Banach space.

2. A *"trigonometric polynomial"* on $\mathbb{T}$ is an expression of the form
$$P \sim \sum_{n=-m}^{m} a_n e^{int} \qquad \text{where} \qquad m \in \mathbb{N}. \tag{1}$$
   - The number $n \in \mathbb{N}$ above are called the frequencies of P.
   - The largest integer n such that $|a_n| + |a_{-n}| \neq 0$ is called the *"the degree"* of P.

   Since (1) is finite sum, it represents a function, which we denote again by P, defined for each $t \in \mathbb{T}$ by
$$P(t) = \sum_{n=-m}^{m} a_n e^{int} \qquad \text{where} \qquad m \in \mathbb{N}, t \in \mathbb{T}. \tag{2}$$

3. (**Fourier coefficients**) Let P be defined by (2), we can compute the coefficients a_n by the formula
$$a_n = \frac{1}{2\pi}\int_{\mathbb{T}} P(t)e^{-int}\,dt. \qquad \left(\frac{1}{2\pi}\int_{\mathbb{T}} e^{ijt}\,dt = \begin{cases} 1 & \text{if } j = 0, \\ 0 & \text{if } j \neq 0. \end{cases}\right) \tag{3}$$
We shall consider trigonometric polynomials as both formal expressions and functions.

4. A *trigonometric series* on $\mathbb{T}$ is an expression of the form
$$S \sim \sum_{n=-\infty}^{\infty} a_n e^{int} \qquad \text{with the conjugate series is} \qquad \tilde{S} \sim \sum_{n=-\infty}^{\infty} (-i\operatorname{sgn}(n) a_n)\, e^{-int} \tag{4}$$
where $\operatorname{sgn}(n) = 0$ if $n = 0$ and $\operatorname{sgn}(n) = n/|n|$ otherwise.

5. Let $f \in L^1(\mathbb{T})$, motivated from (3) we define the nth Fourier coefficient of f by
$$\mathcal{F}[f](n) = \widehat{f}(n) = \frac{1}{2\pi}\int_{\mathbb{T}} f(t)e^{-int}\,dt \qquad \text{for} \qquad n \in \mathbb{Z}. \tag{5}$$
The *Fourier series* $S[f]$ of a function $f \in L^1(\mathbb{T})$ is the trigonometric series
$$S[f] \sim \sum_{n=-\infty}^{\infty} \widehat{f}(n)e^{int}.$$
The series conjugate to $S[f]$ will be denoted by $\tilde{S}[f]$ and referred to as the conjugate Fourier series of f. We say a trigonometric series is a Fourier series if it is the Fourier series of some $f \in L^1(\mathbb{T})$. It is easy to see that for $f, g \in L^1(\mathbb{T})$ then we have the basic properties as following:

(a) $(\widehat{f+g})(n) = \widehat{f}(n) + \widehat{g}(n)$, and for any complex number α then $(\widehat{\alpha f})(n) = \alpha \widehat{f}(n)$.

(b) If $\overline{f}$ is the conjugate of f, i.e., $\overline{f}(t) = \overline{f(t)}$ then $(\widehat{\overline{f}})(n) = \overline{\widehat{f}(-n)}$.

(c) Denote $f_s(t) = (\tau_s f)(t) = f(t-s)$ for $s \in \mathbb{T}$, then $\widehat{f_s}(n) = \widehat{f}(n)e^{-ins}$.

(d) $|\widehat{f}(n)| \leq \frac{1}{2\pi}\int_{\mathbb{T}} |f(t)|\, dt = \|f\|_{L^1(\mathbb{T})}$. Thus if $f_j \longrightarrow f$ in $L^1(\mathbb{T})$ then $\widehat{f_j}(n) \longrightarrow \widehat{f}(n)$ uniformly.

6. A relation between Fourier coefficient of $f \in L^1(\mathbb{T})$ and its anti-derivative is given by:

Theorem 1.1. *If $f \in L^1(\mathbb{T})$ with $\widehat{f}(0) = 0$, then the function $F(t) = \int_0^t f(s)\, ds$ is continuous, 2π-periodic and $\widehat{F}(n) = \frac{1}{in}\widehat{f}(n)$ for $n \neq 0$.*

Proof. It is obvious that F is (absolutely) continuous. The periodicity of F follows from

$$F(t+2\pi) - F(t) = \int_t^{t+2\pi} f(s)\, ds = 2\pi \widehat{f}(0) = 0.$$

For the second part, we fist assume $f \in C(\mathbb{T}) \cap L^1(\mathbb{T})$, then $F \in C^1(\mathbb{T})$ by the fundamental theorem of calculus and hence we can use integration by part to get

$$\widehat{F}(n) = \frac{1}{2\pi}\int_0^{2\pi} F(t)e^{-int}\, dt = \frac{1}{in}\widehat{f}(n).$$

Now if $f \in L^1(\mathbb{T})$, we can find $f_j \in C(\mathbb{T})$ such that $\|f_j - f\|_{L^1(\mathbb{T})} \longrightarrow 0$, then clearly

$$\frac{1}{2\pi}\int_0^{2\pi} |F_j(t) - F(t)|\, dt \leq \int_0^t |f_j(s) - f(s)|\, ds \leq 2\pi \|f_j - f\|_{L^1(\mathbb{T})} \longrightarrow 0$$

thus $\widehat{F_j}(n) \longrightarrow \widehat{F}(n)$ uniformly, so $\widehat{F}(n) = \lim_{j\longrightarrow\infty} \frac{1}{in}\widehat{f_j}(n) = \frac{1}{in}\widehat{f}(n)$ since $\widehat{f_j}(n) \longrightarrow \widehat{f}(n)$ uniformly. ☐

7. (**Convolution**) Before going to define the convolution on $\mathbb{T}$, we need the following theorem.

Theorem 1.2. *Let $f, g \in L^1(\mathbb{T})$, then for a.e. $t \in \mathbb{T}$ the function $s \longmapsto f(t-s)g(s)$ is integrable on $\mathbb{T}$, and if we define*

$$h(t) = \frac{1}{2\pi}\int_{\mathbb{T}} f(t-s)g(s)\, ds \qquad \Longrightarrow \qquad h \in L^1(\mathbb{T}) \quad \textit{with} \quad \|h\|_{L^1(\mathbb{T})} \leq \|f\|_{L^1(\mathbb{T})}\|g\|_{L^1(\mathbb{T})}$$

and $\widehat{h}(n) = \widehat{f}(n)\widehat{g}(n)$ for all $n \in \mathbb{Z}$.

Proof. The function $F : (t,s) \longmapsto f(t-s)g(s)$ is clearly measurable as a function of (t,s), we have

$$\frac{1}{2\pi}\int_0^{2\pi}\left(\frac{1}{2\pi}\int_0^{2\pi} |f(t-s)g(s)|\, dt\right) ds = \frac{1}{2\pi}\int_0^{2\pi} |g(s)| \cdot \|f\|_{L^1(\mathbb{T})}\, ds = \|f\|_{L^1(\mathbb{T})}\|g\|_{L^1(\mathbb{T})}.$$

By Tonelli's theorem, $F(t,s) \in L^1(\mathbb{T}\times\mathbb{T})$, and hence by Fubini's theorem we have $s \longmapsto f(t-s)g(s)$ is integrable as a function of s for a.e. $t \in \mathbb{T}$, and the order of integration can be switched as

$$\begin{aligned}\|h\|_{L^1(\mathbb{T})}\frac{1}{2\pi}\int_0^{2\pi} |h(t)|\, dt &\leq \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{1}{2\pi}\int_0^{2\pi} |f(t-s)g(s)|\, ds\right) dt \\ &= \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{1}{2\pi}\int_0^{2\pi} |f(t-s)g(s)|\, dt\right) ds \leq \|f\|_{L^1(\mathbb{T})}\|g\|_{L^1(\mathbb{T})}.\end{aligned}$$

Finally we have

$$\begin{aligned}\widehat{h}(n) = \frac{1}{2\pi}\int_0^{2\pi} h(t)e^{-int}\, dt &= \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{1}{2\pi}\int_0^{2\pi} f(t-s)g(s)\, ds\right)e^{-int}\, dt \\ &= \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{1}{2\pi}\int_0^{2\pi} f(t-s)e^{-in(t-s)}\, dt\right)g(s)e^{-ins}\, ds = \widehat{f}(n)\widehat{g}(n)\end{aligned}$$

where all the change in the order of integration is justified by Fubini's theorem. ☐

From that, we define the convolution $f*g$ of two function $f,g \in L^1(\mathbb{T})$ to be another $L^1(\mathbb{T})$ function:

$$(f*g)(t) = \frac{1}{2\pi}\int_0^{2\pi} f(t-s)g(s)\,ds \qquad \text{has} \qquad \widehat{f*g}(n) = \widehat{f}(n)\widehat{g}(n).$$

Theorem 1.3. *The convolution operation in $L^1(\mathbb{T})$ is commutative, associative, and distributive (with respect to the addition).*

Proof. For $f,g \in L^1(\mathbb{T})$ by changing of variable we obtain

$$(f*g)(t) = \frac{1}{2\pi}\int_0^{2\pi} f(t-s)g(s)\,ds = \frac{1}{2\pi}\int_{t-2\pi}^{t} g(t-u)f(u)\,du = \frac{1}{2\pi}\int_0^{2\pi} g(t-u)f(u)\,du = (g*f)(t).$$

Now if $f,g,h \in L^1(\mathbb{T})$ changing of variable we have

$$\begin{aligned}
\big((f*g)*h\big)(t) &= \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{1}{2\pi}\int_0^{2\pi} f(t-s-u)g(u)\,du\right)h(s)\,ds \\
(w = s+u) \quad &= \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{1}{2\pi}\int_0^{2\pi} f(t-w)g(w-s)\,dw\right)h(s)\,ds \\
&= \frac{1}{2\pi}\int_0^{2\pi} f(t-w)\left(\frac{1}{2\pi}\int_0^{2\pi} g(w-s)h(s)\,ds\right)dw = \big(f*(g*h)\big)(t).
\end{aligned}$$

Finally the distribution law $(f+g)*h = f*h + g*h$ is obvious from the definition. □

Theorem 1.4. *Assume $f \in L^1(\mathbb{T})$ and $\varphi(t) = e^{int}$ for some $n \in \mathbb{N}$, then we have $(f*\varphi)(t) = \widehat{f}(n)e^{int}$.*

Proof. Since $f,\varphi \in L^1(\mathbb{T})$ we have $f*\varphi \in L^1(\mathbb{R})$, and $(f*\varphi)(t) = \frac{1}{2\pi}\int_0^{2\pi} f(t-s)e^{ins}\,ds = \widehat{f}(n)e^{int}$. □

As a corollary, if $f \in L^1(\mathbb{T})$ then

$$P(t) = \sum_{-m}^{m} a_n e^{int} \qquad \Longrightarrow \qquad (P*f)(t) = \sum_{-m}^{m} a_n \widehat{f}(n)e^{int}.$$

1.2 Summability kernels and homogeneous Banach spaces on $\mathbb{T}$

1. We shall see that $\widehat{f}$ determines f uniquely and we show how we can find f if we know $\widehat{f}$. First of all let's recall the two important properties of $L^1(\mathbb{T})$, that are

   (H1) (Translation-invariant) If $f \in L^1(\mathbb{T})$ and $s \in \mathbb{T}$ then $t \longrightarrow f_s(t) \in L^1(\mathbb{T})$ and $\|f_s\|_{L^1(\mathbb{T})} = \|f\|_{L^1(\mathbb{T})}$.

   (H2) (Continuity of translation w.r.t L^1-norm) For $f \in L^1(\mathbb{T})$ and $s \in \mathbb{T}$ then $\lim_{s\longrightarrow 0}\|f_s - f\|_{L^1(\mathbb{T})} = 0$.

   The property (H2) follows from the fact that it is true for continuous functions, and by a density argument (continuous functions are dense in $L^1(\mathbb{T})$) we obtain the result.

2. (**Integration of vector-valued functions**) Consider a Banach space $(X, \|\cdot\|)$ and F be a X-valued function, defined and continuous on a compact interval $[a,b] \subset \mathbb{R}$. We define the (Riemann) integral of F on $[a,b]$ in a manner completely analogous to that used in the case of numerical functions, namely for any partition

$$P_N = \{x_0 = a < x_1 < \ldots < x_{N+1} = b\} \qquad \text{we define} \qquad S_{P_N} = \sum_{j=0}^{N}\big(x_{j+1} - x_j\big)F(x_j).$$

   The integral is defined by

$$\int_a^b F(x)\,dx = \lim_{N\longrightarrow\infty} S_{P_N}$$

where the limit is taken in X-norm, and the subdivision $\{x_j : j = 0,1,\ldots,N+1\}$ becomes finer and finer, i.e., as $N \longrightarrow \infty$ we have $\max_{1\le j\le N} |x_{j+1} - x_j| \longrightarrow 0$. The existence of such limit follows by a simple argument , by constructing a sequence of partial sums which is Cauchy (taking the common refinement of two partitions).

3. A *"summability kernel"* is a sequence $\{\zeta_n\}$ of continuous 2π-periodic functions satisfying:

   (S1) $\frac{1}{2\pi}\int_0^{2\pi} \zeta_n(t)\,dt = 1$.

   (S2) $\sup_{n\in\mathbb{N}} \|\zeta_n\|_{L^1(\mathbb{T})} \le C$.

   (S3) For all $0 < \delta < \pi$ we have $\lim_{n\longrightarrow\infty} \int_\delta^{2\pi-\delta} |\zeta_n(t)|\,dt = 0$.

   A positive summability kernel is one such that $\zeta_n(t) \ge 0$ for all t and n. We consider also families ζ_r depending on a continuous parameter r instead of the discrete n. We state the following lemma in a general setting with vector-valued functions.

Lemma 1.5. *Let $(X,\|\cdot\|)$ is a Banach space and $\varphi : \mathbb{T} \longrightarrow (X,\|\cdot\|)$ is continuous, then for any summability kernel $\{\zeta_n\}$ we have*

$$\lim_{n\longrightarrow\infty} \frac{1}{2\pi}\int_0^{2\pi} \zeta_n(t)\varphi(t)\,dt = \varphi(0).$$

Proof. Since φ is continuous on compact set $\mathbb{T}$, it is norm-bounded $\|\varphi(t)\| \le C$ for all $t \in \mathbb{T}$. For $\varepsilon > 0$, there exists $\delta > 0$ such that if $t \in \mathbb{T}$ and $|t| < \delta$ then $|\varphi(t) - \varphi(0)| < \varepsilon$, then we have

$$\frac{1}{2\pi}\int_0^{2\pi} \zeta_n(t)\Big(\varphi(t)-\varphi(0)\Big)\,dt = \frac{1}{2\pi}\int_{\{t\in\mathbb{T}:|t|<\delta\}} \zeta_n(t)\Big(\varphi(t)-\varphi(0)\Big)\,dt + \frac{1}{2\pi}\int_\delta^{2\pi-\delta} \zeta_n(t)\Big(\varphi(t)-\varphi(0)\Big)\,dt.$$

On the other hand

$$\left|\frac{1}{2\pi}\int_{\{t\in\mathbb{T}:|t|<\delta\}} \zeta_n(t)\Big(\varphi(t)-\varphi(0)\Big)\,dt\right| \le 2\varepsilon$$

and by (S3) we have

$$\left|\frac{1}{2\pi}\int_\delta^{2\pi-\delta} \zeta_n(t)\Big(\varphi(t)-\varphi(0)\Big)\,dt\right| \le \frac{C}{\pi}\int_\delta^{2\pi-\delta} |\zeta_n(t)|\,dt$$

as $n \longrightarrow \infty$, which concludes our result. □

As a consequence we have

Theorem 1.6. *Let $f \in L^1(\mathbb{T})$ and $\{\zeta_n\}$ be a summability kernel, then*

$$f = \lim_{n\longrightarrow\infty} \int_0^{2\pi} \zeta_n(t) f_t(\cdot)\,dt \qquad in \qquad L^1(\mathbb{T}).$$

Proof. Let $(X,\|\cdot\|) = \Big(L^1(\mathbb{T}), \|\cdot\|_{L^1(\mathbb{T})}\Big)$ and $\varphi(s) = f_s(\cdot)$ for $s \in \mathbb{T}$, the result follows from lemma 1.5. □

The vector-valued integral above can be understood in the usual sense, by the following lemma.

Lemma 1.7. *Let $\zeta \in C(\mathbb{T})$ (we can relax this condition) and $f \in L^1(\mathbb{T})$ then*

$$\frac{1}{2\pi}\int_0^{2\pi} \zeta(t) f_t(\cdot)\,dt \equiv (\zeta * f)(\cdot)$$

as functions in $L^1(\mathbb{T})$, where on the left hand side we have the vector-valued integral.

Proof 1. Assume first that $f \in C(\mathbb{T})$, we have

$$\frac{1}{2\pi}\int_0^{2\pi} \zeta(t) f_t(\cdot)\, dt = \lim_{\{s_j\}\longrightarrow 0} \frac{1}{2\pi} \sum_j (s_{j+1}-s_j)\zeta(s_j) f_{s_j}$$

where the limit is taken in $L^1(\mathbb{T})$-norm and $\{s_j\} \longrightarrow 0$ means the subdivision $\{s_j\}$ of $[0,2\pi)$ becomes finer and finer. Note that if $f \in C(\mathbb{T})$ then by Riemann sum approximation as usual (which holds for continuous functions, that's why we need $\zeta \in C(\mathbb{T})$) we have

$$(\zeta * f)(t) = \frac{1}{2\pi}\int_0^{2\pi} \zeta(s) f(t-s)\, ds = \lim_{\{s_j\}\longrightarrow 0} \frac{1}{2\pi} \sum_j (s_{j+1}-s_j)\zeta(s_j) f(t-s_j)$$

uniformly for $t \in \mathbb{T}$, thus the lemma is proved if $f \in C(\mathbb{T})$. The case $f \in L^1(\mathbb{T})$ follows by a density argument since $C(\mathbb{T})$ is dense in $L^1(\mathbb{T})$ under the L^1-norm. □

Proof 2. The proof is quite simple if we use some measure theory facts instead of approximating the integral in the Riemann sense. For a continuous function $\varphi : [a,b] \longrightarrow (X, \|\cdot\|)$, then for any $\Lambda \in X^*$ we have

$$\Lambda\left(\int_a^b \varphi(t)\, dt\right) = \int_a^b \Lambda \circ \varphi(t)\, dt$$

and the fact that if $\Lambda(x) = \Lambda(y)$ for all $\Lambda \in X^*$ implies $x \equiv y$, based on a simple application of Hahn-Banach theorem. For $\Lambda \in L^1(\mathbb{T})^*$, since $(L^1)^* = L^\infty$, there exists a unique $g \in L^\infty(\mathbb{T})$ such that $\Lambda(\varphi) = \int_0^{2\pi} g(t)\varphi(t)\, dt$ for all $\varphi \in L^1(\mathbb{T})$, thus

$$\begin{aligned}
\Lambda\left(\frac{1}{2\pi}\int_0^{2\pi} \zeta(t) f_t(\cdot)\, dt\right) &= \frac{1}{2\pi}\int_0^{2\pi} \zeta(t)\Lambda\big(f(\cdot - t)\big)\, dt \\
&= \frac{1}{2\pi}\int_0^{2\pi} \zeta(t)\left(\int_0^{2\pi} g(u) f(u-t)\, du\right) dt \\
\text{(Fubini's theorem)} \quad &= \frac{1}{2\pi}\int_0^{2\pi} g(u)\left(\int_0^{2\pi} \zeta(t) f(u-t)\, dt\right) du = \frac{1}{2\pi}\int_0^{2\pi} g(u)\big(\zeta * f\big)(u)\, du = \Lambda(\zeta * f).
\end{aligned}$$

Thus the proof is complete, note that in this way we don't need to use the continuity of ζ. □

Using this lemma, for any summability kernel $\{\zeta_n\}$ we have $\zeta_n * f \longrightarrow f$ in $L^1(\mathbb{T})$.

4. (**Fejer's kernel**) The Fejer's kernel is defined by

$$\mathcal{K}_n(t) = \sum_{j=-n}^{n} \left(1 - \frac{|j|}{n+1}\right) e^{ijt}.$$

It is clear that $\|\mathcal{K}_n\|_{L^1(\mathbb{T})} = 1$, the first and the third properties of a summability kernel is verified using the formula $\sin(a) - \sin(b) = 2\cos\left(\frac{a+b}{2}\right)\sin\left(\frac{a-b}{2}\right)$. Indeed, let $\mathcal{D}_n(t) = \sum_{j=-n}^{n} e^{ijt}$, which is called "**Dirichlet kernel**" (not a summability kernel), we have

$$\mathcal{D}_n(t) = \sum_{j=-n}^{n} e^{ijt} = 1 + 2\sum_{j=1}^{n} \cos(jt) = 1 + \frac{\left(\sum_{j=1}^{n}\left[\sin\left(j+\frac{1}{2}\right)t - \sin\left(j-\frac{1}{2}\right)t\right]\right)}{\left(\sin\frac{t}{2}\right)} = \frac{\sin\left(n+\frac{1}{2}\right)t}{\sin\frac{t}{2}}.$$

Thus

$$(n+1)\mathcal{K}_n(t) = \sum_{j=-n}^{n}\big(n+1-|j|\big)e^{ijt} = \sum_{k=0}^{n} \mathcal{D}_k(t) = \left(\sin\frac{t}{2}\right)^{-1} \operatorname{Im}\left[e^{\frac{it}{2}} \sum_{k=0}^{n} e^{ikt}\right] = \left(\sin\frac{t}{2}\right)^{-1} \operatorname{Im}\left[e^{\frac{it}{2}} \frac{1-e^{i(n+1)t}}{1-e^{it}}\right].$$

It is easy to compute the last sum at get

$$\mathcal{K}_n(t) = \frac{1}{n+1}\left(\sin\frac{t}{2}\right)^{-1} \operatorname{Im}\left[\frac{1-e^{i(n+1)t}}{e^{-\frac{it}{2}} - e^{\frac{it}{2}}}\right] = \frac{1}{n+1}\left(\frac{1-\cos(n+1)t}{2\sin^2\left(\frac{t}{2}\right)}\right) = \frac{1}{n+1}\left(\frac{\sin\left(\frac{n+1}{2}\right)t}{\sin\frac{t}{2}}\right)^2.$$

We adhere to the generally used notation and write $\sigma_n(f) = \mathcal{K}_n * f$ and $\sigma_n(f)(t) = (\mathcal{K}_n * f)(t)$. It is clear from theorem 1.4 that

$$\sigma_n(f)(t) = (\mathcal{K}_n * f)(t) = \sum_{j=-n}^{n} \left(1 - \frac{|j|}{n+1}\right) \widehat{f}(j) e^{ijt}. \tag{6}$$

We already knew that $\sigma_n(f) \longrightarrow f$ in $L^1(\mathbb{T})$ for every $f \in L^1(\mathbb{T})$. Note that $\sigma_n(f)$ is a trigonometric polynomial, thus from this fact we deduce that trigonometric polynomials are dense in $L^1(\mathbb{T})$.

Theorem 1.8 (Uniqueness). *If $f \in L^1(\mathbb{T})$ has $\widehat{f}(n) = 0$ for all $n \in \mathbb{Z}$ then $f \equiv 0$.*

Proof. It is obvious from (6) and the fact that $\sigma_n(f) \longrightarrow f$ in $L^1(\mathbb{T})$. □

Theorem 1.9 (Riemann-Lebesgue lemma). *If $f \in L^1(\mathbb{T})$ then $\lim_{|n| \longrightarrow \infty} \widehat{f}(n) = 0$. Moreover, if K is a compact subset of $L^1(\mathbb{T})$ then $\lim_{|n| \longrightarrow \infty} \left(\sup_{f \in K} |\widehat{f}(n)| \right) = 0$.*

Proof. Let P be a trigonometric polynomial with $\|f - P\|_{L^1(\mathbb{T})} < \varepsilon$, then for $n \in \mathbb{Z}$ with $|n| > \deg(P)$ we have $\widehat{P}(n) = 0$, hence $|\widehat{f}(n)| = |(\widehat{f - P})(n)| \leq \|f - P\|_{L^1(\mathbb{T})} < \varepsilon$. If K is a compact subset of $L^1(\mathbb{T})$ and $\varepsilon > 0$, there exists a finite number of trigonometric polynomials $P_1, \ldots, P_m$ such that for any $f \in K$ there exists $j \in \{1, \ldots, m\}$ such that $\|f - P_j\|_{L^1(\mathbb{T})} < \varepsilon$. The argument follows similarly as before with $|n| > \max\{\deg P_j : j = 1, 2, \ldots, m\}$. □

In summary, $\{\mathcal{K}_n\}$ is a positive summability kernel which possess the following properties:

$$\lim_{n \longrightarrow \infty} \left(\sup_{\delta < t < 2\pi - \delta} \mathcal{K}_n(t) \right) = 0 \qquad \text{for any} \qquad 0 < \delta < \pi \tag{F1}$$

and

$$\mathcal{K}_n(t) = \mathcal{K}_n(-t). \tag{F2}$$

5. For $f \in L^1(\mathbb{T})$ we denote by $S_n(f)$ the n partial sum of $S[f]$, that is

$$S_n(f)(t) = S_n(f, t) = \sum_{j=-n}^{n} \widehat{f}(j) e^{ijt}, \qquad \text{i.e.,} \qquad S_n(f) = \mathcal{D}_n * f.$$

We can see that

$$\sigma_n(f) = \frac{1}{n+1} \sum_{k=0}^{n} S_n(f),$$

which are the Cesàro means of $S_n(f)$. Cesàro mean theorem says that if $S_n(f) \longrightarrow g$ in $L^1(\mathbb{T})$ as $n \longrightarrow \infty$ then $\sigma_n(f) \longrightarrow g$ in $L^1(\mathbb{T})$ as $n \longrightarrow \infty$ as well, which follows that $f = g$. Since the Dirichlet kernel $\{\mathcal{D}_n\}$ doesn't satisfy (S2) or (S3), this explains why the problem of convergence for Fourier series is so much harder than the problem of summability.

6. (**Homogeneous Banach spaces on** $\mathbb{T}$) A homogeneous Banach space on $\mathbb{T}$ is a linear subspace $B \subset L^1(\mathbb{T})$ having a norm $\|\cdot\|_B \geq \|\cdot\|_{L^1}$ under which it is a Banach space, and having the following properties:

(H1) (Translation-invariant) $f \in B$ and $s \in \mathbb{T}$ implies $f_s \in B$ and $\|f_s\|_B = \|f\|_B$.

(H2) (Continuity of translation) For all $f \in B$ and $s, t \in \mathbb{T}$ we have $\lim_{t \longrightarrow s} \|f_t - f_s\|_B = 0$.

If we have a space B satisfying (H1) and we want to show it satisfies (H2) as well, it is sufficient to check the continuity of the translation on a dense subset of B.

Lemma 1.10. *Let $B \subset L^1(\mathbb{T})$ be a Banach space satisfying (H1). Denote by B_c the set of all $f \in B$ such that $s \longmapsto f_s$ is a continuous B-valued function, then B_c is a closed subspace of B.*

Proof. Assume $f \in \overline{B_c}$ where the closure is taken in $(B, \|\cdot\|_B)$. Given $\varepsilon > 0$, there exists $g \in B_c$ such that $\|f - g\|_B < \varepsilon$, then

$$\|f_s - f\|_B \leq \|f_s - g_s\|_B + \|g_s - g\|_B + \|g - f\|_B = 2\|f - g\|_B + \|g_s - g\|_B < 2\varepsilon + \|g_s - g\|_B$$

which can be made less then 3ε if we choose s small enough. ☐

Examples of homogeneous Banach spaces on $\mathbb{T}$,

(a) $C(\mathbb{T})$-the space of all continuous 2π-periodic functions with the norm

$$\|f\|_u = \|f\|_\infty = \max_{t\in\mathbb{T}} |f(t)|.$$

(b) $C^n(\mathbb{T})$-the subspace of $C(\mathbb{T})$ of all n-times continuously differentiable functions ($n \in \mathbb{N}$) with the norm

$$\|f\|_{C^n(\mathbb{T})} = \sum_{k=0}^{n} \frac{1}{k!} \max_{t\in\mathbb{T}} |f^{(k)}(t)|.$$

(c) $L^p(\mathbb{T})$, $1 \leq p < \infty$-the subspace of $L^1(\mathbb{T})$ consisting of all the functions f for which $\int_{\mathbb{T}} |f(t)|^p \, dt$ is finite with the norm

$$\|f\|_{L^p(\mathbb{T})} = \left(\frac{1}{2\pi} \int_0^{2\pi} |f(t)|^p \, dt \right)^{\frac{1}{p}}.$$

Checking (H2) for (a),(b) is equivalent to the fact that continuous functions on $\mathbb{T}$ are uniformly continuous, while checking (H2) for (c) is similar to the L^1-case. Now we extend some results to the homogeneous Banach spaces on $\mathbb{T}$.

Theorem 1.11. *Let B be a homogeneous Banach space on $\mathbb{T}$, let $f \in B$ and $\{\zeta_n\}$ be a summability kernel, then $\|\zeta_n * f - f\|_B \longrightarrow 0$ as $n \longrightarrow \infty$.*

Proof. By definition we have

$$\lim_{\{s_j\}\longrightarrow 0} \left\| \frac{1}{2\pi} \int_0^{2\pi} \zeta(t) f_t(\cdot) \, dt - \frac{1}{2\pi} \sum_j (s_{j+1} - s_j) \zeta(s_j) f_{s_j} \right\|_B$$

where the limit is taken in B-norm and $\{s_j\} \longrightarrow 0$ means the subdivision $\{s_j\}$ of $[0, 2\pi)$ becomes finer and finer. Since $\|\cdot\|_{L^1} \leq \|\cdot\|_B$, it happens that

$$\underbrace{\frac{1}{2\pi} \int_0^{2\pi} \zeta(t) f_t(\cdot) \, dt}_{B\text{-valued}} \equiv \underbrace{\frac{1}{2\pi} \int_0^{2\pi} \zeta(t) f_t(\cdot) \, dt}_{L^1(\mathbb{T})\text{-valued}}.$$

By lemma 1.7 they all equal to $\zeta * f$ as a function in B. The conclusion now follows from lemma 1.5 with $\varphi(s) = f_s = f(\cdot - s)$. ☐

Theorem 1.12. *Let B be a homogeneous Banach space on $\mathbb{T}$, then the trigonometric polynomials in B are everywhere dense.*

Proof. For every $f \in B$ we have $\sigma_n(f) \longrightarrow f$ in $(B, \|\cdot\|_B)$, and since $\sigma_n(f)$ is a trigonometric polynomial in B, we have the conclusion. ☐

7. (**de la Vallée Poussin kernel**) The de la Vallée Poussin kernel is defined by

$$\mathcal{V}_n(t) = 2\mathcal{K}_{2n+1}(t) - \mathcal{K}_n(t).$$

It is obvious that $\{\mathcal{V}_n\}$ is a summability kernel from the fact that $\{\mathcal{K}_n\}$ is a summability kernel, it is a polynomial of degree $2n+1$ having the property that $\widehat{\mathcal{V}_n}(j) = 1$ if $|j| \leq n+1$. It is therefore useful when we want to approximate a function f by polynomials having the same Fourier coefficients as f over prescribed intervals (namely $\mathcal{V}_n * f$).

8. (**Poisson kernel**) For $0 < r < 1$ out

$$P_r(t) = \sum_{j\in\mathbb{Z}} r^{|j|} e^{ijt} = 1 + 2\sum_{j=1}^{\infty} r^j \cos(jt) = \frac{1-r^2}{1-2r\cos t + r^2}.$$

1.3 Point-wise convergence of $\sigma_n(f)$

1. We have already known that if $f \in L^1(\mathbb{T})$ then $\sigma_n(f) \longrightarrow f$ in the topology of any homogeneous Banach space that contains f. In particular if $f \in C(\mathbb{T})$ then $\sigma_n(f) \longrightarrow f$ uniformly. In case f is not continuous, we have to reexamine the integrals defining $\sigma_n(f)$ for point-wise convergence.

Theorem 1.13 (Fejér). *Let $f \in L^1(\mathbb{T})$.*

(a) Assume (Fejér condition)

$$\lim_{h\longrightarrow 0}\Big(f(t_0+h) + f(t_0-h)\Big) \quad \textit{exists, which can be } \pm\infty$$

then

$$\lim_{n\longrightarrow\infty} \sigma_n(f)(t_0) = \frac{1}{2}\lim_{h\longrightarrow 0}\Big(f(t_0+h)+f(t_0-h)\Big).$$

In particular, if t_0 is a point of continuity of f then $\sigma_n(f)(t_0) \longrightarrow f(t_0)$.

(b) If every point of a closed interval I is a point of continuity for f, $\sigma_n(f)(t) \longrightarrow f(t)$ uniformly on I.

(c) If for a.e. t, $m \le f(t)$ then $m \le \sigma_n(f)(t)$. If for a.e. t, $f(t) \le M$, then $\sigma_n(f)(t) \le M$.

Proof. We assume first that $\tilde{f}(t_0) = \lim_{h\longrightarrow 0} \frac{f(t_0+h)+f(t_0-h)}{2}$ is finite. From (F2) we have

$$\begin{aligned}\sigma_n(f)(t_0) - \tilde{f}(t_0) &= \frac{1}{2\pi}\int_{\mathbb{T}} \mathcal{K}_n(s)\Big(f(t_0-s) - \tilde{f}(t_0)\Big)\, ds \\ &= \frac{1}{\pi}\int_{[0,\delta]\cup[\delta,\pi)} \mathcal{K}_n(s)\left(\frac{f(t_0+s)+f(t_0-s)}{2} - \tilde{f}(t_0)\right) ds.\end{aligned}$$

Given $\varepsilon > 0$, we choose $\delta > 0$ such that $|h| < \delta$ implies $\left|\frac{f(t_0+h)+f(t_0-h)}{2} - \tilde{f}(t_0)\right| < \varepsilon$, then we have

$$\frac{1}{\pi}\int_0^{\delta} \mathcal{K}_n(s)\left|\frac{f(t_0+s)+f(t_0-s)}{2} - \tilde{f}(t_0)\right| ds < \varepsilon. \tag{7}$$

Now from (F1) we can choose $n_0 \in \mathbb{N}$ such that $n > n_0$ implies $\sup_{\delta<t<2\pi-\delta} \mathcal{K}_n(t) < \varepsilon$, which implies

$$\frac{1}{\pi}\int_{\delta}^{\pi} \mathcal{K}_n(s)\left|\frac{f(t_0+s)+f(t_0-s)}{2} - \tilde{f}(t_0)\right| ds < \varepsilon\|f - \tilde{f}(t_0)\|_{L^1(\mathbb{T})}. \tag{8}$$

From (7) and (8) we deduce that

$$\Big|\sigma_n(f) - \tilde{f}(t_0)\Big| < \varepsilon + \varepsilon\|f - \tilde{f}(t_0)\|_{L^1(\mathbb{T})}$$

which proves part (a) when $\tilde{f}(t_0)$ is finite. It is easy to see that the same argument holds when $\tilde{f}(x_0) = \pm\infty$. For part (b), if f is continuous at every points in a closed interval I then f is uniformly continuous on I, then we can modify the proof above as given $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|h| < \delta \qquad \Longrightarrow \qquad \sup_{t\in I}\left|\frac{f(t-h)+f(t+h)}{2} - \tilde{f}(t)\right| < \varepsilon$$

and the argument above can be applied again to get (b). Part (c) follows from the fact that $\mathcal{K}_n$ is positive and $\|\mathcal{K}_n\|_{L^1(\mathbb{T})} = 1$, indeed if $m \le f$ a.e. then

$$\sigma_n(f)(t) - m = \frac{1}{2\pi}\int_{\mathbb{T}} \mathcal{K}_n(s)\Big(f(t-s) - m\Big)\, ds \ge 0$$

and similarly for the case $f \le M$. □

2. As corollary, if $f \in L^1(\mathbb{T})$ is continuous at t_0 and if the Fourier series of f converges at t_0 then its sum is $f(t_0)$.

3. The theorem still valid if we replace $\sigma_n(f)$ by $\zeta_n * f$ where $\{\zeta_n\}$ is a summability kernel which satisfies two properties (F1) and (F2). In particular, the Possion kernel satisfies all of these requirements and the statement of Fejér theorem remains valid if we replace $\sigma_n(f)$ by the Abel means of the Fourier series of f,

4. The Fejér's condition

$$\tilde{f}(t_0) = \frac{1}{2}\lim_{h\longrightarrow 0}\Big(f(t_0+h)+f(t_0-h)\Big) \tag{9}$$

implies that

$$\lim_{h\longrightarrow 0}\frac{1}{h}\int_0^h \left|\frac{f(t_0+s)+f(t_0-s)}{2}-\tilde{f}(t_0)\right| ds = 0. \tag{10}$$

The condition (10) is far less restrictive and more natural for summable functions, since it doesn't change if we modify f on a set of measure zero.

Theorem 1.14 (Lebesgue)**.** *If (10) holds then $\sigma_n(f)(t_0) \longrightarrow f(t_0)$. In particular $\sigma_n(f)(t) \longrightarrow f(t)$ a.e.*

Proof. We have

$$\sigma_n(f)(t_0) - \tilde{f}(t_0) = \frac{1}{\pi}\int_{[0,\delta]\cup[\delta,\pi)} \mathcal{K}_n(s)\left(\frac{f(t_0+s)+f(t_0-s)}{2}-\tilde{f}(t_0)\right) ds.$$

Recall that

$$\mathcal{K}_n(s) = \frac{1}{n+1}\left(\frac{\sin(n+1)\frac{s}{2}}{\sin\frac{s}{2}}\right)^2 \le \min\left\{n+1, \frac{\pi^2}{(n+1)s^2}\right\}$$

where we have used $\sin\frac{x}{2} \ge \frac{x}{\pi}$ for $0 \le x \le \pi$. From that we have

$$\frac{1}{\pi}\int_\delta^\pi \mathcal{K}_n(s)\left|\frac{f(t_0+s)+f(t_0-s)}{2}-\tilde{f}(t_0)\right| ds \le \min\left\{n+1, \frac{\pi^2}{(n+1)\delta^2}\right\}\frac{\|f-\tilde{f}(t_0)\|_{L^1(\mathbb{T})}}{\pi}$$

which converges to 0 if $(n+1)\delta^2 \longrightarrow +\infty$ as $n \longrightarrow \infty$. Let's pick $\delta = n^{-1/4}$, we have left to show that

$$\lim_{n\longrightarrow\infty}\frac{1}{\pi}\int_0^{n^{-1/4}} \mathcal{K}_n(s)\left(\frac{f(t_0+s)+f(t_0-s)}{2}-\tilde{f}(t_0)\right) ds = 0. \tag{11}$$

Let's define for simplicity the function

$$\Phi(h) = \int_0^h \left|\frac{f(t_0+s)+f(t_0-s)}{2}-\tilde{f}(t_0)\right| ds \quad \text{then} \quad \lim_{h\longrightarrow 0}\frac{\Phi(h)}{h} = 0.$$

For given $\varepsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that $\Phi(s) < \varepsilon s$ for $0 < s < n^{-1/4}$ for all $n \ge n_0$. Also

$$\frac{1}{\pi}\int_0^{n^{-1}} \mathcal{K}_n(s)\left|\frac{f(t_0+s)+f(t_0-s)}{2}-\tilde{f}(t_0)\right| ds \le \frac{n+1}{\pi}\int_0^{\frac{1}{n}} \left|\frac{f(t_0+h)+f(t_0-h)}{2}-\tilde{f}(t_0)\right| ds \longrightarrow 0$$

as $n \longrightarrow \infty$ by (10). Now

$$\frac{1}{\pi}\int_{n^{-1}}^{n^{-1/4}} \mathcal{K}_n(s)\left|\frac{f(t_0+s)+f(t_0-s)}{2}-\tilde{f}(t_0)\right| ds \le \frac{\pi}{n+1}\int_{n^{-1}}^{n^{-1/4}} \left|\frac{f(t_0+h)+f(t_0-h)}{2}-\tilde{f}(t_0)\right| \frac{1}{s^2}\, ds$$

Using a generalized version of the integration by parts formula we obtain

$$\begin{aligned}\frac{\pi}{n+1}\int_{n^{-1}}^{n^{-1/4}} \left|\frac{f(t_0+h)+f(t_0-h)}{2}-\tilde{f}(t_0)\right| \frac{1}{s^2}\, ds &= \frac{\pi}{n+1}\left(\frac{\Phi(s)}{s^2}\right)\Bigg|_{s=n^{-1}}^{s=n^{-1/4}} + \frac{2\pi}{n+1}\int_{n^{-1}}^{n^{-1/4}} \frac{\Phi(s)}{s^3}\, ds \\ &\le 2\pi\varepsilon\left(\frac{n}{n+1}\right) + \frac{2\pi\varepsilon}{n+1}\int_{n^{-1}}^{n^{-1/4}} \frac{1}{s^2}\, ds < 6\pi\varepsilon.\end{aligned}$$

The proof is complete. □

As a consequence, if the Fourier series of $f \in L^1(\mathbb{T})$ converges on a set E of positive measure, its sum coincides with f almost everywhere on E. In particular, if a Fourier series converges to zero almost everywhere, all its coefficients must vanish.

1.4 The order of magnitude of Fourier coefficients

Two things we have known about the size of Fourier coefficients are if $f \in L^1(\mathbb{T})$ then $\|\widehat{f}\|_{L^\infty(\mathbb{T})} \leq \|\widehat{f}\|_{L^1(\mathbb{T})}$ and the Riemann-Lebesgue lemma: $\lim_{|n| \longrightarrow \infty} \widehat{f}(n) = 0$.

1. Can the Riemann-Lebesgue lemma be improved to provide a certain rate of vanishing of $\widehat{f}(n)$ as $|n| \longrightarrow \infty$? The answer is no.

   Theorem 1.15. *Let $\{a_n\}_{n\in\mathbb{Z}}$ be a even sequence of non-negative numbers tending to zero at infinity. Assume that for $n > 0$ we have*
   $$a_{n-1} + a_{n+1} - 2a_n \geq 0. \tag{12}$$
   Then there exists a non-negative $f \in L^1(\mathbb{T})$ such that $\widehat{f}(n) = a_n$.

2. A basic difference between sine-series and cosine-series is given by:

   Theorem 1.16. *If $f \in L^1(\mathbb{T})$ and $\widehat{f}(|n|) = -\widehat{f}(-|n|) \geq 0$ for all $n \in \mathbb{Z}$ then*
   $$\sum_{n=1}^{\infty} \frac{1}{n} \widehat{f}(n) < \infty.$$

   Proof. Assume $\widehat{f}(0) = 0$, let $F(t) = \int_0^t f(s)\,ds$, by theorem 1.1 we have $F \in C(\mathbb{T})$ with $\widehat{F}(n) = \frac{1}{in}\widehat{f}(n)$ for $n \neq 0$. Since F is continuous, we can apply Fejer's theorem 1.13 to obtain
   $$\lim_{m\longrightarrow\infty} \sigma_m(F)(0) = \lim_{m\longrightarrow\infty} \sum_{-m}^{m} \left(1 - \frac{n}{m+1}\right) \frac{1}{in} \widehat{f}(n) = F(0).$$
   I.e.,
   $$\lim_{m\longrightarrow\infty} 2\sum_{n=1}^{m} \left(1 - \frac{n}{m+1}\right) \frac{1}{n} \widehat{f}(n) = i\big(F(0) - \widehat{F}(0)\big) = -i\widehat{F}(0).$$
   Since $\frac{1}{n}\widehat{f}(n) \geq 0$ for $n = 1, 2, \ldots$ the proof is complete. □

3. We now turn to some simple results about the order of magnitude of Fourier coefficients of functions satisfying various smoothness conditions.

   Theorem 1.17. *If $f \in L^1(\mathbb{T})$ is absolutely continuous, then $\widehat{f}(n) = o\left(\frac{1}{n}\right)$ as $|n| \longrightarrow \infty$.*

   Proof. f' exists and $f' \in L^1(\mathbb{T})$ with $\widehat{f}(n) = \frac{1}{in}\widehat{f'}(n)$, and thus $\widehat{f}(n) = o\left(\frac{1}{n}\right)$ since $\widehat{f'}(n) \longrightarrow 0$ as $|n| \longrightarrow \infty$ by Riemann-Lebesgue lemma. □

   Similarly, if f is k-times differentiable and $f^{(k-1)}$ is absolutely continuous then $\widehat{f}(n) = o(n^{-k})$ as $|n| \longrightarrow \infty$. Similarly, we have:

   Theorem 1.18. *If f is k-times differentiable and $f^{(k-1)}$ is absolutely continuous then*
   $$|\widehat{f}(n)| \leq \min_{0\leq j\leq k} \frac{\|f^j\|_{L^1(\mathbb{T})}}{|n|^j}$$
   In particular, if $f \in C^\infty(\mathbb{T})$ then
   $$|\widehat{f}(n)| \leq \min_{0\leq j} \frac{\|f^j\|_{L^1(\mathbb{T})}}{|n|^j}$$

1.5 Fourier coefficients of linear functionals

1. Let B be a homogeneous Banach space on $\mathbb{T}$ and let's assume that $e^{int} \in B$ for all $n \in \mathbb{Z}$, we denote by B^* be the dual space of B. The Fourier coefficients of a functional $\mu \in B^*$ are defined by

$$\widehat{\mu}(n) = \overline{\mu(e^{int})} = \overline{\langle e^{int}, \mu\rangle}, \qquad n \in \mathbb{Z}. \tag{13}$$

The Fourier series of μ is defined by

$$S[\mu] \sim \sum_{n=-\infty}^{\infty} \widehat{\mu}(n) e^{int}.$$

It is clear that $|\widehat{\mu}(n)| \le \|\mu\|_{B^*} \|e^{in(\cdot)}\|_B$.

2. For $1 \le p < \infty$, recall that $(L^p)^* = L^q$ for $1 < q \le \infty$ where $\frac{1}{p} + \frac{1}{q} = 1$. A function $\mu \in L^q$ can be identified with the linear function

$$\mu : L^p \longmapsto \mathbb{C} \qquad \text{map} \qquad f \longmapsto \mu(f) = \langle f, \mu\rangle_{L^q, L^p} = \frac{1}{2\pi} \int_{\mathbb{T}} f(x)\overline{\mu(x)}\, dx.$$

The definition (13) reads

$$\widehat{\mu}(n) = \overline{\langle e^{int}, \mu\rangle_{L^q, L^p}} = \overline{\frac{1}{2\pi} \int_{\mathbb{T}} e^{int}\overline{\mu(x)}\, dt} = \frac{1}{2\pi} \int_{\mathbb{T}} \mu(x) e^{-int}\, dt$$

which is consistent with our previous definition of Fourier coefficient for a function.

3. (Parseval's formula)

Theorem 1.19 (Parseval's formula)**.** *Let $f \in B$ and $\mu \in B^*$, then*

$$\langle f, \mu\rangle = \lim_{m \longrightarrow \infty} \sum_{-m}^{m} \left(1 - \frac{|n|}{m+1}\right) \widehat{f}(n)\overline{\widehat{\mu}(n)}. \tag{14}$$

Proof. From theorem 1.11 we have $\sigma_n(f) \longrightarrow f$ in B norm, and since $\sigma_n(f)$ is a trigonometric polynomial, which gives us

$$\mu(S_k(f)) = \langle S_k(f), \mu\rangle = \left\langle \sum_{j=-k}^{k} \widehat{f}(j) e^{ijt}, \mu \right\rangle = \sum_{j-=-k}^{k} \widehat{f}(j) \langle e^{ijt}, \mu\rangle = \sum_{j-=-k}^{k} \widehat{f}(j)\overline{\widehat{\mu}(j)}.$$

Thus we have

$$\sigma_n(f) = \frac{1}{n+1} \sum_{k=0}^{n} S_n(f) \quad \Longrightarrow \quad \langle \sigma_n(f), \mu\rangle = \frac{1}{n+1} \sum_{k=0}^{n} \sum_{j-=-k}^{k} \widehat{f}(j)\overline{\widehat{\mu}(j)} = \sum_{-n}^{n} \left(1 - \frac{|j|}{n+1}\right) \widehat{f}(j)\overline{\widehat{\mu}(j)}.$$

Taking the limit as $n \longrightarrow \infty$ we obtain the result. □

If the series on the right hand side of (14) converges then

$$\langle f, \mu\rangle = \lim_{m \longrightarrow \infty} \sum_{-\infty}^{\infty} \left(1 - \frac{|n|}{m+1}\right) \widehat{f}(n)\overline{\widehat{\mu}(n)}. \tag{15}$$

From that we have the uniqueness theorem

Theorem 1.20 (Uniqueness)**.** *If $\mu \in B^*$ and $\widehat{\mu}(n) = 0$ for all $n \in \mathbb{Z}$ then $\mu \equiv 0$.*

4. For $\mu \in B^*$ we define

$$S_n(\mu)(\cdot) = \sum_{-n}^{n} \widehat{\mu}(j) e^{ijt} \qquad (\sim \mathcal{D}_n * \mu),$$

$$\sigma_n(\mu)(\cdot) = \sum_{-n}^{n} \left(1 - \frac{|j|}{n+1}\right) \widehat{\mu}(j) e^{ijt} \qquad (\sim \mathcal{K}_n * \mu).$$

We still have

$$\sigma_n(\mu) = \frac{1}{n+1} \sum_{k=0}^{n} S_n(\mu).$$

They are elements of B^* by the actions

$$\langle f, S_n(\mu)(\cdot)\rangle = \frac{1}{2\pi}\int_{\mathbb{T}} f(t)\overline{S_n(\mu)(t)}\,dt = \sum_{-n}^{n} \overline{\widehat{\mu}(j)} \left(\frac{1}{2\pi}\int_{\mathbb{T}} f(t) e^{-ijt}\,dt\right) = \sum_{-n}^{n} \widehat{f}(j)\overline{\widehat{\mu}(j)}$$

and similarly

$$\langle f, \sigma_n(\mu)(\cdot)\rangle = \frac{1}{2\pi}\int_{\mathbb{T}} f(t)\overline{\sigma_n(\mu)(t)}\,dt = \sum_{-n}^{n} \left(1 - \frac{|j|}{n+1}\right) \widehat{f}(j)\overline{\widehat{\mu}(j)}$$

for all $f \in B$. We have some remarks:

(a) From the Parseval's formula 1.19 for any $\mu \in B^*$ then $\sigma_n(\mu) \overset{*}{\rightharpoonup} \mu$ in the weak* topology of B^*. If $B = C(\mathbb{T})$ then as a measure, $\langle f, \sigma_n(f)\rangle = \frac{1}{2\pi}\int_{\mathbb{T}} f\, \overline{d\sigma_n(f)}$ which means

$$d\sigma_n(\mu) = \sum_{-n}^{n} \left(1 - \frac{|j|}{n+1}\right) \widehat{\mu}(j) e^{ijt}\,dt.$$

We also observe that from that $\sigma_n(\mu) \ll m$ where m is the Lebesgue measure.

(b) The linear operator $\mathbb{S}_n : B \longrightarrow B$ maps $f \longmapsto S_n(f)$ is bounded, since for $f \in B$ and recall that $e^{int} \in B$,

$$\|S_n(f)\|_B = \left\| \sum_{j=-n}^{n} \widehat{f}(j) e^{ijt} \right\|_B \le \sum_{j=-n}^{n} |\widehat{f}(j)| \cdot \|e^{ijt}\|_B \le \left(\sum_{-n}^{n} \|e^{ijt}\|_B \right) \|f\|_{L^1} \le \left(\sum_{-n}^{n} \|e^{ijt}\|_B \right) \|f\|_B.$$

(c) The linear operator $\mathbb{S}_n^* : B^* \longrightarrow B^*$ map $\mu \longmapsto S_n(\mu)$ is the adjoint operator of $\mathbb{S}_n$, since for any $\mu \in B^*$ we have for all $f \in B$ then

$$\mathbb{S}_n^*(\mu)(f) = \langle f, S_n(\mu)\rangle = \sum_{j=-n}^{n} \widehat{f}(j)\overline{\widehat{\mu}(j)}\widehat{\mu}(j) = \langle S_n(f), \mu\rangle = \mu \circ (\mathbb{S}_n)(f)$$

and thus $\mathbb{S}_n^* \in B^{**}$ with $\|\mathbb{S}_n^*\|_{B^{**}} = \|\mathbb{S}_n\|_{B^*}$.

(d) Similarly, $\Sigma_n : B \longrightarrow B$ maps $f \longmapsto \sigma_n(f)$ belongs to B^* and $\Sigma_n^* : B^* \longrightarrow B^*$ maps $\mu \longmapsto \sigma_n(\mu)$ is the adjoint of Σ_n, thus $\|\Sigma_n^*\|_{B^*_*} = \|\Sigma_n\|_{B^*}$. Indeed we have $\|\Sigma_n^*\|_{B^{**}} = \|\Sigma_n\|_{B^*} = 1$ since

$$\|\sigma_n(f)\|_B = \left\| \frac{1}{2\pi}\int_{\mathbb{T}} \mathcal{K}_n(t) f(\cdot - t)\,dt \right\|_B \le \frac{1}{2\pi}\int_{\mathbb{T}} \mathcal{K}_n(t)\|f\|_B\,dt = \|\mathcal{K}\|_{L^1(\mathbb{T})}\|f\|_B = \|f\|_B$$

thus $\|\Sigma_n\|_{B^*} \le 1$. On the other hand by testing with $e^{i0t} = 1$ in B we have

$$\|\sigma_n(1)\|_B = \left\| e^{i0t} \frac{1}{2\pi}\int_{\mathbb{T}} \mathcal{K}_n(t)\,dt \right\|_B = \left| \frac{1}{2\pi}\int_{\mathbb{T}} \mathcal{K}_n(t)\,dt \right| \cdot \|e^{i0t}\|_B = \|e^{i0t}\|_B.$$

Thus $\|\Sigma_n\|_{B^*} \ge 1$.

Theorem 1.21. *If $\mu \in B^*$, the linear operator $\Sigma_n^* : B^* \longrightarrow B^*$ maps*

$$\mu \longmapsto \Sigma_n^*(\mu) = \sigma_n(\mu) \qquad \textit{which has its Fourier series is} \qquad \sum_{-n}^{n}\left(1 - \frac{|j|}{n+1}\right)\widehat{\mu}(j)e^{ijt}$$

satisfies $\|\Sigma_n^*\|_{B^{**}} = 1$. In particular $\|\sigma_n(\mu)\|_{B^*} \le \|\mu\|_{B^*}$ for all $n \in \mathbb{Z}$.

5. Parseval's formula enables us to characterize sequences of Fourier coefficients of linear functionals.

Theorem 1.22. *Let B be a homogeneous Banach space on $\mathbb{T}$. Assume that $e^{int} \in B$ for all $n \in \mathbb{N}$. Let $\{a_n\}_{n\in\mathbb{Z}}$ be a sequence of complex numbers, then the following conditions are equivalent:*

(a) $\exists\ \mu \in B^*$, $\|\mu\|_{B^*} \le C$ such that $\widehat{\mu}(n) = a_n$ for all $n \in \mathbb{Z}$.

(b) For all trigonometric polynomial P then

$$\left|\sum_{n\in\mathbb{Z}} \widehat{P}(n)\overline{a_n}\right| \le C\|P\|_B.$$

Proof. Assume (a) holds, then for a trigonometric polynomial P, we can assume

$$P(t) = \sum_{-m}^{m} c_j e^{ijt} \qquad \Longrightarrow \qquad \widehat{P}(n) = \begin{cases} c_n & \text{if} \quad -m \le n \le m \\ 0 & \text{elsewhere.} \end{cases}$$

Thus

$$\left|\sum_{n\in\mathbb{Z}} \widehat{P}(n)\overline{a_n}\right| = \left|\sum_{-m}^{m} \widehat{P}(n)\overline{a_n}\right| = \left|\langle P, \mu\rangle\right| \le C\|P\|_B$$

by Parseval's formula 1.19 we have

$$\langle P, \mu\rangle = \lim_{m\longrightarrow\infty} \sum_{-m}^{m}\left(1 - \frac{|n|}{m+1}\right)\widehat{P}(n)\overline{a_n} = \sum_{-m}^{m} \widehat{P}(n)\overline{a_n}.$$

Now assume (b)holds, we can define the linear bounded functional on the set of trigonometric polynomials in B by

$$\Lambda : P \longmapsto \sum_{n\in\mathbb{Z}} \widehat{P}(n)\overline{a_n}.$$

Since the set of all trigonometric polynomials is dense in B, Λ extends uniquely to $\Lambda \in B^*$, then clearly $\|\Lambda\|_{B^*} \le C$ and clearly

$$\widehat{\Lambda}(n) = \overline{\langle e^{int}, \Lambda\rangle} = \overline{\overline{a_n}} = a_n.$$

□

Corollary 1.23. *A trigonometric series $S \sim \sum_{n\in\mathbb{Z}} a_n e^{int}$ is the Fourier series of some $\mu \in B^*$, $\|\mu\|_{B^*} \le C$ if and only if $\|\sigma_m(S)\|_{B^*} \le C$ for all m, here $\sigma_m(S)$ denotes the element in B^* which has the Fourier series is*

$$\sum_{-m}^{m}\left(1 - \frac{|j|}{m+1}\right)a_j e^{ijt}.$$

Proof. If $\mu \in B^*$ with $\|\mu\|_{B^*} \le C$ has its Fourier series is $S[\mu] \sim \sum_{n\in\mathbb{Z}} a_n e^{int}$ then $\widehat{\mu}(n) = a_n$ for all $n \in \mathbb{Z}$, then from theorem 1.21 we have

$$\sigma_m(S) = \sigma_m(\mu) \qquad \Longrightarrow \qquad \|\sigma_m(S)\|_{B^*} = \|\sigma_m(\mu)\|_{B^*} = \|\Sigma_m^*(\mu)\|_{B^*} \le \|\Sigma_m^*\|_{B^{**}}\|\mu\|_{B^*} \le C.$$

Conversely, if $\|\sigma_m(S)\|_{B^*} \le C$ for all $m \in \mathbb{Z}$ then by Banach-Alaoglu theorem, there exists $\mu \in B^*$ such that $\sigma_m(S) \overset{*}{\rightharpoonup} \mu$ in the weak* topology of B^* as $m \longrightarrow \infty$ (upto sub-sequence). It is clear that

$\|\mu\|_{B^*} \leq C$ as well and by theorem 1.22 we have for all trigonometric polynomial P then

$$\begin{aligned}\langle P, \mu\rangle &= \lim_{m \longrightarrow \infty} \langle P, \sigma_m(S)\rangle \\ &= \lim_{m \longrightarrow \infty} \frac{1}{2\pi} \int_{\mathbb{T}} P(t)\overline{\sigma_m(S)}\, dt = \lim_{m \longrightarrow \infty} \sum_{-m}^{m} \left(1 - \frac{|j|}{m+1}\right) \widehat{P}(j)\overline{a_j} = \sum_{-\deg P}^{\deg P} \widehat{P}(j)\overline{a_j}.\end{aligned}$$

On the other hand the Parseval's formula reads

$$\langle P, \mu\rangle = \lim_{m \longrightarrow \infty} \sum_{-m}^{m} \left(1 - \frac{|j|}{m+1}\right) \widehat{P}(j)\overline{\widehat{\mu}(j)} = \sum_{-\deg P}^{\deg P} \widehat{P}(j)\overline{\widehat{\mu}(j)}.$$

This is true for all trigonometric polynomials P, hence the result follows $\widehat{\mu}(n) = a_n$ for all $n \in \mathbb{Z}$. □

6. In the case $B = C(\mathbb{T})$, the dual space B^* is identified with $M(\mathbb{T})$-the space of all Borel measures on $\mathbb{T}$ by mean of the coupling (Rieze's representation theorem)

$$\langle f, \mu\rangle = \int_{\mathbb{T}} f\, \overline{d\mu} \qquad \text{for all} \qquad f \in C(\mathbb{T}).$$

We shall refer to Fourier coefficients of measures as Fourier-Stieltjes coefficients and to Fourier series of measures as Fourier-Stieltjes series. The mapping

$$f \longmapsto \frac{1}{2\pi} f(t)\, dt \qquad \text{is an isometric embedding of} \qquad L^1(\mathbb{T}) \text{ in } M(\mathbb{T}).$$

Observe that if $\mu = \frac{1}{2\pi} f\, dt$ then

$$\widehat{\mu}(n) = \overline{\langle e^{int}, \mu\rangle} = \frac{1}{2\pi} \overline{\int_{\mathbb{T}} e^{int}\overline{f(t)}\, dt} = \frac{1}{2\pi} \int_{\mathbb{T}} f(t) e^{-int}\, dt = \widehat{f}(n).$$

7. A measure μ is positive if $\mu(E) \geq 0$ for every measurable set E, or equivalently $\int_{\mathbb{T}} f\, d\mu \geq 0$ for all non-negative $f \in C(\mathbb{T})$. If $\mu \ll m$, i.e $\mu = \frac{1}{2\pi} f(t)\, dt$ for some $f \in L^1(\mathbb{T})$, then μ is positive if and only if $f(t) \geq 0$ almost everywhere.

Theorem 1.24. *A series $S \sim \sum_{n\in\mathbb{Z}} a_n e^{int}$ is the Fourier-Stieltjes series of a positive measure if and only if for all $n \in \mathbb{Z}$ and $t \in \mathbb{T}$*

$$\sigma_n(S)(t) = \sum_{-n}^{n} \left(1 - \frac{|j|}{n+1}\right) a_j e^{ijt} \geq 0.$$

Proof. If there exists $\mu \in M(\mathbb{T})$ such that $S = S(\mu)$ and $\mu \geq 0$ then $\widehat{\mu}(n) = a_n$ for all $n \in \mathbb{Z}$ and if $f \in C(\mathbb{T})$ with $f \geq 0$ then

$$\langle f, \sigma_n(\mu)\rangle = \frac{1}{2\pi} \int_{\mathbb{T}} f(t)\overline{\sigma_n(\mu)(t)}\, dt = \sum_{-n}^{n} \left(1 - \frac{|j|}{n+1}\right) \widehat{f}(j)\overline{\widehat{\mu}(j)}$$

while $f \geq 0$ implies $\sigma_n(f) = \mathcal{K}_n * f = \frac{1}{2\pi} \int_{\mathbb{T}} \mathcal{K}_n(s) f(\cdot - s)\, ds \geq 0$ as well since $\mathcal{K}_n \geq 0$, thus

$$0 \leq \langle \sigma_n(f), \mu\rangle = \sum_{-n}^{n} \left(1 - \frac{|j|}{n+1}\right) \widehat{f}(j)\langle e^{ijt}, \mu\rangle = \sum_{-n}^{n} \left(1 - \frac{|j|}{n+1}\right) \widehat{f}(j)\overline{\widehat{\mu}(j)}$$

and thus (this can be explained also in term of adjoint operator)

$$\langle f, \sigma_n(\mu)\rangle = \langle \sigma_n(f), \mu\rangle \geq 0$$

for all non-negative $f \in C(\mathbb{T})$, hence $\sigma_n(\mu) \geq 0$ on $\mathbb{T}$. Conversely if $\sigma_n(S)(t) \geq 0$ on $\mathbb{T}$, as member of $M(\mathbb{T}) = C(\mathbb{T})^*$ we have

$$\|\sigma_n(S)\|_{M(\mathbb{T})} = \frac{1}{2\pi} \int_{\mathbb{T}} \sigma_n(S)(t)\, dt = a_0$$

for all $n \in \mathbb{Z}$, thus theorem 1.23 implies that there exists $\mu \in M(\mathbb{T})$ with $\|\mu\|_{M(\mathbb{T})} = a_0$ such that $\widehat{\mu}(n) = a_n$, i.e., $S = S(\mu)$, and clearly by weak$*$ convergence $\sigma_n(\mu) \overset{*}{\rightharpoonup} \mu$ in $M(\mathbb{T})$ we have

$$\langle f, \mu \rangle = \lim_{n \longrightarrow \infty} \langle f, \sigma_n(S) \rangle \geq 0$$

for any non-negative $f \in C(\mathbb{T})$. □

The condition $\sigma_n(S)(t) \geq 0$ for all $n \in \mathbb{Z}$ can be replaced by $\sigma_n(S)(t) \geq 0$ for infinitely many n's.

8. (**Characterization Fourier-Stieltjes coefficients of positive measures as positive definite sequences**) A numerical sequence $\{a_n\}_{n\in\mathbb{Z}}$ is *"positive definite"* if for any sequence $\{z_n\}$ of complex numbers having only a finite number of non-zero terms we have

$$\sum_{n,m} a_{n-m} z_n \overline{z_m} \geq 0.$$

It is obvious that with the sequence $z_0 = 1$ and $z_n = 0$ elsewhere we obtain $a_0 \geq 0$.

Theorem 1.25 (Herglotz). *A numerical sequence $\{a_n\}_{n\in\mathbb{Z}}$ is positive definite if and only if there exists a positive measure $\mu \in M(\mathbb{T})$ such that $\widehat{\mu}(n) = a_n$ for all $n \in \mathbb{Z}$.*

Proof. If $a_n = \widehat{\mu}(n)$ for some positive $\mu \in M(\mathbb{T})$ then

$$a_n = \widehat{\mu}(n) = \overline{\langle e^{int}, \mu \rangle} = \int_{\mathbb{T}} e^{-int} \, \overline{d\mu}$$

and hence for such a sequence $\{z_n\}$ only has finitely many non-zero terms we have

$$\sum_{m,n} a_{n-m} z_n \overline{z_m} = \sum_{m,n} \int_{\mathbb{T}} e^{-int} e^{imt} z_n \overline{z_m} \, \overline{d\mu} = \int_{\mathbb{T}} \left| \sum_n e^{-int} z_n \right|^2 d\mu \geq 0.$$

Conversely, if $\{a_n\}_{n\in\mathbb{Z}}$ is a positive definite sequence, we write $S \sim \sum_{n\in\mathbb{Z}} a_n e^{int}$. For any $N \in \mathbb{Z}$ and $t \in \mathbb{T}$ we define

$$z_n = \begin{cases} e^{int} & \text{if } |n| \leq N, \\ 0 & \text{if } |n| > N. \end{cases}$$

Then we have

$$\sum_{m,n\in\mathbb{Z}} a_{m-n} z_n \overline{z_m} = \sum_{m=-N}^{N} \sum_{n=-N}^{N} a_{m-n} e^{i(n-m)t}.$$

Let $k = n - m$ and re-write the formula above in terms of sum in k, we obtain

$$0 \leq \sum_{m,n\in\mathbb{Z}} a_{m-n} z_n \overline{z_m} = \sum_{k=-2N}^{2N} (2N+1-|k|) a_k e^{ikt} = (2N+1) \sum_{k=-2N}^{2N} \left(1 - \frac{|k|}{2N+1}\right) a_k e^{ikt} = (2N+1)\sigma_{2N}(S)(t).$$

It is true for all $t \in \mathbb{T}$ and for infinitely many n's, thus theorem 1.24 concludes that there exists a positive measure $\mu \in M(\mathbb{T})$ such that $S = S(\mu)$. □

Theorem 1.26. *If $\{a_n\}$ is positive definite then $|a_n| \leq a_0$, $a_{-N} = \overline{a_N}$ and $\left\{a_n - \frac{a_{n-1}+a_{n+1}}{2}\right\}$ is positive definite.*

Proof. Take $z_0 = 1, z_N = z$ and $z_n = 0$ elsewhere, we have

$$\sum_{m,n} a_{n-m} z_n \overline{z_m} = a_0 \left(1 + |z|^2\right) + a_N z + a_{-N} \overline{z} \geq 0 \qquad \text{for all} \qquad z \in \mathbb{C}.$$

Set $z = 1$ we have $2a_0 + a_N + a_{-N} \geq 0$, thus $a_N + a_{-N} \in \mathbb{R}$. Set $z = i$ we have $i(a_N - a_{-N}) \geq 0$, which means $a_{-N} = \overline{a_N}$. Thus we can take $z \in \mathbb{C}$ such that

$$za_N = -|a_N| \qquad \Longrightarrow \qquad 2a_0 - 2|a_N| \geq 0 \qquad \Longrightarrow \qquad |a_N| \leq 0 \qquad \text{for all } N \in \mathbb{Z}.$$

This fact can be obtained obviously from Herglotz theorem 1.25, since with the positive measure μ satisfies $\widehat{\mu}(n) = a_n$ then

$$|a_n| = |\widehat{\mu}(n)| \leq \int_{\mathbb{T}} |\overline{d\mu}| = \mu(\mathbb{T}) = \mu(0) = a_0.$$

Finally let $d\mu_1 = e^{it}\, d\mu$ and $d\mu_{-1} = e^{-it}\, d\mu$ be measures in $M(\mathbb{T})$, we then have $\widehat{\mu_1}(n) = \widehat{\mu}(n+1) = a_{n+1}$ and $\widehat{\mu_{-1}}(n) = \widehat{\mu}(n-1) = a_{n-1}$, let

$$\nu = \mu - \frac{\mu_1 + \mu_{-1}}{2} \qquad \Longrightarrow \qquad d\nu = \left(1 - \frac{e^{it} + e^{-it}}{2}\right) d\mu = (1 - \cos t)\, d\mu \geq 0$$

and clearly $\widehat{\nu}(n) = a_n - \frac{a_{n-1} + a_{n+1}}{2}$. As a consequence, we have

$$|\widehat{\nu}(n)| = \left|a_n - \frac{a_{n-1} + a_{n+1}}{2}\right| \leq |\widehat{\nu}(0)| = \left|a_0 - \frac{a_1 + a_{-1}}{2}\right| = a_0 - \mathrm{Re}(a_1).$$

□

9. (**Universal multipliers - convolution**)

Theorem 1.27 (Universal multipliers)**.** *Let B be a homogeneous Banach space on $\mathbb{T}$ (contains e^{int}) and $\mu \in M(\mathbb{T})$. There exists a unique linear operator Λ on B having the properties:*

(i) $\|\Lambda\|_{L(B,B)} \leq \|\mu\|_{M(\mathbb{T})}$.

(ii) $\widehat{\Lambda f}(n) = \widehat{\mu}(n)\widehat{f}(n)$ *for all* $f \in B$.

Proof. If an operator $\Lambda \in L(B,B)$ satisfies (i) and (ii) then for any trigonometric polynomial

$$P(t) = \sum_{n=-m}^{m} \widehat{P}(n)e^{int}, \qquad \text{i.e.,} \qquad \widehat{P}(n) = 0 \qquad \text{for} \qquad |n| > m$$

the corresponding Fourier series of the element $\Lambda P \in B$ is

$$S[\Lambda P] \sim \sum_{-\infty}^{+\infty} \widehat{\Lambda P}(n)e^{int} = \sum_{-m}^{m} \widehat{\mu}(n)\widehat{P}(n)e^{int}.$$

As $B \subset L^1(\mathbb{T})$, the uniqueness theorem 1.8 concludes that the action of Λ is uniquely determined on the set of trigonometric polynomials on B, and hence on B since (i) and the fact that the set of trigonometric polynomials is dense in B. For the existence of Λ, let's define

$$\Lambda P(t) = \sum_{-m}^{m} \widehat{\mu}(n)\widehat{P}(n)e^{int} \qquad \text{for any trigonometric polynomial } P \text{ in } B.$$

It is clear that Λ defines a linear operator on the set of trigonometric polynomials on B, we have left to show that $\|\Lambda\|_{B^*} \leq \|\mu\|_{M(\mathbb{T})}$. We observe that if $\mu \ll m$, i.e., $d\mu = \frac{1}{2\pi}g(t)\, dt$ for some $g \in C(\mathbb{T})$ then since $\widehat{\mu}(n) = \widehat{g}(n)$ for all $n \in \mathbb{Z}$, we obtain

$$\Lambda P(t) = \sum_{-m}^{m} \widehat{g}(n)\widehat{P}(n)e^{int} = \sum_{-m}^{m} \left(\widehat{g * P}\right)(n)e^{-int} = g * P(t)$$

and by lemma 1.7 we have $g * P$ can be seen as the B-valued integral

$$g * P = \frac{1}{2\pi}\int_{\mathbb{T}} g(s)P_s(\cdot)\, ds \qquad \Longrightarrow \qquad \|\Lambda P\|_B = \|g * P\|_B \leq \left(\frac{1}{2\pi}\int_{\mathbb{T}} |g(s)|\, ds\right)\|P\|_B = \|\mu\|_{M(\mathbb{T})} \cdot \|P\|_B$$

and thus $\|\Lambda\|_{B^*} \le \|\mu\|_{M(\mathbb{T})}$. In the general case, recall that

$$d\sigma_n(\mu) = \frac{1}{2}g_n(t)\,dt \qquad \text{where} \qquad g_n(t) = \sum_{-n}^{n}\left(1 - \frac{|j|}{n+1}\right)\widehat{\mu}(j)e^{ijt} \in C(\mathbb{T})$$

satisfies $\sigma_n(\mu) \overset{*}{\rightharpoonup} \mu$ in the weak* topology of $M(\mathbb{T})$, also $\|\sigma_n(\mu)\|_{M(\mathbb{T})} \le \|\mu\|_{M(\mathbb{T})}$ by theorem 1.21. Finally by theorem 1.4 we have

$$(g_n * P)(t) = \sum_{-n}^{n}\left(1 - \frac{|j|}{n+1}\right)\widehat{\mu}(j)\widehat{P}(j)e^{ijt} = \sum_{-m}^{m}\left(1 - \frac{|j|}{n+1}\right)\widehat{\mu}(j)\widehat{P}(j)e^{ijt}$$

if we choose $n > |m| = \deg P$. Thus

$$\|g_n * P - \Lambda P\|_B = \left\|\sum_{-m}^{m}\left(\frac{|j|}{n+1}\right)\widehat{\mu}(j)\widehat{P}(j)e^{int}\right\|_B = \frac{1}{n+1}\left\|\sum_{-m}^{m}|j|\widehat{\mu}(j)\widehat{P}(j)e^{int}\right\|_B \longrightarrow 0$$

as $n \longrightarrow \infty$. Thus since $\|g_n * P\|_B \le \|\mu\|_{M(\mathbb{T})}$ for all $n \in \mathbb{N}$ we obtain $\|\Lambda P\|_B \le \|\mu\|_{M(\mathbb{T})}$, and the extension to all $f \in B$ is obvious since the set of trigonometric polynomials in B is dense in B. □

Corollary 1.28. *Let $f \in B$ and $\mu \in M(\mathbb{T})$, then $\left\{\widehat{\mu}(n)\widehat{f}(n)\right\}$ is the sequence of Fourier coefficients of some function in B.*

In view of these above result, we shall write $\mu * f$ instead of Λf, and refer to it as the convolution of μ and f.

10. (Convolution of a measures and a linear functional using Fourier series) For $\mu \in M(\mathbb{T})$ we define $\mu^{\#} \in M(\mathbb{T})$ by

$$\mu^{\#}(E) = \overline{\mu(-E)} \quad \text{for all Borel sets } E, \quad \text{or equivalently, by} \quad \int_{\mathbb{T}} f(t)\,d\mu^{\#} = \int_{\mathbb{T}} f(-t)\,\overline{d\mu} \quad \text{for all } f \in C(\mathbb{T}).$$

It is clear that

$$\widehat{\mu^{\#}}(n) = \overline{\widehat{\mu}(n)} \qquad \text{for} \qquad n \in \mathbb{Z}.$$

Let $\Lambda \in L(B,B)$ be the operator which maps $f \longmapsto \Lambda f = \mu * f$ previously. If $\Lambda^* \in L(B^*,B^*)$ be the adjoint of Λ, then for any $\nu \in B^*$ and $f \in B$ by Parseval's formula we have

$$\lim_{n\longrightarrow\infty}\sum_{-n}^{n}\left(1 - \frac{|j|}{n+1}\right)\widehat{f}(j)\widehat{\mu}(j)\overline{\widehat{\nu}(j)} = \langle \Lambda f, \nu\rangle = \langle f, \Lambda^*\nu\rangle = \lim_{n\longrightarrow\infty}\sum_{-n}^{n}\left(1 - \frac{|j|}{n+1}\right)\widehat{f}(j)\overline{\widehat{\Lambda^*\nu}(j)}.$$

Since it is true for all $f \in B$, by testing with trigonometric polynomials in B we deduce that

$$\widehat{\Lambda^*\nu}(n) = \overline{\widehat{\mu}(n)}\widehat{\nu}(n) = \widehat{\mu^{\#}}(n)\widehat{\nu}(n) \qquad \text{for all} \qquad n \in \mathbb{Z}.$$

In other words, $\Lambda^*\nu$ is the element of B^* which has its Fourier series is

$$S[\Lambda^*\nu] \sim \sum_{n\in\mathbb{Z}}\widehat{\mu^{\#}}(n)\widehat{\nu}(n)e^{int}.$$

We denote this element as $\Lambda^*\nu = \mu^{\#} * \nu$. We have proved the following theorem.

Theorem 1.29. *Let B be a homogeneous Banach space on $\mathbb{T}$ (contains e^{int}) and B^* its dual. If $\mu \in M(\mathbb{T})$ and $\nu \in B^*$, then there exists a unique element in B^*, denoted by $\mu * \nu$ which has its Fourier series is*

$$S[\mu * \nu] \sim \sum_{n\in\mathbb{Z}}\widehat{\mu}(n)\widehat{\nu}(n)e^{int}.$$

Moreover, $\|\mu * \nu\|_B \le \|\mu\|_{M(\mathbb{T})}\|\nu\|_B$.

In particular, for two measures $\mu, \nu \in M(\mathbb{T})$ there exists a unique measure $\mu * \nu \in M(\mathbb{T})$ which has its Fourier series is (Fourier-Stieltjes series of the measure)

$$S[\mu * \nu] \sim \sum_{n\in\mathbb{Z}} \widehat{\mu}(n)\widehat{\nu}(n)e^{int}.$$

11. Of course we can define the convolution of two measures $\mu, \nu \in M(\mathbb{T})$ in a direct way. For any $f \in C(\mathbb{T})$, the integral

$$I(f) = \iint_{\mathbb{T}^2} f(t+s)\, d\mu(t) d\nu(s)$$

is well-defined and $f \longmapsto I(f)$ defines a bounded linear functional on $C(\mathbb{T})$ since $|I(f)| \le \|\mu\|_{M(\mathbb{T})}\|\nu\|_{M(\mathbb{T})}$. By Rieze's representation theorem there exists a unique measure $\lambda \in M(\mathbb{T})$ such that

$$I(f) = \int_{\mathbb{T}} f(t)\, d\lambda(t) = \int_{\mathbb{T}^2} f(t+s)\, d\mu(t)\, d\nu(s) \qquad \text{for all} \qquad f \in C(\mathbb{T}).$$

By taking $f(t) = e^{int}$ we obtain $\widehat{\lambda}(n) = \widehat{\mu}(n)\widehat{\nu}(n)$ for $n \in \mathbb{Z}$, thus $\lambda = \mu * \nu$. In other words,

$$\int_{\mathbb{T}} f\, d(\mu * \nu) = \int_{\mathbb{T}^2} f(t+s)\, d\mu(t)\, d\nu(s) \qquad \text{for all} \qquad f \in C(\mathbb{T})$$

or by taking a sequence of continuous functions which converges to χ_E for a closed set E, we have

$$(\mu * \nu)(E) = \int_{\mathbb{T}} \mu(E-s)\, d\nu(s) \qquad \text{for all Borel set } E.$$

By regularity it is true for all Borel set E.

12. A measure $\mu \in M(\mathbb{T})$ is discrete if $\mu = \sum_{j=1}^n a_j \delta_{s_j}$ where $\{a_j\}$ are complex numbers.

Lemma 1.30. *If* $\nu = \sum_{j=1}^n a_j \delta_{s_j}$ *then* $\|\nu\|_{M(\mathbb{T})} = \sum_{j=1}^n |a_j|$.

Proof. First of all, if $\nu = a\delta_0$ for $a \in \mathbb{C}\backslash\{0\}$ then with $\mu = |a|\delta_0$ we have $\nu \ll \mu$, thus by Radon-Nikodym theorem

$$d\nu = \frac{d\nu}{d\mu}\, d\mu \quad\Longrightarrow\quad \nu(E) = \int_E \frac{d\nu}{d\mu}\, d\mu = \begin{cases} a & \text{if} \quad 0 \in E, \\ 0 & \text{if} \quad 0 \notin E \end{cases} \quad\Longrightarrow\quad \int_E \left(\frac{d\nu}{d\mu} - \frac{a}{|a|}\right) d\mu = 0$$

for all Borel set E, which implies $\frac{d\nu}{d\mu} = \frac{a}{|a|}$ μ-a.e. and hence the total variation of ν is, by definition

$$d|\nu| = \left|\frac{a}{|a|}\right| d\mu = |a|\delta_0 \qquad\Longrightarrow\qquad \|a\delta_0\|_{M(\mathbb{T})} = |a|.$$

For the general case, we can assume $\{s_j\}_{j=1}^n$ are disjoint. Let $\mu = \sum_{j=1}^n |a_j|\delta_{s_j}$ then clearly $\nu \ll \mu$, so

$$d\nu = \frac{d\nu}{d\mu}\, d\mu \quad\Longrightarrow\quad \nu(E) = \int_E \frac{d\nu}{d\mu} = \int_E \left(\sum_{j=1}^n \frac{a_j}{|a_j|}\chi_E(a_j)\right) d\mu \quad\Longrightarrow\quad \frac{d\nu}{d\mu} = \sum_{j=1}^n \frac{a_j}{|a_j|}\chi_{\{a_j\}} \qquad \mu-\text{a.e.}$$

Thus

$$d|\nu| = \left|\sum_{j=1}^n \frac{a_j}{|a_j|}\chi_{\{a_j\}}\right| d\mu \qquad\Longrightarrow\qquad \|\nu\|_{M(\mathbb{T})} = \sum_{j=1}^n |a_j|.$$

□

A measure $\mu \in M(\mathbb{T})$ is continuous if $\mu(\{t\}) = 0$ for every $t \in \mathbb{T}$, equivalently μ is continuous if

$$\lim_{\eta\longrightarrow 0} \int_{t-\eta}^{t+\eta} d|\mu| = 0 \qquad \text{for every } t \in \mathbb{T}.$$

Theorem 1.31. *Every measure $\mu \in M(\mathbb{T})$ can be decomposed to a sum $\mu = \mu_c + \mu_d$ where μ_c is continuous and μ_d is discrete.*

If $\mu \in M(\mathbb{T})$ is a continuous measure then for any $\nu \in M(\mathbb{T})$, from the formula

$$(\mu * \nu)(E) = \iint_{\mathbb{T}^2} \mu(E - s)\, d\nu(s)$$

we deduce that $\mu * \nu$ is continuous. Since $\delta_a * \delta_b = \delta_{a+b}$, if $\mu = \sum_{j=1}^{n} a_j \delta_{s_j}$ and $\nu = \sum_{j=1}^{n} b_j \delta_{t_j}$ then

$$\mu * \nu = \sum_{j,k=1}^{n} a_j b_k \delta_{s_j + t_k}.$$

Let $\mu = \mu_c + \mu_d$ and $\mu^\# = \mu_c^\# + \mu_d^\#$ be the decompositions to continuous and discrete parts, we have

$$\mu * \mu^\# = \underbrace{\left(\mu_c * \mu_c^\# + \mu_c * \mu_d^\# + \mu_d * \mu_c^\#\right)}_{(\mu*\mu^\#)_c} + \left(\mu_d * \mu_d^\#\right).$$

Assume that $\mu_d = \sum_{j=1}^{n} a_j \delta_j$, then $\mu_d^\# = \sum_{j=1}^{n} \overline{a_j} \delta_{-s_j}$ and thus $\mu * \mu^\#(\{0\}) = \sum_{j=1}^{n} |a_j|^2$.

Lemma 1.32. *Let $\mu \in M(\mathbb{T})$, then*

$$\mu * \mu^\# \{0\} = \sum_{t \in \mathbb{T}} |\mu(\{t\})|^2.$$

In particular, μ is continuous if and only if $(\mu * \mu^\#)(\{0\}) = 0$.

The discrete part of a measure $\mu \in M(\mathbb{T})$ can be recovered from its Fourier-Stieltjes series.

Theorem 1.33. *Let $\mu \in M(\mathbb{T})$ and $t \in \mathbb{T}$, then*

$$\mu(\{t\}) = \lim_{m \longrightarrow \infty} \frac{1}{2m+1} \sum_{-m}^{m} \widehat{\mu}(n) e^{int}.$$

Proof. For $t \in \mathbb{T}$, the function

$$\varphi_m(s) = \frac{1}{2m+1} \mathcal{D}_m(t - s) = \frac{1}{2m+1} \sum_{-m}^{m} e^{-ins} e^{int}$$

is bounded by 1 and tends to zero uniformly outside any neighborhood of t. Now the measure

$$\nu = \mu - \mu(\{t\})\delta_t$$

satisfies $\nu(\{t\}) = 0$. Let's recall that the total variance of complex measures can be computed by

$$|\nu|(E) = \sup \left\{ \sum_{j=1}^{n} |\nu(E_i)| : n \in \mathbb{N}, E_1, \dots, E_n \text{ disjoint}, E = \bigcup_{j=1}^{n} E_j \right\}.$$

From that we obtain $|\nu|(\{t\}) = 0$. Thus by dominated convergence theorem we have

$$\lim_{\eta \longrightarrow 0} \int_{t-\eta}^{t+\eta} d|\nu| = \lim_{\eta \longrightarrow 0} \int_{\mathbb{T}} \chi_{(t-\eta, t+\eta)}\, d|\nu| = |\nu|(\{t\}) = 0.$$

Thus we have

$$\langle \varphi_m, \mu - \mu(\{t\})\delta_t \rangle = \langle \varphi_m, \nu \rangle = \int_{\mathbb{T}} \varphi_m\, \overline{d\nu} = \int_{(t-\eta, t+\eta)} \varphi_m\, \overline{d\nu} + \int_{\mathbb{T} \setminus (t-\eta, t+\eta)} \varphi_m\, \overline{d\nu} \longrightarrow 0$$

as $m \longrightarrow \infty$ since φ_m is bounded by 1 and converges to 0 uniformly away from t. Since

$$\begin{aligned}
\langle \varphi_m, \mu - \mu(\{t\})\delta_t \rangle = \langle \varphi_m, \mu \rangle - \int_{\mathbb{T}} \varphi_m \overline{\mu(\{t\})}\, d\delta_t \\
= \frac{1}{2m+1} \sum_{-m}^{m} \langle e^{-ins}, \mu \rangle e^{int} - \overline{\mu(\{t\})} \\
= \frac{1}{2m+1} \sum_{-m}^{m} \overline{\widehat{\mu}(-n)} e^{int} - \overline{\mu(\{t\})} = \frac{1}{2m+1} \sum_{-m}^{m} \overline{\widehat{\mu}(n)} e^{-int} - \overline{\mu(\{t\})}
\end{aligned}$$

we have

$$\overline{\langle \varphi_m, \mu - \mu(\{t\})\delta_t \rangle} = \frac{1}{2m+1} \sum_{-m}^{m} \widehat{\mu}(n) e^{int} - \mu(\{t\})$$

and the result follows from the fact that $\langle \varphi_m, \mu - \mu(\{t\}) \rangle \longrightarrow 0$ as $m \longrightarrow \infty$. ☐

Corollary 1.34 (Wiener)**.** *Let $\mu \in M(\mathbb{T})$ then*

$$\sum_{t \in \mathbb{T}} |\mu(\{t\})|^2 = \lim_{m \longrightarrow \infty} \frac{1}{2m+1} \sum_{-m}^{m} |\widehat{\mu}(n)|^2.$$

In particular, μ is continuous if and only if

$$\lim_{m \longrightarrow \infty} \frac{1}{2m+1} \sum_{-m}^{m} |\widehat{\mu}(n)|^2 = 0.$$

Proof. Apply theorem 1.33 to $\mu * \mu^{\#}$ at $t = 0$ we obtain the result. ☐

2 Fourier transform

- We denote $L^1(\mathbb{R})$ means $L^1(\mathbb{R}, m)$ where m is the Lebesgue measure on $\mathbb{R}$, $\mathcal{S}(\mathbb{R})$ the space of Schwartz functions on $\mathbb{R}$ and $C_0(\mathbb{R})$ the set of functions f that vanishes at infinity, i.e., $\{x : |f(x)| \geq \varepsilon\}$ is compact for all $\varepsilon > 0$.
- The Fourier transform follows the following convention

$$\mathcal{F}f(\xi) = \widehat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-i\xi x}\, dx \qquad \text{for} \qquad \xi \in \widehat{\mathbb{R}}$$

where $\widehat{\mathbb{R}}$ is the space of frequencies, another copy of $\mathbb{R}$. The inversion formula is (under some mild conditions)

$$f(x) = \mathcal{F}^{-1}[\widehat{f}] = \frac{1}{2\pi} \int_{\widehat{\mathbb{R}}} \widehat{f}(\xi) e^{i\xi x}\, d\xi \qquad \text{for} \qquad x \in \mathbb{R}.$$

- The Fejer's kernel is $\{\mathcal{K}_\lambda : \lambda > 0\}$ where $\mathcal{K}_\lambda(x) = \lambda \mathcal{K}(\lambda x)$ with

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

We have $\|\mathcal{K}_\lambda\|_{L^1(\mathbb{R})} = 1$ for all $\lambda > 0$, it is a summability kernel with $\mathcal{K}_\lambda \longrightarrow \delta_0$ as $\lambda \longrightarrow \infty$. Its Fourier transform is

$$\widehat{\mathcal{K}}_\lambda(\xi) = \left(1 - \frac{|\xi|}{\lambda}\right) \chi_{[-\lambda,\lambda]}(\xi).$$

The Fejer's kernel does not belong to $\mathcal{S}(\mathbb{R})$, but is is infinitely differentiable. As tempered distributions, $\lambda^{-1}\mathcal{K}_\lambda \longrightarrow \frac{1}{2\pi}$ in $\mathcal{S}'(\mathbb{R})$ as $\lambda \longrightarrow 0$, thus $\lambda^{-1}\widehat{\mathcal{K}}_\lambda \longrightarrow \delta_0$ in $\mathcal{S}'(\mathbb{R})$ as $\lambda \longrightarrow 0$. It is indeed true that $\lambda^{-1}\mathcal{K}_\lambda \longrightarrow \frac{1}{2\pi}$ uniformly on compact sets of $\mathbb{R}$.

- For $f \in L^1(\mathbb{R})$ and $m \in \mathbb{Z}$ we have

$$\int_{\mathbb{T}} \sum_{-m}^{m} |f(t-2\pi n)|\, dt = \sum_{-m}^{m} \int_{\mathbb{T}+2\pi n} |f(s)|\, ds \qquad \Longrightarrow \qquad \int_{\mathbb{T}} \sum_{-\infty}^{\infty} |f(t-2\pi n)|\, dt = \|f\|_{L^1(\mathbb{R})}$$

by monotone convergence theorem. Thus the series $\sum_{-\infty}^{\infty} f(t-2\pi n)$ is finite a.e. in $\mathbb{T}$, which implies it converges absolutely for a.e. $t \in \mathbb{T}$, hence the function

$$\varphi(t) := 2\pi \sum_{n\in\mathbb{Z}} f(t-2\pi n) \qquad \left(\|\varphi\|_{L^1(\mathbb{T})} = \frac{1}{2\pi}\int_{\mathbb{T}} |\varphi(t)|\, dt\right)$$

is well-defined as a (2π-periodic) function in $L^1(\mathbb{T})$ with $\|\varphi\|_{L^1(\mathbb{T})} \le \|f\|_{L^1(\mathbb{R})}$. For $n \in \mathbb{Z}$ then

$$\begin{aligned}\widehat{\varphi}(n) &= \frac{1}{2\pi}\int_{\mathbb{T}} \varphi(t) e^{-int}\, dt \\ &= \sum_{j=-\infty}^{\infty} \int_{\mathbb{T}} f(t-2\pi j) e^{-int}\, dt = \sum_{j=-\infty}^{\infty} \int_{\mathbb{T}+2\pi j} f(s) e^{-ins}\, dt = \int_{\mathbb{R}} f(s) e^{-ins}\, ds = \widehat{f}(n).\end{aligned}$$

If we denote $f_\lambda(x) = \lambda f(\lambda x)$, and $\varphi_\lambda(t) = 2\pi \sum_{n\in\mathbb{Z}} f_\lambda(t-2\pi n)$ then similarly $\widehat{\varphi_\lambda}(n) = \widehat{f_\lambda}(n) = \widehat{f}\left(\frac{n}{\lambda}\right)$.

2.1 Fourier-Stieltjes transforms

We denote by $M(\mathbb{R})$ the space of all finite Borel measures on $\mathbb{R}$, it is a normed space with the total mass norm on $M(\mathbb{R})$ is defined by $\|\mu\|_{M(\mathbb{R})} = \int_{\mathbb{R}} 1\, d|\mu| = |\mu|(\mathbb{R})$. Recall that $\left(M(\mathbb{R}), \|\cdot\|_{M(\mathbb{R})}\right)$ is identified with the dual space of $C_0(\mathbb{R})$ by means of the coupling ($\|\cdot\|_{M(\mathbb{R})}$ is identified with the dual norm)

$$\langle f, \mu\rangle_{M(\mathbb{R})} = \int_{\mathbb{R}} f\, \overline{d\mu}, \qquad f \in C_0(\mathbb{R}), \qquad \mu \in M(\mathbb{R}).$$

It is clear that the above formula defines μ as a linear functional on a larger space $\mathrm{BC}(\mathbb{R})$. The weak* topology on $M(\mathbb{R})$ is called the "vague topology", which is defined by $\mu_n \overset{*}{\rightharpoonup} \mu$ in $M(\mathbb{R})$ iff $\langle f, \mu_n\rangle \longrightarrow \langle f, \mu\rangle$ as $n \longrightarrow \infty$ for all $f \in C_0(\mathbb{R})$ (we suppress the subscript $M(\mathbb{R})$ in the product).

1. The mapping $f \longmapsto f\, dm$ identifies $L^1(\mathbb{R}, m)$ with a closed subspace of $\left(M(\mathbb{R}), \|\cdot\|_{M(\mathbb{R})}\right)$, since if $f_n \longrightarrow f$ in $L^1(\mathbb{R}, m)$ then for $\mu_n = f_n\, dm$ and $\mu = f\, dm$ we have

$$\|\mu_n - \mu\|_{M(\mathbb{R})} = \int_{\mathbb{R}} 1\, d|\mu_n - \mu| = \int_{\mathbb{R}} |f_n - f|\, dm = \|f_n - f\|_{L^1} \longrightarrow 0 \qquad \text{as} \qquad n \longrightarrow \infty.$$

2. The convolution of a measure $\mu \in M(\mathbb{R})$ and a function $\varphi \in C_0(\mathbb{R}, \mathbb{C})$ is a function defined by

$$(\mu * \varphi)(x) = \int_{\mathbb{R}} \varphi(x-y)\, d\mu(y).$$

Since $|\mu|(\mathbb{R}) < \infty$, it is clear that $\|\mu * \varphi\|_u \le \|\mu\|_{M(\mathbb{R})} \cdot \|\varphi\|_u$ thus the formula above is well-defined.

Lemma 2.1. *$\mu * \varphi \in C_0(\mathbb{R})$ for all $\mu \in M(\mathbb{R})$ and $\varphi \in C_0(\mathbb{R})$.*

Proof. The uniform continuity of $\varphi \in C_0(\mathbb{R}) \subset C_c(\mathbb{R})$ implies that $\mu * \varphi$ is continuous. For $\varepsilon > 0$, let's define the compact set A_ε to be

$$A_\varepsilon = \left\{ z \in \mathbb{R} : |\varphi(z)| \ge \frac{\varepsilon}{2|\mu|(\mathbb{R})} \right\}.$$

Since A_ε is compact, there exists $n \in \mathbb{N}$ such that $A_\varepsilon \subset [-n, n]$. Also it is clear that $\varphi \in L^1(|\mu|)$, hence there exists $m \in \mathbb{N}$ such that

$$\int_{\mathbb{R}\setminus[-m,m]} |\varphi(z)|\, d|\mu(z)| < \frac{\varepsilon}{2}.$$

For $|x| > m + n$ then $(x - A_\varepsilon) \cap [-m, m] = \emptyset$, thus

$$\int_{\mathbb{R}} |\varphi(x-y)|\, d|\mu(y)| = \int_{x-A_\varepsilon} |\varphi(x-y)|\, |d\mu(y)| + \int_{\mathbb{R}\setminus(x-A_\varepsilon)} |\varphi(x-y)|\, |d\mu(y)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2|\mu|(\mathbb{R})}|\mu|(\mathbb{R}) = \varepsilon.$$

Hence $\{x \in \mathbb{R} : |\mu * \varphi(x)| \geq \varepsilon\} \subset [-(m+n), m+n]$ which is compact since it is closed already. □

3. The convolution of two measures $\mu, \nu \in M(\mathbb{R})$ is another measure defined by

$$\begin{aligned}(\mu * \nu)(E) &= \int_{\mathbb{R}} \mu(E-y)\, d\nu(y) \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} \chi_{E-y}(x)\, d\mu(x)\, d\nu(y) = \int_{\mathbb{R}}\int_{\mathbb{R}} \chi_E(x+y)\, d\mu(x)\, d\nu(y).\end{aligned}$$

for every Borel set $E \subset \mathbb{R}$. It is clear that $\mu * \nu \in M(\mathbb{R})$ and $\|\mu * \nu\|_{M(\mathbb{R})} \leq \|\mu\|_{M(\mathbb{R})} \cdot \|\nu\|_{M(\mathbb{R})}$. We can generalize the above formula to the following.

Lemma 2.2. *For any bounded Borel measurable function $h : \mathbb{R} \longrightarrow \mathbb{R}$, we have*

$$\int_{\mathbb{R}} h(x)\, d(\mu * \nu)(x) = \int_{\mathbb{R}}\int_{\mathbb{R}} h(y+z)\, d\mu(y)\, d\nu(z).$$

Proof. First of all the claim is true for all characteristic function χ_E where E is a Borel set, thus by linearity it is true for all (Borel) simple functions. Write $h = h^+ - h^-$ where h^+, h^- are non-negative bounded Borel measurable functions, the finite properties of μ, ν an $\mu * \nu$ implies that they $h^\pm$ is $(\mu*\nu)$-integrable, and $(y,z) \longmapsto h(y+z)$ is $\mu \otimes \nu$-integrable as well (under the product measure). The general case follows by these observations and the monotone convergence theorem. □

4. We define the Fourier-Stieltjes transform of a measure $\mu \in M(\mathbb{R})$ to be a function by

$$\boxed{\mathcal{F}\mu(\xi) = \widehat{\mu}(\xi) = \overline{\left\langle e^{i\xi(\cdot)}, \mu \right\rangle}_{M(\mathbb{R})} = \int_{\mathbb{R}} e^{-i\xi x}\, d\mu(x) \qquad \text{for all} \qquad \xi \in \widehat{\mathbb{R}}.}$$

If μ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$, say $\mu = f(x)\, dx$ for some $f \in L^1(\mathbb{R}, m)$, then clearly $\widehat{\mu}(\xi) = \widehat{f}(\xi)$. Many properties of L^1-Fourier transforms are shared by Fourier-Stieltjes transforms:

(a) If $\mu \in M(\mathbb{R})$ then clearly $|\widehat{\mu}(\xi)| \leq \|\mu\|_{M(\mathbb{R})}$.

(b) $\mathcal{F} : M(\mathbb{R}) \longrightarrow \mathrm{BUC}(\widehat{\mathbb{R}})$. Indeed, for $\xi, \eta \in \widehat{\mathbb{R}}$, we have

$$\left|\widehat{\mu}(\xi+\eta) - \widehat{\mu}(\xi)\right| = \left|\int_{\mathbb{R}} e^{-i\xi x}\left(e^{-i\eta x} - 1\right)\, d|\mu(x)|\right| \leq \int_{\mathbb{R}} |e^{-i\eta x} - 1|\, d|\mu(x)|.$$

The integral on the right hand side is independent of ξ, and $|e^{-i\eta x} - 1| \leq 2 \in L^1(\mathbb{R}, |\mu(x)|)$, thus dominated convergence theorem can be applied to deduce that:

$$\lim_{\eta\longrightarrow 0}\left(\sup_{\xi\in\mathbb{R}}\left|\widehat{\mu}(\xi+\eta) - \widehat{\mu}(\xi)\right|\right) \leq \lim_{\eta\longrightarrow 0}\int_{\mathbb{R}} |e^{-i\eta x} - 1|\, d|\mu(x)| = \int_{\mathbb{R}}\left(\lim_{\eta\longrightarrow 0}|e^{-i\eta x} - 1|\right)\, d|\mu(x)| = 0.$$

(c) For $\mu, \nu \in M(\mathbb{R})$ then $\widehat{\mu * \nu}(\xi) = \widehat{\mu}(\xi)\widehat{\nu}(\xi)$ for any $\xi \in \widehat{\mathbb{R}}$. It follows by lemma 2.2 by

$$\begin{aligned}\widehat{\mu * \nu}(\xi) &= \int_{\mathbb{R}} e^{-i\xi x}\, d(\mu * \nu)(x) \\ &= \int_{\mathbb{R}} e^{-i\xi(y+z)}\, d\mu(y)\, d\nu(z) = \left(\int_{\mathbb{R}} e^{-i\xi y}\, d\mu(y)\right)\left(\int_{\mathbb{R}} e^{-i\xi z}\, d\nu(z)\right) = \widehat{\mu}(\xi)\widehat{\nu}(\xi).\end{aligned}$$

(d) We have an analog of Parseval's formula.

Theorem 2.3 (Parseval's formula)**.** *Let $\mu \in M(\mathbb{R})$ and $f \in L^1(\mathbb{R}, m) \cap C(\mathbb{R})$ such that $\widehat{f} \in L^1(\widehat{\mathbb{R}}, m)$, then*

$$\int_{\mathbb{R}} f(x)\, d\mu(x) = \frac{1}{2\pi} \int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\widehat{\mu}(-\xi)\, d\xi. \tag{16}$$

Proof. Note that $L^1(\mathbb{R}) \cap C(\mathbb{R}) \subset C_0(\mathbb{R})$, thus the integral on the left of (16) makes sense. Indeed, for each $\varepsilon > 0$ the set $\{x \in \mathbb{R} : |f(x)| \geq \varepsilon\}$ has measure at most $\varepsilon^{-1}\|f\|_{L^1}$, thus it must be a bounded set and hence it is compact in $\mathbb{R}$ since it is closed. We can use the Fourier inversion formula to deduce that

$$f(x) = \frac{1}{2\pi} \int_{\widehat{\mathbb{R}}} \widehat{f}(\xi) e^{i\xi x}\, d\xi$$

and hence since $\widehat{f} \in L^1(\widehat{\mathbb{R}})$, Fubini's theorem reads

$$\begin{aligned}\int_{\mathbb{R}} f(x)\, d\mu(x) &= \frac{1}{2\pi} \int_{\mathbb{R}} \left(\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi) e^{i\xi x}\, d\xi \right) d\mu(x) \\ &= \frac{1}{2\pi} \int_{\widehat{\mathbb{R}}} \widehat{f}(\xi) \left(\int_{\mathbb{R}} e^{i\xi x} d\mu(x) \right) d\xi = \frac{1}{2\pi} \int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\widehat{\mu}(-\xi)\, d\xi.\end{aligned}$$

□

Another way to rewrite this result is

$$\int_{\mathbb{R}} f(x)\, \overline{d\mu(x)} = \frac{1}{2\pi} \int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\overline{\widehat{\mu}(\xi)}\, d\xi.$$

(e) (Uniqueness theorem)If $\widehat{\mu}(\xi) = 0$ for all $\xi \in \widehat{\mathbb{R}}$, then $\mu = 0$.

A departure from the theory of L^1-Fourier transforms is the falling of the Riemann-Lebesgue lemma: the Fourier-Stieltjes transform of a measure μ need not vanish at infinity.

5. The assumption $\widehat{f} \in L^1(\widehat{\mathbb{R}})$ justifies the change of order of integration by Fubini's theorem; however it is not really needed. In particular, we have the following theorems.

Theorem 2.4. *If $f \in L^1(\mathbb{R}) \cap C(\mathbb{R})$ then:*

$$\int_{\mathbb{R}} f(x)\, d\mu(x) = \lim_{\lambda \longrightarrow \infty} \frac{1}{2\pi} \int_{-\lambda}^{\lambda} \left(1 - \frac{|\xi|}{\lambda}\right) \widehat{f}(\xi)\widehat{\mu}(-\xi)\, d\xi.$$

Proof. Recall the Fejer's kernel $\mathcal{K}_\lambda(x) = \lambda\mathcal{K}(\lambda x)$ satisfies $\widehat{\mathcal{K}}_\lambda(\xi) = \left(1 - \frac{|\xi|}{\lambda}\right)\chi_{[-\lambda,\lambda]}(\xi)$. Now we have $f_\lambda = \mathcal{K}_\lambda * f \in L^1(\mathbb{R}) \cap C(\mathbb{R})$ and $\widehat{f_\lambda} = \widehat{K}_\lambda \cdot \widehat{f} \in L^1(\widehat{\mathbb{R}})$, thus Parseval's formula 2.3 implies

$$\int_{\mathbb{R}} f_\lambda(x)\, d\mu(x) = \frac{1}{2\pi} \int_{\widehat{\mathbb{R}}} \widehat{f_\lambda}(\xi)\widehat{\mu}(-\xi)\, d\xi \quad \Longleftrightarrow \quad \int_{\mathbb{R}} (\mathcal{K}_\lambda * f)(x)\, d\mu(x) = \frac{1}{2\pi} \int_{\widehat{\mathbb{R}}} \widehat{K}_\lambda(\xi)\widehat{f}(\xi)\widehat{\mu}(-\xi)\, d\xi$$

Since $\mathcal{K}_\lambda * f \longrightarrow f$ everywhere since $f \in C(\mathbb{R})$ as $\lambda \longrightarrow \infty$ (summability kernel), and since $L^1(\mathbb{R}) \cap C(\mathbb{R}) \subset C_0(\mathbb{R})$, we have $\|\mathcal{K}_\lambda * f\|_{L^\infty(\mathbb{R})} \leq \|\mathcal{K}_\lambda\|_{L^1(\mathbb{R})}\|f\|_{L^\infty} = \|f\|_{L^\infty} \in L^1(\mathbb{R}, \mu)$, thus the dominated convergence theorem applies to $(\mathbb{R}, \mu)$ gives us the desired formula. □

As a corollary we have:

Corollary 2.5. *If $f \in L^1(\mathbb{R}) \cap C(\mathbb{R})$ such that $\widehat{f}(\xi)\widehat{\mu}(-\xi) \in L^1(\widehat{\mathbb{R}})$ then (16) holds true.*

Proof. Using the same technique $\mathcal{K}_\lambda * f$, then if $\widehat{f}(\xi)\widehat{\mu}(-\xi) \in L^1(\widehat{\mathbb{R}})$ in the limit we obtain (16) by dominated convergence theorem. □

6. The problem of characterizing Fourier-Stieltjes transforms among $\mathrm{BUC}(\widehat{\mathbb{R}})$ is very hard. One immediate result we can have is if $f \in \mathcal{F}L^1$, say $f = \widehat{g}$, then with $\mu = g(x)\,dx$ we have $f(\xi) = \widehat{g}(\xi) = \widehat{\mu}(\xi)$.

Theorem 2.6. *If $\mu \in M(\mathbb{R})$ and $\mathcal{V}_\lambda$ is de la Vallée Poussin's kernel $\mathcal{V}_\lambda = 2\mathcal{K}_{2\lambda} - \mathcal{K}_\lambda$ then $\mu * \mathcal{V}_\lambda \in L^1(\mathbb{R})$ and $\widehat{\mu * \mathcal{V}_\lambda}(\xi) = \widehat{\mu}(\xi)$ for $|\xi| \le \lambda$.*

Proof. First of all observe that $\|\mathcal{V}_\lambda\|_{L^1} \le 2\|\mathcal{K}_{2\lambda}\|_{L^1} + \|\mathcal{K}_\lambda\|_{L^1} = 3$. We have $(x,y) \longmapsto \mathcal{V}_\lambda(x-y)$ from $(\mathbb{R}, m) \times (\mathbb{R}, \mu) \longrightarrow \mathbb{R}$ is measurable and satisfies

$$\int_{\mathbb{R}} |\mathcal{V}_\lambda(x-y)|\,dx \le \|\mathcal{V}_\lambda\|_{L^1} \quad \text{for all } y \in \mathbb{R}, \qquad \text{and} \qquad \int_{\mathbb{R}}\int_{\mathbb{R}} |\mathcal{V}_\lambda(x-y)|\,dx\,d\mu(y) \le 3\|\mu\|_{M(\mathbb{R})}.$$

Thus Tonelli's theorem concludes that $(x,y) \longmapsto \mathcal{V}_\lambda(x-y) \in L^1\Big((\mathbb{R}, m) \times (\mathbb{R}, \mu)\Big)$, hence by Fubini's theorem we have $\mu * \mathcal{V}_\lambda \in L^1(\mathbb{R})$ since

$$\int_{\mathbb{R}} \Big|\mu * \mathcal{V}_\lambda(x)\Big|\,dx = \int_{\mathbb{R}}\int_{\mathbb{R}} \Big|\mathcal{V}_\lambda(x-y)\Big|\,d\mu(y)\,dx = \int_{\mathbb{R}}\int_{\mathbb{R}} \Big|\mathcal{V}_\lambda(x-y)\Big|\,dx\,d\mu(y) \le 3\|\mu\|_{M(\mathbb{R})}.$$

The rest is straight-forward from the Fourier transform of $\mathcal{V}_\lambda$, which is

$$\widehat{\mathcal{V}_\lambda}(\xi) = \begin{cases} 1 & |\xi| \le \lambda, \\ 2 - \frac{|\xi|}{\lambda} & \lambda \le |\xi| \le 2\lambda, \\ 0 & 2\lambda \le |\xi|. \end{cases} \qquad \Longrightarrow \qquad \widehat{\mu * \mathcal{V}_\lambda}(\xi) = \widehat{\mathcal{V}_\lambda}(\xi)\widehat{\mu}(\xi) = \mu(\xi)$$

if $|\xi| \le \lambda$. The proof is complete. □

A further characterization is given below.

Theorem 2.7. *Let $\varphi \in C(\widehat{\mathbb{R}})$, define*

$$\Phi_\lambda(x) = \frac{1}{2\pi}\int_{-\lambda}^{\lambda} \left(1 - \frac{|\xi|}{\lambda}\right)\varphi(\xi)e^{i\xi x}\,d\xi.$$

Then φ is a Fourier-Stieltjes transform iff $\Phi_\lambda \in L^1(\mathbb{R})$ for all $\lambda > 0$ and $\|\Phi_\lambda\|_{L^1(\mathbb{R})}$ is bounded as $\lambda \longrightarrow \infty$.

Proof. If $\varphi = \widehat{\mu}$ for some $\mu \in M(\mathbb{R})$, then $\varphi = \widehat{\mu} \in \mathrm{BUC}(\widehat{\mathbb{R}})$, thus $\widehat{\mathcal{K}}_\lambda(\cdot)\widehat{\mu}(\cdot) \in L^1(\widehat{\mathbb{R}})$ and $\mathcal{K}_\lambda * \mu \in L^1(\mathbb{R})$ by an analog to the argument in the proof of theorem 2.6. The Fourier inversion formula reads

$$(\mathcal{K}_\lambda * \mu)(x) = \frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{\mathcal{K}}_\lambda(\xi)\widehat{\mu}(\xi)e^{i\xi x}\,d\xi = \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\left(1 - \frac{|\xi|}{\lambda}\right)\varphi(\xi)e^{i\xi x}\,d\xi = \Phi_\lambda(x).$$

Therefore $\Phi_\lambda = \mathcal{K}_\lambda * \mu$ for all $\lambda > 0$, and clearly $\|\Phi_\lambda\|_{L^1} \le \|\mathcal{K}_\lambda\|_{L^1}\|\mu\|_{M(\mathbb{R})} = \|\mu\|_{M(\mathbb{R})}$.

For the converse, for each $\lambda > 0$ we can define the corresponding measure $\mu_\lambda = \Phi_\lambda(x)\,dx \in M(\mathbb{R})$. Since $\|\mu_\lambda\|_{M(\mathbb{R})} \le C$ for all $\lambda > 0$, Banach-Alaoglu's theorem implies that there exists a sequence $\lambda_n \longrightarrow \infty$ and $\mu \in M(\mathbb{R})$ such that $\mu_{\lambda_n} \overset{*}{\rightharpoonup} \mu$ in $M(\mathbb{R})$. Recall that $\widehat{\mu_\lambda}(\xi) = \widehat{\Phi_\lambda}(\xi)$, and furthermore $\widehat{\mathcal{K}}_\lambda(\xi)\varphi(\xi) \in L^1(\widehat{\mathbb{R}})$ reads

$$\mathcal{F}\Big[\widehat{\mathcal{K}}_\lambda(\cdot)\varphi(\cdot)\Big](-x) = \mathcal{F}^{-1}\Big[\widehat{\mathcal{K}}_\lambda(\cdot)\varphi(\cdot)\Big](x) = \Phi_\lambda(x) \in L^1(\mathbb{R})$$

which, by the Fourier inversion formula gives us $\widehat{\Phi}_\lambda(\xi) = \widehat{\mathcal{K}}_\lambda(\xi)\varphi(\xi)$ for $\xi \in \widehat{\mathbb{R}}$. We also obtain

$$\Big\|\widehat{\mathcal{K}}_\lambda(\xi)\varphi(\xi)\Big\|_u = \Big\|\widehat{\Phi}_\lambda\Big\|_u \le \|\Phi_\lambda\|_{L^1} \le C \qquad \Longrightarrow \qquad \|\varphi\|_u \le C$$

by sending $\lambda \longrightarrow \infty$. In order to show $\varphi = \widehat{\mu}$, it suffices to show that (since both φ and $\widehat{\mu}$ are continuous)

$$\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\widehat{\mu}(\xi)\,d\xi = \int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\varphi(\xi)\,d\xi \qquad \text{for all} \qquad \widehat{f} \in C_c^\infty(\widehat{\mathbb{R}}). \tag{17}$$

For such test functions $\widehat{f}$ we have $f \in \mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R}) \cap C_0(\mathbb{R})$. The key tool is Parseval's identity (16), indeed we have

$$\int_{\mathbb{R}} f(x)\, d\mu_\lambda(x) = \frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\widehat{\mu}_\lambda(-\xi)\, d\xi = \frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\widehat{\Phi}_\lambda(-\xi)\, d\xi = \frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\widehat{\mathcal{K}}_\lambda(-\xi)\varphi(-\xi)\, d\xi.$$

Let $\lambda \longrightarrow \infty$ along the sequence λ_n and using $\mu_{\lambda_n} \overset{*}{\rightharpoonup} \mu$ on the left and dominated convergence theorem on the right (thanks to φ is uniformly bounded) we obtain

$$\int_{\mathbb{R}} f(x)\, d\mu(x) = \frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\varphi(-\xi)\, d\xi \qquad \Longrightarrow \qquad \frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\widehat{\mu}(-\xi)\, d\xi = \frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\varphi(-\xi)\, d\xi$$

by using the Parseval's identity again for the left hand side. Thus (17) is justified and the proof is complete. □

Note that the application of Parseva's formula above is typical and is the standard way to utilize the weak* limit in $M(\mathbb{R})$. Nothing like that was needed in the case of $M(\mathbb{T})$ since weak* convergence in $M(\mathbb{R})$ implies point-wise convergence of the Fourier-Stieltjes coefficients (the exponentials belong to $C(\mathbb{T})$ of which $M(\mathbb{T})$ is the dual). The exponentials on $\mathbb{R}$ do not belong to $C_0(\mathbb{R})$ and it is false that weak* convergence in $M(\mathbb{R})$ implies pointwise convergence of the Fourier-Stieltjes transforms. However the argument above gives:

Lemma 2.8. *Let $\mu_n \overset{*}{\rightharpoonup} \mu$ in $M(\mathbb{R})$ such that $\widehat{\mu}_n(\xi) \longrightarrow \varphi(\xi)$ point-wise for some $\varphi \in C(\widehat{\mathbb{R}})$, then $\widehat{\mu} = \varphi$.*

Proof. For a test function $\widehat{f} \in C_c^\infty(\widehat{\mathbb{R}})$, recall that $f \in \mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R}) \cap C_0(\mathbb{R})$ and thus the Parseval's formula reads

$$\int_{\mathbb{R}} f(x)\, d\mu_n(x) = \frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\widehat{\mu}_n(-\xi)\, d\xi \qquad \Longrightarrow \qquad \int_{\mathbb{R}} f(x)\, d\mu(x) = \frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\varphi(-\xi)\, d\xi$$

by sending $n \longrightarrow \infty$. The result follows from Parseval's formula again. □

A related result is the following:

Lemma 2.9. *If X is a LCH space and $\{\mu_n\} \subset M(X)$, $\mu_n \longrightarrow \mu$ vaguely, and $\|\mu_n\| \longrightarrow \|\mu\|$, then we have $\int_X f\, d\mu_n \longrightarrow \int_X f\, d\mu$ for every $f \in \mathrm{BC}(X)$. Moreover, the hypothesis $\|\mu_n\| \longrightarrow \|\mu\|$ cannot be omitted.*

7. A similar application of Parseval's formula gives the following useful criterion:

Theorem 2.10. *A function φ defined and continuous on $\widehat{\mathbb{R}}$, is a Fourier-Stieltjes transform if and only if there exists a constant C such that*

$$\left|\frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\varphi(-\xi)\, d\xi\right| \le C \sup_{x\in\mathbb{R}} |f(x)| \tag{18}$$

for every $f \in L^1(\mathbb{R})$ such that $\widehat{f}$ has compact support.

Proof. First of all, let $\mathcal{T}(\mathbb{R}) = \{f \in L^1(\mathbb{R}) : \widehat{f} \in C_c(\widehat{\mathbb{R}})\}$ then $\mathcal{T}(\mathbb{R}) \subset L^1(\mathbb{R}) \cap C_0(\mathbb{R})$ by Riemann-Lebesgue lemma. By Parseval's formula

$$\int_{\mathbb{R}} f(x)\, d\mu(x) = \frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\widehat{\mu}(-\xi)\, d\xi \qquad \Longrightarrow \qquad \left|\frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\varphi(-\xi)\, d\xi\right| \le \|\mu\| \sup_{x\in\mathbb{R}} |f(x)|.$$

Conversely, assuming (18) holds true for all $f \in \mathcal{T}(\mathbb{R})$. Let's define the linear functional

$$\Lambda : \mathcal{T}(\mathbb{R}) \longrightarrow \mathbb{C} \qquad \text{maps} \qquad f \longmapsto \frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\varphi(-\xi)\, d\xi \qquad \text{with} \qquad |\Lambda(f)| \le C\|f\|_u. \tag{19}$$

We claim that $\mathcal{T}(\mathbb{R})$ is dense in $C_0(\mathbb{R})$. Indeed, given any $f \in C_0(\mathbb{R})$, we can find $f_n \in C_c^\infty(\mathbb{R})$ with $\|f_n - f\|_u \longrightarrow 0$. For each $f_n \in C_c^\infty(\mathbb{R})$, in turn we have $\widehat{f_n} \in \mathcal{S}(\mathbb{R}) \subset L^1(\widehat{\mathbb{R}})$, there exists $\widehat{g_n} \in C_c^\infty(\widehat{\mathbb{R}})$ such that $\|\widehat{g_n}-\widehat{f_n}\|_{L^1(\widehat{\mathbb{R}})} < 2^{-n}$, which implies that $\|g_n - f_n\|_u < 2^{-n}$, thus we have shown that $g_n \in \mathcal{T}(\mathbb{R})$ and $g_n \longrightarrow f$ uniformly. Therefore the linear functional Λ defined in (19) can be extended uniquely to $\Lambda : C_0(\mathbb{R}) \longrightarrow \mathbb{C}$ with the same bound. By Rieze's representation theorem, there exists a unique measure $\mu \in M(\mathbb{R})$ such that $\|\mu\| = \|\Lambda\| \le C$ and

$$\Lambda(f) = \int_{\mathbb{R}} f\, d\mu = \frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\varphi(-\xi)\, d\xi \quad \text{for} \quad f \in \mathcal{T}(\mathbb{R}).$$

By Parseval's formula we deduce that $\widehat{u} = \varphi$ and the proof is complete. □

One observation we didn't use in the proof is that, if we assume (18) holds for $f \in \mathcal{T}(\mathbb{R})$, we can deduce first that $\varphi \in L^\infty(\widehat{\mathbb{R}})$. Indeed, it is obvious that $\mathcal{F}[\mathcal{T}(\mathbb{R})]$ is dense in $C_0(\widehat{\mathbb{R}})$, since $C_c^\infty(\widehat{\mathbb{R}})$ is dense in $C_0(\widehat{\mathbb{R}})$ and $\mathcal{F}^{-1}[C_c^\infty(\widehat{\mathbb{R}})] \subset \mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R})$. Thus the linear map

$$\Phi : C_c(\widehat{\mathbb{R}}) \longrightarrow \mathbb{C} \quad \text{maps} \quad f \longmapsto \int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\varphi(-\xi)\, d\xi \quad \text{with} \quad |\Phi(f)| \le C\|f\|_u \le \|\widehat{f}\|_{L^1(\widehat{\mathbb{R}})}$$

extends uniquely to a bounded linear functional on $L^1(\widehat{\mathbb{R}})$. The Rieze's representation $(L^1)^* = L^\infty$ gives us the unique $u \in L^\infty(\widehat{\mathbb{R}})$ such that $\varphi(\xi) = u(-\xi)$ and hence $\varphi \in L^\infty(\widehat{\mathbb{R}})$.

The family of function f such that (18) holds true can be taken in many ways. We need a collection of functions $\{f : f \in \mathcal{J}\}$ such that they are dense in $C_0(\mathbb{R})$ and $\{\widehat{f} : f \in \mathcal{J}\}$ is dense in $C_0(\widehat{\mathbb{R}})$, for example $\{f : \widehat{f} \in C_c^\infty(\widehat{\mathbb{R}})\}$ or $\{f : f \in C_c^\infty(\mathbb{R})\}$.

8. With measures on $\mathbb{R}$ we can associate measures on $\mathbb{T}$ simply by integrating 2π-periodic functions. Formally, if E is a Borel set on $\mathbb{T}$, which is identified with $(-\pi, \pi]$, we denote by $E_n = E + 2\pi n$ and $\tilde{E} = \bigcup_{n\in\mathbb{Z}} E_n$. If $\mu \in M(\mathbb{R})$ we define
$$\mu_{\mathbb{T}}(E) = \mu(\tilde{E}).$$
It is clear that $\mu_{\mathbb{T}}$ is a measure on $\mathbb{T}$ and that identifying continuous functions on $\mathbb{T}$ with 2π-periodic functions on $\mathbb{R}$ gives us
$$\int_{\mathbb{R}} f(x)\, dx = \int_{\mathbb{T}} f(t)\, dt.$$
The mapping $\mu \longmapsto \mu_{\mathbb{T}}$ is an operator of norm 1 from $M(\mathbb{R})$ onto $M(\mathbb{T})$. It also follows that for $n \in \mathbb{Z}$ then $\widehat{\mu}(n) = \widehat{\mu_{\mathbb{T}}}(n)$, thus the restriction of a Fourier-Stieltjes transform to $\mathbb{Z}$ gives a sequence of Fourier-Stieltjes coefficients.

Theorem 2.11. *A function φ defined and continuous on $\widehat{\mathbb{R}}$ is a Fourier-Stieltjes transform if and only if there exists $C > 0$ such that for all $\lambda > 0$, $\{\varphi(\lambda n)\}_{n\in\mathbb{Z}}$ are the Fourier-Stieltjes coefficients of a measure $\mu_{\mathbb{T}} \in M(\mathbb{T})$ with $\|\mu_{\mathbb{T}}\|_{M(\mathbb{T})} \le C$.*

Proof. If $\varphi = \widehat{\mu}$ for some $\mu \in M(\mathbb{R})$ then $\varphi(n) = \widehat{\mu}(n) = \widehat{\mu_{\mathbb{T}}}(n)$ for all $n \in \mathbb{Z}$, and $\|\mu_{\mathbb{T}}\|_{M(\mathbb{T})} \le \|\mu\|_{M(\mathbb{R})}$. Let's denote by μ_λ the measure in $M(\mathbb{R})$ satisfying

$$\int_{\mathbb{R}} f(x)\, d\mu_\lambda(x) = \int_{\mathbb{R}} f(\lambda x)\, d\mu(x) \qquad \text{for all} \qquad f \in C_0(\mathbb{R})$$

then we have $\|\mu_\lambda\|_{M(\mathbb{R})} = \|\mu\|_{M(\mathbb{R})}$ for all $\lambda > 0$, and clearly $\widehat{\mu_\lambda}(\xi) = \widehat{\mu}(\lambda\xi) = \varphi(\lambda\xi)$ for $\xi \in \widehat{\mathbb{R}}$. Thus after transferring to a measure in $M(\mathbb{T})$ we obtain

$$\mathcal{F}\Big[(\mu_\lambda)_{\mathbb{T}}\Big](n) = \widehat{\mu_\lambda}(n) = \varphi(\lambda n)$$

thus $\{\varphi(\lambda n)\}_{n\in\mathbb{Z}}$ are the Fourier-Stieltjes coefficients $(\mu_\lambda)_{\mathbb{T}} \in M(\mathbb{T})$ with $\|(\mu_\lambda)_{\mathbb{T}}\|_{M(\mathbb{T})} \le \|\mu\|_{M(\mathbb{R})}$.

Conversely, if there exists $C > 0$ such that for all $\lambda > 0$ we have $\{\varphi(\lambda n)\}_{n\in\mathbb{Z}}$ are the Fourier-Stieltjes coefficients $\mu_\lambda \in M(\mathbb{T})$ with $\|\mu_\lambda\|_{M(\mathbb{T})} \le C$, then we want to estimate the integral

$$\frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\varphi(-\xi)\, d\xi \qquad \text{for} \qquad f \in L^1(\mathbb{R}), \widehat{f} \in C_c^\infty(\widehat{\mathbb{R}})$$

in order to use theorem 2.10. Let's assume supp $\widehat{f} \subset [-R,R]$ and $\lambda = R/m$ with $m \in \mathbb{N}$, we can divide $[-R,R]$ into $2m$ siub-interval of length λ, and we can approximate the integral by Riemann's sum, that is given $\varepsilon > 0$, there exists $m \in \mathbb{N}$ large so that

$$\left| \frac{1}{2\pi} \int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\varphi(-\xi)\, d\xi - \frac{\lambda}{2\pi} \sum_{n=-m}^{m} \widehat{f}(n\lambda)\varphi(-n\lambda) \right| < \varepsilon. \tag{20}$$

Recall that if $f \in L^1(\mathbb{R})$ then for

$$\psi_\lambda(t) = \sum_{-\infty}^{\infty} f\left(\frac{t-2\pi n}{\lambda}\right) \qquad \text{we have} \qquad \psi_\lambda \in L^1(\mathbb{T}) \quad \text{with} \quad \widehat{\psi_\lambda}(n) = \frac{\lambda}{2\pi}\widehat{f}(n\lambda).$$

Thus if $\varphi = \widehat{\mu_\lambda}$ for some $\mu_\lambda \in M(\mathbb{T})$ with $\|\mu_\lambda\|_{M(\mathbb{T})} \leq C$ then

$$\frac{\lambda}{2\pi} \sum_{n=-m}^{m} \widehat{f}(n\lambda)\varphi(-n\lambda) = \sum_{-m}^{m} \widehat{\psi_\lambda}(n)\widehat{\mu_\lambda}(-n)$$

The Parseval's formula for $\overline{\mu_\lambda} \in M(\mathbb{T})$ reads

$$\langle \psi_\lambda, \overline{\mu_\lambda} \rangle_{M(\mathbb{T})} = \lim_{k \longrightarrow \infty} \sum_{-k}^{k} \left(1 - \frac{|n|}{k+1}\right) \widehat{\psi_\lambda}(n)\widehat{\mu_\lambda}(-n) = \sum_{-m}^{m} \widehat{\psi_\lambda}(n)\widehat{\mu_\lambda}(-n) = \frac{\lambda}{2\pi} \sum_{n=-m}^{m} \widehat{f}(n\lambda)\varphi(-n\lambda).$$

Thus we obtain

$$\left| \frac{\lambda}{2\pi} \sum_{n=-m}^{m} \widehat{f}(n\lambda)\varphi(-n\lambda) \right| \leq \|\mu_\lambda\|_{M(\mathbb{T})} \sup_{t \in \mathbb{T}} |\psi_\lambda(t)| \leq C \sup_{t \in \mathbb{T}} |\psi_\lambda(t)| \tag{21}$$

for all $\lambda > 0$. Since $\widehat{f} \in C_c^\infty(\widehat{\mathbb{R}})$, we have $f = \mathcal{F}^{-1}(\widehat{f}) \in \mathcal{S}(\mathbb{R})$ which decays very fast at $|x| \longrightarrow \infty$, thus if we choose λ small enough we obtain

$$\sup_{t \in \mathbb{T}} |\psi_\lambda(t)| \leq \sup_{x \in \mathbb{R}} |f(x)| + \varepsilon. \tag{22}$$

This fact together with (20) and (21) implies

$$\left| \frac{1}{2\pi} \int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\varphi(-\xi)\, d\xi \right| < C \sup_{x \in \mathbb{R}} |f(x)| + (C+1)\varepsilon.$$

Since $\varepsilon > 0$ is chosen arbitrary, we obtain the result from theorem 2.10. ☐

The estimate (22) can be proved precisely as following. Since $\widehat{f} \in \mathcal{S}(\widehat{\mathbb{R}})$ we have $f \in \mathcal{S}(\mathbb{R})$ as well, thus

$$\sup_{x \in \mathbb{R}} (2\pi + |x|)^2 |f(x)| \leq C \qquad \Longrightarrow \qquad |f(x)| \leq \frac{C}{(2\pi + |x|)^2} \quad \text{for all} \quad x \in \mathbb{R}.$$

For $t \in \mathbb{T} \sim [0, 2\pi)$ and $n \neq 0$ we have

$$\left| f\left(\frac{t-2\pi n}{\lambda}\right) \right| \leq \frac{C\lambda^2}{\big(2\pi + |t - 2\pi n|\big)^2} \leq \frac{C\lambda^2}{\big(2\pi + 2\pi|n| - t\big)^2} \leq \frac{C\lambda^2}{4\pi^2|n|^2}.$$

Hence

$$|\psi_\lambda(t)| \leq \left| f\left(\frac{t}{\lambda}\right) \right| + \sum_{n \neq 0} \left| f\left(\frac{t-2\pi n}{\lambda}\right) \right| \leq \sup_{x \in \mathbb{R}} |f(x)| + \frac{C\lambda^2}{4\pi^2} \sum_{n \neq 0} \frac{1}{n^2} = \sup_{x \in \mathbb{R}} |f(x)| + \frac{C\lambda^2}{24}$$

and thus the result follows when we choose λ small enough.

9. Parseval's formula also offers an obvious criterion for determining when a function φ is the Fourier-Stieltjes transform of a positive measure. The analog to theorem 2.10 is

Theorem 2.12. *A function $\varphi \in BC(\widehat{\mathbb{R}})$ is the Fourier-Stieltjes transform of a positive measure on $\mathbb{R}$ if and only if*

$$\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\varphi(-\xi)\, d\xi \geq 0 \tag{23}$$

for every $f \in \mathcal{P}(\mathbb{R}) = \{f \geq 0, f \in C_c^\infty(\mathbb{R})\}$.

Proof. If $\varphi = \widehat{\mu}$ for some positive measure μ on $\mathbb{R}$, then Parseval's formula implies (23) obviously. Conversely if (23) holds true for all $f \in \mathcal{P}(\mathbb{R})$, then it is also true for all non-negative functions $\{f \in L^1(\mathbb{R}), \widehat{f} \in L^1(\widehat{\mathbb{R}})\}$.

- Let $\zeta \in C_c^\infty(\mathbb{R})$ be the standard mollifier, i.e., $0 \leq \zeta \leq 1$, supp $\zeta \subset (-1,1)$, $\zeta = \frac{1}{2\pi}$ in a neighborhood of the origin, and $\int_{\mathbb{R}} \zeta(x)\, dx = 1$. For each $\varepsilon > 0$ let $\zeta_\varepsilon(x) = \varepsilon^{-1}\zeta\left(\varepsilon^{-1}x\right)$, then

$$\varepsilon\zeta_\varepsilon(x) = \zeta\left(\frac{x}{\varepsilon}\right) \quad \Longrightarrow \quad \varepsilon\widehat{\zeta_\varepsilon}(\xi) = \varepsilon\widehat{\zeta}(\varepsilon\xi) \quad \Longrightarrow \quad \int_{\widehat{\mathbb{R}}} \varepsilon\widehat{\zeta_\varepsilon}(\xi)\, d\xi = \varepsilon\int_{\widehat{\mathbb{R}}} \widehat{\zeta}(\varepsilon\xi)\, d\xi = 2\pi\zeta(0) = 1$$

  by the Fourier inversion formula. In other words, $\{\varepsilon\widehat{\zeta_\varepsilon}\}_{\varepsilon\longrightarrow\infty}$ forms a summability sequence, hence for every $\widehat{f} \in C_c(\widehat{\mathbb{R}}) \subset L^1(\widehat{\mathbb{R}})$ we have

$$\left\|\widehat{f} * (\varepsilon\widehat{\zeta_\varepsilon}) - \widehat{f}\right\|_{L^1(\widehat{\mathbb{R}})} \longrightarrow 0 \qquad \text{as} \qquad \varepsilon \longrightarrow \infty. \tag{24}$$

- For a non-negative $f \in L^1(\mathbb{R}) \cap C^\infty(\mathbb{R})$ with $\widehat{f} \in C_c(\widehat{\mathbb{R}})$ we have $f(x)\zeta\left(\frac{x}{\varepsilon}\right) = f\,(\varepsilon\zeta_\varepsilon) \in C_c^\infty(\mathbb{R})$ and is non-negative, thus (23) reads

$$\int_{\widehat{\mathbb{R}}} \left(\widehat{f} * \varepsilon\widehat{\zeta_\varepsilon}\right)(\xi)\varphi(-\xi)\, d\xi \geq 0 \qquad \text{for all} \qquad \varepsilon > 0.$$

  From that and (24) as $\varepsilon \longrightarrow \infty$, (23) is true for non-negative $f \in L^1(\mathbb{R}) \cap C^\infty(\mathbb{R})$ with $\widehat{f} \in C_c(\widehat{\mathbb{R}})$.

- Finally if $f \in L^1(\mathbb{R})$ with $\widehat{f} \in L^1(\widehat{\mathbb{R}})$, then let $\eta \in C_c^\infty(\widehat{\mathbb{R}})$ be the standard symmetric mollifier, i.e., $0 \leq \eta \leq 1$, supp $\eta \subset (-1,1)$, $\eta = 1$ in a neighborhood of the origin, and $\int_{\widehat{\mathbb{R}}} \eta(\xi)\, d\xi = 2\pi$. For each $\varepsilon > 0$ let $\eta_\varepsilon(\xi) = \varepsilon^{-1}\eta\left(\varepsilon^{-1}\xi\right)$, then

$$f_\varepsilon = f * \mathcal{F}^{-1}(\varepsilon\eta_\varepsilon) \in C^\infty(\mathbb{R}) \cap L^1(\mathbb{R}) \qquad \text{and} \qquad \widehat{f_\varepsilon}(\xi) = \widehat{f}(\xi)\eta\left(\frac{\xi}{\varepsilon}\right) \in C_c(\widehat{\mathbb{R}}).$$

  The result from the previous step implies

$$\int_{\widehat{\mathbb{R}}} \widehat{f}(\xi)\eta\left(\frac{\xi}{\varepsilon}\right)\varphi(-\xi)\, d\xi \geq 0 \qquad \text{for all} \qquad \varepsilon > 0.$$

  As $\varepsilon \longrightarrow \infty$, by the dominated convergence theorem we obtain (23) is true for all non-negative $f \in L^1(\mathbb{R})$ with $\widehat{f} \in L^1(\widehat{\mathbb{R}})$.

Going back to our problem, recalling that with the Fejer's kernel $\{\mathcal{K}_\lambda\}_{\lambda>0}$ we have $\{\lambda^{-1}\mathcal{K}_\lambda\}_{\lambda>0}$ satisfies $\lambda^{-1}\widehat{\mathcal{K}}_\lambda \longrightarrow \delta_0$ in $\mathcal{S}'(\widehat{\mathbb{R}})$ and thus (if φ is a Schwartz function)

$$\lim_{\lambda\longrightarrow 0} \int_{\widehat{\mathbb{R}}} \lambda^{-1}\widehat{\mathcal{K}}_\lambda(\xi)\varphi(-\xi)\, d\xi = \varphi(0) \tag{25}$$

and then using this fact to proof an identity that theorem 2.10 requires. The identity (25) is indeed true even if we only have $\varphi \in \text{BC}(\widehat{\mathbb{R}})$ (actually we only need φ is continuous), since

$$\begin{aligned}
\frac{1}{\lambda}\int_{\widehat{\mathbb{R}}} \widehat{\mathcal{K}}_\lambda(\xi)\varphi(-\xi)\, d\xi &= \frac{1}{\lambda}\int_{-\lambda}^{\lambda} \left(1 - \frac{|\xi|}{\lambda}\right)\varphi(-\xi)\, d\xi \\
&= \frac{1}{\lambda}\int_0^{\lambda} \left(1 - \frac{\xi}{\lambda}\right)\Big(\varphi(\xi) + \varphi(-\xi)\Big)\, d\xi \\
&= \int_0^1 (1-\eta)\Big(\varphi(\lambda\eta) + \varphi(-\lambda\eta)\Big)\, d\eta \longrightarrow 2\varphi(0)\left(\int_0^1 (1-\eta)\, d\eta\right) = \varphi(0)
\end{aligned}$$

by the dominated convergence theorem. Now we show that

$$\left|\frac{1}{2\pi}\int_{\widehat{\mathbb{R}}}\widehat{f}(\xi)\varphi(-\xi)\,d\xi\right| \le 2\varphi(0)\left(\sup_{x\in\mathbb{R}}|f(x)|\right) \qquad \text{for} \qquad f\in C_c^\infty(\mathbb{R},\mathbb{C}) \tag{26}$$

then the proof will be complete by theorem 2.10. Let's consider a real-valued function $f\in C_c^\infty(\mathbb{R},\mathbb{R})$, since $\mathcal{K}(\lambda x)\longrightarrow \frac{1}{2\pi}$ as $\lambda\longrightarrow 0$ uniformly on compact sets, for given $\varepsilon>0$ there exists $\lambda_0>0$ small enough such that for all $\lambda\le\lambda_0$ then

$$\sup_{x\in\operatorname{supp} f}\left|\mathcal{K}(\lambda x)-\frac{1}{2\pi}\right|\le\varepsilon \qquad\Longrightarrow\qquad \frac{1}{2\pi}-\varepsilon\le\mathcal{K}(\lambda x)\le\frac{1}{2\pi}+\varepsilon \qquad \text{for all} \qquad x\in\operatorname{supp}(f)$$

which implies that

$$\begin{cases}\frac{f(x)}{2\pi}\le(\mathcal{K}(\lambda x)+\varepsilon)\sup|f| & \text{if}\quad f(x)\ge 0\\ \frac{f(x)}{2\pi}\le(\mathcal{K}(\lambda x)-\varepsilon)\sup|f| & \text{if}\quad f(x)\le 0\end{cases} \qquad\Longrightarrow\qquad f(x)\le 2\pi\Big(\mathcal{K}(\lambda x)+\varepsilon\Big)\sup|f| \quad \text{for all}\quad x\in\mathbb{R}.$$

In other words we have $2\pi\sup|f|\Big(\lambda^{-1}\mathcal{K}_\lambda(x)+\varepsilon\Big)-f(x)$ is a non-negative function which belongs to $C^\infty(\mathbb{R})\cap L^1(\mathbb{R})$ with its Fourier transform belongs to $L^1(\widehat{\mathbb{R}})$, thus (23) (applying to a bigger class of functions) reads

$$\int_{\widehat{\mathbb{R}}}2\pi\sup|f|\Big(\lambda^{-1}\widehat{\mathcal{K}}_\lambda(\xi)+\varepsilon 2\pi\delta_0\Big)\varphi(-\xi)\,d\xi\ge\int_{\widehat{\mathbb{R}}}\widehat{f}(\xi)\varphi(-\xi)\,d\xi$$

i.e.,

$$\frac{1}{2\pi}\int_{\widehat{\mathbb{R}}}\widehat{f}(\xi)\varphi(-\xi)\,d\xi\le\left(2\pi\varepsilon\varphi(0)+\int_{\widehat{\mathbb{R}}}\lambda^{-1}\widehat{\mathcal{K}}_\lambda(\xi)\varphi(-\xi)\,d\xi\right)\sup|f|.$$

Let $\lambda\longrightarrow 0$ and using (25) we obtain

$$\frac{1}{2\pi}\int_{\widehat{\mathbb{R}}}\widehat{f}(\xi)\varphi(-\xi)\,d\xi\le\varphi(0)(2\pi\varepsilon+1)\sup_{x\in\mathbb{R}}|f(x)|.$$

Let $\varepsilon\longrightarrow 0$ and replace f by $-f$ we obtain (26), after applying the same argument to the real part and the imaginary part of a complex-valued function f. □

10. An analog of theorem 2.11 but with positive measure in $M(\mathbb{T})$ is:

Theorem 2.13. *A function $\varphi\in C(\widehat{\mathbb{R}})$ is the Fourier-Stieltjes transform of a positive measure if and only for all $\lambda>0$, $\{\varphi(\lambda n)\}_{n\in\mathbb{Z}}$ are the Fourier-Stieltjest coefficients of a positive measure on $\mathbb{T}$.*

Proof. If $\varphi=\widehat{\mu}$ for some positive measure $\mu\in M(\mathbb{R})$ then $\varphi(n)=\widehat{\mu}(n)=\widehat{\mu_{\mathbb{T}}}(n)$ for all $n\in\mathbb{Z}$, and $\|\mu_{\mathbb{T}}\|_{M(\mathbb{T})}\le\|\mu\|_{M(\mathbb{R})}$. Let's denote by μ_λ the measure in $M(\mathbb{R})$ satisfying

$$\int_{\mathbb{R}}f(x)\,d\mu_\lambda(x)=\int_{\mathbb{R}}f(\lambda x)\,d\mu(x) \qquad \text{for all} \qquad f\in C_0(\mathbb{R})$$

then it is clear that μ_λ is positive and $\|\mu_\lambda\|_{M(\mathbb{R})}=\|\mu\|_{M(\mathbb{R})}$ for all $\lambda>0$, and $\widehat{\mu_\lambda}(\xi)=\widehat{\mu}(\lambda\xi)=\varphi(\lambda\xi)$ for $\xi\in\widehat{\mathbb{R}}$. Thus after transferring to a measure in $M(\mathbb{T})$ we obtain

$$\mathcal{F}\Big[(\mu_\lambda)_{\mathbb{T}}\Big](n)=\widehat{\mu_\lambda}(n)=\varphi(\lambda n)$$

thus $\{\varphi(\lambda n)\}_{n\in\mathbb{Z}}$ are the Fourier-Stieltjes coefficients of a positive measure $(\mu_\lambda)_{\mathbb{T}}\in M(\mathbb{T})$ with $\|(\mu_\lambda)_{\mathbb{T}}\|_{M(\mathbb{T})}\le\|\mu\|_{M(\mathbb{R})}$.

Conversely, if for all $\lambda>0$ we have $\{\varphi(\lambda n)\}_{n\in\mathbb{Z}}$ are the Fourier-Stieltjes coefficients of a positive measure $\mu_\lambda\in M(\mathbb{T})$, i.e., $\varphi(\lambda n)=\widehat{\mu_\lambda}(n)$ with $\mu_\lambda\ge 0$ in $M(\mathbb{T})$, then clearly $\varphi(0)=\widehat{\mu_\lambda}(0)=\|\mu_\lambda\|_{M(\mathbb{T})}$ for all $\lambda>0$. By theorem 2.11 there exists a measure $\mu\in M(\mathbb{R})$ such that $\varphi=\widehat{\mu}$. We have left to show that $\mu \geq 0$. Let's follow the procedure above for the "only if" part. Let $\nu_\lambda \in M(\mathbb{R})$ be the measure satisfying

$$\int_{\mathbb{R}} f(x)\, d\nu_\lambda(x) = \int_{\mathbb{R}} f(\lambda x)\, d\mu(x) \qquad \text{for all} \qquad f \in C_0(\mathbb{R})$$

then it is clear that $\|\nu_\lambda\|_{M(\mathbb{R})} = \|\mu\|_{M(\mathbb{R})}$ for all $\lambda > 0$, and $\widehat{\nu_\lambda}(\xi) = \widehat{\mu}(\lambda\xi) = \varphi(\lambda\xi)$ for $\xi \in \widehat{\mathbb{R}}$. Thus after transferring to a measure in $M(\mathbb{T})$ we obtain

$$\mathcal{F}[(\nu_\lambda)_{\mathbb{T}}](n) = \widehat{\nu_\lambda}(n) = \varphi(\lambda n) = \widehat{\mu_\lambda}(n)$$

for all $n \in \mathbb{Z}$. The uniqueness of Fourier-Stieltjes series in $M(\mathbb{T})$ implies that $\mu_\lambda \equiv (\nu_\lambda)_{\mathbb{T}} \geq 0$, and thus $\nu_\lambda \geq 0$. Hence

$$\int_{\mathbb{R}} f(x)\, d\nu_\lambda(x) = \int_{\mathbb{R}} f(\lambda x) d\mu(x) \geq 0 \qquad \text{for all} \qquad f \in C_0(\mathbb{R}), f \geq 0.$$

From that we obtain $\mu \geq 0$ and the proof is complete. □

11. A complex-valued function φ defined on $\widehat{\mathbb{R}}$ is said to be "*positive definite*" if, for every choice of $\xi_1, \ldots, \xi_m \in \widehat{\mathbb{R}}$ and complex numbers $z_1, \ldots, z_m$ we have

$$\sum_{1 \leq j,k \leq m} \varphi(\xi_j - \xi_k) z_j \overline{z_k} \geq 0.$$

Immediate consequences of this condition are $\varphi(-\xi) = \overline{\varphi(\xi)}$ and $|\varphi\xi)| \leq \varphi(0)$ for all $\xi \in \widehat{\mathbb{R}}$ if φ is positive definite.

Theorem 2.14 (Bochner)**.** *A function φ defined on $\widehat{\mathbb{R}}$ is a Fourier-Stieltjes transform of a positive measure if and only if it is positive definite and continuous.*

Proof. If $\varphi = \widehat{\mu}$ for some $\mu \geq 0$ in $M(\mathbb{R})$, then clearly φ is continuous, and for $\xi_1, \ldots, \xi_m \in \widehat{\mathbb{R}}$ and complex numbers $z_1, \ldots, z_m$ we have

$$\sum_{1 \leq j,k \leq m} \varphi(\xi_j - \xi_k) z_j \overline{z_k} = \sum_{1 \leq j,k \leq m} \left(\int_{\mathbb{R}} e^{-i\xi_j x} e^{i\xi_k x}\, d\mu(x) \right) z_j \overline{z_k} = \int_{\mathbb{R}} \left| \sum_{1 \leq j \leq m} z_j e^{-i\xi_j x} \right|^2 d\mu(x) \geq 0.$$

Conversely, if φ is positive definite then for any $\lambda > 0$ we have $\{\varphi(\lambda n)\}_{n \in \mathbb{Z}}$ is a positive definite sequence. By Herglotz's theorem, there exists a positive measure $\mu_\lambda \in M(\mathbb{T})$ such that $\varphi(\lambda n) = \widehat{\mu_\lambda}(n)$ for all $n \in \mathbb{Z}$, which implies that $\varphi = \widehat{\mu}$ for some positive measure $\mu \in M(\mathbb{R})$. □

12. Let $\mu \in M(\mathbb{R})$, let's define the measure $\mu^\# \in M(\mathbb{R})$ by

$$\mu^\#(E) = \overline{\mu(-E)} \qquad \text{for every Borel set} \qquad E \subset \mathbb{R}$$

or equivalently

$$\int_{\mathbb{R}} f(x)\, d\mu^\#(x) = \int_{\mathbb{R}} f(-x)\, \overline{d\mu(x)} \qquad \text{for every} \qquad f \in C_0(\mathbb{R}) \qquad (\text{or BC}(\mathbb{R})).$$

It is clear that $\widehat{\mu^\#}(\xi) = \overline{\widehat{\mu}(\xi)}$ for all $x \in \widehat{\mathbb{R}}$, thus $\widehat{\mu * \mu^\#}(\xi) = |\widehat{\mu}(\xi)|^2$ for all $x \in \widehat{\mathbb{R}}$. A measure $\mu \in M(\mathbb{R})$ is continuous if $\mu(\{x\}) = 0$ for every $t \in \mathbb{R}$, equivalently, μ is continuous if

$$\lim_{\eta \longrightarrow 0} \int_{x-\eta}^{x+\eta} |d\mu| = 0 \qquad \text{for every } x \in \mathbb{R}.$$

Theorem 2.15. *Every measure $\mu \in M(\mathbb{R})$ can be decomposed to a sum $\mu = \mu_c + \mu_d$ where μ_c is continuous and μ_d is discrete.*

If $\mu \in M(\mathbb{R})$ is a continuous measure then for any $\nu \in M(\mathbb{R})$, from the formula

$$(\mu * \nu)(E) = \iint_{\mathbb{R}^2} \mu(E-s)\, d\nu(s)$$

we deduce that $\mu * \nu$ is continuous. Since $\delta_a * \delta_b = \delta_{a+b}$, if $\mu = \sum_{j=1}^n a_j \delta_{s_j}$ and $\nu = \sum_{j=1}^n b_j \delta_{t_j}$ then

$$\mu * \nu = \sum_{j,k=1}^{n} a_j b_k \delta_{s_j + t_k}.$$

Let $\mu = \mu_c + \mu_d$ and $\mu^\# = \mu_c^\# + \mu_d^\#$ be the decompositions to continuous and discrete parts, we have

$$\mu * \mu^\# = \underbrace{\left(\mu_c * \mu_c^\# + \mu_c * \mu_d^\# + \mu_d * \mu_c^\#\right)}_{(\mu * \mu^\#)_c} + \left(\mu_d * \mu_d^\#\right).$$

Assume that $\mu_d = \sum_{j=1}^n a_j \delta_j$, then $\mu_d^\# = \sum_{j=1}^n \overline{a_j} \delta_{-s_j}$ and thus $\mu * \mu^\#(\{0\}) = \sum_{j=1}^n |a_j|^2$. Thus we have proved the following:

Lemma 2.16. *Let $\mu \in M(\mathbb{R})$, then*

$$\mu * \mu^\# \{0\} = \sum_{x \in \mathbb{R}} |\mu(\{x\})|^2.$$

In particular, μ is continuous if and only if $(\mu * \mu^\#)(\{0\}) = 0$.

13. An analog of Wiener's theorem but with measures in $M(\mathbb{R})$ is the following:

Theorem 2.17. *Let $\mu \in M(\mathbb{R})$, then the discrete part of μ can be recovered by*

$$\mu(\{x\}) = \lim_{\lambda \longrightarrow \infty} \frac{1}{2\lambda} \int_{-\lambda}^{\lambda} \widehat{\mu}(\xi) e^{i\xi x}\, d\xi.$$

As a consequence, we have

$$\sum_{x \in \mathbb{R}} |\mu(\{x\})|^2 = \lim_{\lambda \longrightarrow \infty} \frac{1}{2\lambda} \int_{-\lambda}^{\lambda} |\widehat{\mu}(\xi)|^2\, d\xi.$$

In particular, a necessary and sufficient condition for the continuity of μ is

$$\lim_{\lambda \longrightarrow \infty} \frac{1}{2\lambda} \int_{-\lambda}^{\lambda} |\widehat{\mu}(\xi)|^2\, d\xi = 0.$$

Proof. For a fixed $x \in \mathbb{R}$, let's consider

$$\varphi_\lambda(y) = \frac{1}{2\lambda} \int_{-\lambda}^{\lambda} e^{i\xi(x-y)}\, d\xi \quad \Longrightarrow \quad \sup_{y \in \mathbb{R}} |\varphi_\lambda(y)| \le 1 \quad \text{and} \quad \lim_{\lambda \longrightarrow 0} \varphi_\lambda(y) = 0 \quad \text{uniformly away from } x.$$

Let $\nu = \mu - \mu(\{x\})\delta_x$ then $\nu \in M(\mathbb{R})$ with $\nu(\{x\}) = 0$, which implies $|\nu|(\{x\}) = 0$. Regard ν as a linear functional acting on $\mathrm{BC}(\mathbb{R})$, by the dominated convergence theorem we have

$$\langle \varphi_\lambda(\cdot), \nu \rangle = \int_{\mathbb{R}} \varphi_\lambda(y)\, \overline{d\nu(y)} \longrightarrow 0 \qquad \text{as} \qquad \lambda \longrightarrow \infty. \tag{27}$$

By Fubini's theorem for $(\xi, y) \longmapsto e^{i\xi(x-y)} \chi_{[-\lambda,\lambda]}(\xi) \in L^1\Big((\mathbb{R}, m) \times (\mathbb{R}, \overline{d\mu})\Big)$ we obtain

$$\begin{aligned}
\langle \varphi_\lambda(\cdot), \nu \rangle &= \langle \varphi_\lambda(\cdot), \mu \rangle - \overline{\mu(\{x\})} \\
&= \int_{\mathbb{R}} \left(\frac{1}{2\lambda} \int_{-\lambda}^{\lambda} e^{i\xi(x-y)}\, d\xi \right) \overline{d\mu(y)} - \overline{\mu(\{x\})} \\
&= \frac{1}{2\lambda} \int_{-\lambda}^{\lambda} \left(\int_{\mathbb{R}} e^{-i\xi y}\, \overline{d\mu(y)} \right) e^{i\xi x}\, d\xi - \overline{\mu(\{x\})} = \frac{1}{2\lambda} \int_{-\lambda}^{\lambda} \overline{\widehat{\mu}(-\xi)} e^{i\xi x}\, d\xi - \overline{\mu(\{x\})}.
\end{aligned}$$

Thus we have

$$\overline{\langle \varphi_\lambda(\cdot), \nu \rangle} = \frac{1}{2\lambda}\int_{-\lambda}^{\lambda} \widehat{\mu}(-\xi)e^{-i\xi x}\,d\xi - \mu(\{x\}) = \frac{1}{2\lambda}\int_{-\lambda}^{\lambda} \widehat{\mu}(\xi)e^{i\xi x}\,d\xi - \mu(\{x\})$$

and together with (27) we obtain

$$\mu(\{x\}) = \lim_{\lambda\longrightarrow\infty} \frac{1}{2\lambda}\int_{-\lambda}^{\lambda} \widehat{\mu}(\xi)e^{i\xi x}\,d\xi.$$

Apply this formula for μ being replaced by $\mu * \mu^{\#}$ we obtain

$$\left(\mu * \mu^{\#}\right)(\{0\}) = \lim_{\lambda\longrightarrow\infty} \frac{1}{2\lambda}\int_{-\lambda}^{\lambda} |\widehat{\mu}(\xi)|^2\,d\xi = \sum_{x\in\mathbb{R}} |\mu(\{x\})|^2$$

and thus the proof is complete. □

As a consequence, since the Fejer's kernel $\mathcal{K}_\lambda \in L^1(\mathbb{R})$ and $\widehat{\mathcal{K}}_\lambda \in L^1(\widehat{\mathbb{R}})$, we can use the Parseval's formula to deduce that

$$\begin{aligned}
(\mathcal{K}_\lambda * \mu)(x) &= \int_{\mathbb{R}} \mathcal{K}_\lambda(x-y)\,d\mu(y) \\
&= \frac{1}{2\pi}\int_{\widehat{\mathbb{R}}} \mathcal{F}\Big[\mathcal{K}_\lambda(x-\cdot)\Big](\xi)\widehat{\mu}(-\xi)\,d\xi \\
&= \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\left(1-\frac{|\xi|}{\lambda}\right)\widehat{\mu}(-\xi)e^{-i\xi x}\,d\xi = \frac{1}{2\pi}\int_{-\lambda}^{\lambda}\left(1-\frac{|\xi|}{\lambda}\right)\widehat{\mu}(\xi)e^{i\xi x}\,d\xi.
\end{aligned}$$

Thus

$$\int_{\mathbb{R}} \lambda^{-1}\mathcal{K}_\lambda(x-y)\,d\mu(y) = \frac{1}{2\pi\lambda}\int_{-\lambda}^{\lambda}\left(1-\frac{|\xi|}{\lambda}\right)\widehat{\mu}(\xi)e^{i\xi x}\,d\xi.$$

By the same argument, since $\lambda^{-1}\mathcal{K}_\lambda(x-y) \longrightarrow 0$ uniformly on compact set away from x, and $\lambda^{-1}\mathcal{K}_\lambda(0) = \frac{1}{2\pi}$, send $\lambda \longrightarrow \infty$ we obtain

$$\frac{1}{2\pi}\mu(\{x\}) = \lim_{\lambda\longrightarrow\infty}\frac{1}{2\pi\lambda}\int_{-\lambda}^{\lambda}\left(1-\frac{|\xi|}{\lambda}\right)\widehat{\mu}(\xi)e^{i\xi x}\,d\xi \qquad \Longrightarrow \qquad \mu(\{x\}) = \lim_{\lambda\longrightarrow\infty}\frac{1}{\lambda}\int_{-\lambda}^{\lambda}\left(1-\frac{|\xi|}{\lambda}\right)\widehat{\mu}(\xi)e^{i\xi x}\,d\xi$$

Together with the above theorem, we obtain:

Theorem 2.18. *Let $\mu \in M(\mathbb{R})$, then*

$$\mu(\{x\}) = \lim_{\lambda\longrightarrow\infty}\frac{1}{\lambda}\int_{-\lambda}^{\lambda}\left(1-\frac{|\xi|}{\lambda}\right)\widehat{\mu}(\xi)e^{i\xi x}\,d\xi = \lim_{\lambda\longrightarrow\infty}\frac{1}{\lambda^2}\int_{-\lambda}^{\lambda}|\xi|\widehat{\mu}(\xi)e^{i\xi x}\,d\xi.$$

2.2 Fourier transforms of distributions

We will recall general fact about Schwartz functions in high dimensions even though later on we will only focus on the real line. We denote by $\mathcal{D}(\Omega)$ the space of smooth functions $C_c^\infty(\Omega)$ with compact support in Ω. A distribution in $\mathcal{D}'(\Omega)$ is a linear functional on $\mathcal{D}(\Omega)$, equipped with the weak* topology.

1. (**Schwartz space**) For any non-negative integer N, any multi-index α and $f : \mathbb{R}^n \longrightarrow \mathbb{C}$ we define:

$$\|f\|_{(N,\alpha)} = \sup_{x\in\mathbb{R}^n}(1+|x|)^N\,|\partial^\alpha f(x)| \qquad \text{and} \qquad \mathcal{S}(\mathbb{R}^n) = \Big\{f \in C^\infty(\mathbb{R}^n,\mathbb{C}) : \|f\|_{(N,\alpha)} < \infty \text{ for all } N,\alpha\Big\}.$$

The is a metrizable topology on $\mathcal{S}(\mathbb{R}^n)$ which makes $\mathcal{S}(\mathbb{R}^n)$ a Fréchet space over $\mathbb{C}$.

Lemma 2.19. *If $f \in \mathcal{S}(\mathbb{R}^n)$ then $\partial^\alpha f \in L^p(\mathbb{R}^n)$ for all α and all $p \in [1,\infty]$. Indeed $\partial^\alpha f \in C_0(\mathbb{R}^n)$ for all multi-index α, consequently $\partial^\alpha f$ is uniformly continuous.*

Proof. For any multi-index α, $p \in (1,\infty)$ and $N \in \mathbb{N}$ then

$$|\partial^\alpha f(x)| \le \frac{\|f\|_{N,\alpha}}{(1+|x|)^N} \qquad \Longrightarrow \qquad \int_{\mathbb{R}^n} |\partial^\alpha f(x)|^p \, dx \le \|f\|_{N,\alpha}^p \int_{\mathbb{R}^n} \frac{dx}{(1+|x|)^{Np}} < \infty$$

if we choose $N > \frac{n}{p}$. The case $p = \infty$ is trivial since $(1+|x|) \ge 1$. For $\varepsilon > 0$, we have

$$\{x \in \mathbb{R}^n : |\partial^\alpha f(x)| \ge \varepsilon\} \subset \left\{ x \in \mathbb{R}^n : \frac{\|f\|_{(1,\alpha)}}{1+|x|} \ge \varepsilon \right\} = \left\{ x \in \mathbb{R}^n : |x| \le \varepsilon^{-1} \|f\|_{(1,\alpha)} - 1 \right\}.$$

The latter set is compact in $\mathbb{R}^n$, which implies the result. □

Proposition 2.20. *$\mathcal{S}$ is a Fréchet space with the topology defined by the norms $\|\cdot\|_{(N,\alpha)}$.*

Proof. The topology on $\mathcal{S}$ is generated by a countable sequence of seminorms, thus it is locally convex topological space, Hausdorff and metrizable with an translation-invariant metric. The only nontrivial point we need to check is completeness. If $\{f_k\}$ is a Cauchy sequence in $\mathcal{S}$ then $\|f_k - k_j\|_{(N,\alpha)} \longrightarrow 0$ as $j,k \longrightarrow \infty$ for all N, α. In particular, for each fixed α, take $N = 0$ then

$$\sup_{x\mathbb{R}^n} \left| \partial^\alpha f_j(x) - \partial^\alpha f_k(x) \right| = \|f_j - f_k\|_{(0,\alpha)} \longrightarrow 0$$

as $j,k \longrightarrow \infty$, thus for each $x \in \mathbb{R}^n$ we have $\{\partial^\alpha f_k(x)\}_{k=1}^\infty$ is a Cauchy sequence in $\mathbb{C}$, thus it defines $g_\alpha(x) = \lim_{k\longrightarrow\infty} \partial^\alpha f_k(x)$. To each $\varepsilon > 0$ corresponds $N(\varepsilon) \in \mathbb{N}$ such that

$$\left| \partial^\alpha f_k(x) - \partial^\alpha f_j(x) \right| < \varepsilon \qquad \text{for all} \qquad k, j \ge N(\varepsilon), x \in \mathbb{R}^n.$$

Let $j \longrightarrow \infty$, we deduce that

$$|\partial^\alpha f_k(x) - g_\alpha(x)| < \varepsilon \qquad \text{for all} \qquad k \ge N(\varepsilon), x \in \mathbb{R}^n.$$

In other words $\partial^\alpha f_k \longrightarrow g_\alpha$ uniformly on $\mathbb{R}^n$. Since $\partial^\alpha f_k \in \mathcal{S}(\mathbb{R}^n) \subset C_0(\mathbb{R}^n)$ and $C_0(\mathbb{R}^n)$ is closed in $\mathrm{BC}(\mathbb{R}^n)$ with the uniform metric, it is obvious that $g \in C_0(\mathbb{R}^n)$. Denoting by e_j the vector $(0,\ldots,1,\ldots,0)$ with the 1 in the jth position, we have

$$f_k(x+te_j) - f_k(x) = \int_0^t \partial^{e_j} f_k(x+se_j)\, ds \qquad \Longrightarrow \qquad g_0(x+te_j) - g_0(x) = \int_0^t g_{e_j}(x+se_j)\, ds$$

by letting $k \longrightarrow \infty$. The fundamental theorem of calculus implies that $g_{e_j} = \partial^{e_j} g_0$, and an induction on $|\alpha|$ then yields $g_\alpha = \partial^\alpha g_0$ for all α, thus $g \in \mathcal{S}(\mathbb{R}^n)$ follows easily. Finally recall that

$$\sup_{x\in\mathbb{R}^n} (1+|x|)^N \left| \partial^\alpha f_j(x) - \partial^\alpha g_0(x) \right| \le \|f_j - f_k\|_{(N,\alpha)} + \sup_{x\in\mathbb{R}^n} (1+|x|)^N |\partial^\alpha f_k(x) - \partial^\alpha g_0(x)|$$

for all $k \in \mathbb{N}$. For each $\varepsilon > 0$, choose j,k large such that $\|f_j - f_k\|_{(N,\alpha)} < \frac{\varepsilon}{2}$, then let $k \longrightarrow \infty$ we obtain $\|f_k - g_0\|_{(N,\alpha)} \longrightarrow 0$. □

Another useful characterization of $\mathcal{S}$ is the following.

Proposition 2.21. *If $f \in C^\infty(\mathbb{R}^n)$, then $f \in \mathcal{S}$ iff $x^\beta \partial^\alpha f$ is bounded for all multi-indices α, β iff $\partial^\alpha(x^\beta f)$ is bounded for all multi-indices α, β.*

Proof. Obviously $|x^\beta| \le (1+|x|)^N$ for $|\beta| \le N$. On the other hand, $\sum_{j=1}^n |x_j|^N$ is strictly positive on the unit sphere $|x| = 1$, so it has a positive minimum δ there (note that $\delta \le 1$). If follows that $\sum_{j=1}^n |x_j|^N \ge \delta |x|^N$ for all x since both sides are homogeneous of degree N, and hence

$$(1+|x|)^N \le 2^N \left(1 + |x|^N\right) \le 2^N \left[1 + \delta^{-1} \sum_{j=1}^n |x_j|^N \right] \le 2^N \delta^{-1} \sum_{|\beta|\le N} |x^\beta|.$$

This establishes the first equivalence. The second one follows from the fact that each $\partial^\alpha(x^\beta f)$ is a linear combination of terms of the from $x^\gamma \partial^\delta f$ and vice versa by the product rule. □

2. A very useful fact we need later is that $C_c^\infty(\Omega)$ is dense in $L^p(\Omega)$ for any open set $\Omega \subset \mathbb{R}^n$ and any $1 \le p < \infty$. Since $C_c^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n) \subset L^p(\mathbb{R}^n)$ for all $1 \le p \le \infty$, we deduce that $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^p(\mathbb{R}^n)$ for $1 \le p < \infty$. The result fails for L^∞, however recall that L^∞ is the dual space of L^1 (with respect to norms on these spaces), if we denote $\sigma(L^\infty, L^1)$ to be the weak* topology on L^∞, then we have:

   Theorem 2.22. *The Schwartz space $\mathcal{S}(\mathbb{R}^n)$ is dense in $L^\infty(\mathbb{R}^n)$ with respect to the weak* topology $\sigma(L^\infty, L^1)$. That is, for any $f \in L^\infty(\mathbb{R}^n)$, there exists a sequence of $f_n \in \mathcal{S}(\mathbb{R}^n)$ such that*

$$\lim_{n\longrightarrow\infty} \int_{\mathbb{R}^n} f_n(x)\,\phi(x)\,dx = \int_{\mathbb{R}^n} f(x)\phi(x)\,dx \qquad \textit{for every} \qquad \phi \in L^1(\mathbb{R}^n).$$

   Proof. Take $\eta \in C^\infty(\mathbb{R}^n)$ with $0 \le \eta \le 1$ and $\|\eta\|_{L^1} = 1$, define $\eta_n(x) = n\eta(nx)$ for all $n = 1, 2, \ldots$. For each n let $\xi_n \in C_c^\infty(\mathbb{R}^n)$ such that $0 \le \xi \le 1$, supp $\xi_n \subset B(0, n+1)$ and $\xi \equiv 1$ on $B(0, n)$. Let's

$$f_n = (f * \eta_n)\xi_n \qquad \text{satisfies} \qquad \|f_n\|_{L^\infty} \le \|f * \eta_n\|_{L^\infty} \le \|f\|_{L^\infty}\|\eta_n\|_{L^1} = \|f\|_{L^\infty}.$$

   It is easy to see that since $f \in L^\infty(\mathbb{R}^n)$, $f * \eta_n \in C^\infty(\mathbb{R})$ with $D^\alpha(f * \eta_n) = f * D^\alpha \eta_n$ is bounded for all multi-index α, hence $f_n \in \mathcal{S}(\mathbb{R}^n)$. Since $f * \eta_n \longrightarrow f$ a.e as $n \longrightarrow \infty$, we have $f_n \longrightarrow f$ a.e as $n \longrightarrow \infty$ as well. Now for any $\phi \in L^1(\mathbb{R}^n)$ we have $f_n\phi \longrightarrow f\phi$ a.e and $|f_n\phi| \le \|f\|_1\phi \in L^1(\mathbb{R}^n)$, thus the dominated convergence theorem implies

$$\lim_{n\longrightarrow\infty} \int_{\mathbb{R}^n} f_n(x)\phi(x)\,dx = \int_{\mathbb{R}^n} f(x)\phi(x)\,dx$$

   and the proof is complete. □

3. A tempered distribution on $\mathbb{R}^n$ is a continuous linear functional on $\mathcal{S}(\mathbb{R}^n)$, denoted by $\mathcal{S}'(\mathbb{R}^n) = \mathcal{L}(\mathcal{S}(\mathbb{R}^n), \mathbb{C})$. It comes equipped with the weak* topology, that is, the topology of point-wise convergence in $\mathcal{S}$. By using Hölder inequality with $\mathcal{S} \subset L^p(\mathbb{R}^n)$ for all $1 \le p \le \infty$, every locally L^p function can be identified with a tempered distribution in $\mathcal{S}'(\mathbb{R}^n)$, by the natural pairing

$$f \in L^p_{\text{loc}}(\mathbb{R}^n) \longmapsto \Lambda_f \quad \text{with} \quad \langle \phi, \Lambda_f \rangle = \int_{\mathbb{R}^n} \phi(x)\overline{f(x)}\,dx.$$

   Thus $L^p(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$ for all $1 \le p \le \infty$ (after identifying as distribution). The same is true for the space of finite Borel measure $M(\mathbb{R}^n)$ as well, by the pairing:

$$\mu \in M(\mathbb{R}^n) \longmapsto \Lambda_\mu \qquad \text{with} \qquad \langle \phi, \Lambda_\mu \rangle = \int_{\mathbb{R}^n} \phi(x)\overline{d\mu(x)}.$$

4. For a tempered distribution $\mu \in \mathcal{S}'(\mathbb{R}^n)$, we define its Fourier transform $\widehat{\mu} \in \mathcal{S}'(\widehat{\mathbb{R}}^n)$ by

$$\langle \widehat{\varphi}, \widehat{\mu} \rangle = \langle \varphi, \mu \rangle \qquad \text{for all} \qquad \varphi \in \mathcal{S}(\mathbb{R}^n).$$

   We denote by $\mathcal{F}L^p = \mathcal{F}[L^p(\mathbb{R}^n)]$.

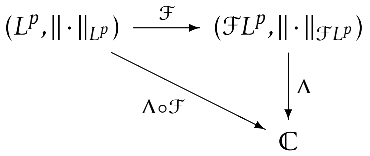
   Theorem 2.23. *For $f \in \mathcal{F}L^p$, let's define $\|\widehat{f}\|_{\mathcal{F}L^p} = \|f\|_{L^p}$, then it defines $(\mathcal{F}L^p, \|\cdot\|_{\mathcal{F}L^p})$ as a Banach space for all $1 \le p \le \infty$.*

   Proof. Since each distribution $\widehat{f}$ acting as $\Lambda_{\widehat{f}}$ on $\mathcal{S}(\mathbb{R}^n)$ by $\left\langle \widehat{\phi}, \widehat{f} \right\rangle = \langle \phi, f \rangle = \int_{\mathbb{R}^n} \phi(x)\overline{f(x)}\,dx$, it is clear that $\|\widehat{f}\|_{\mathcal{F}L^p} = 0$ iff $f = 0$ a.e, which implies $\left\langle \widehat{\phi}, \widehat{f} \right\rangle$ for all $\phi \in \mathcal{S}(\widehat{\mathbb{R}}^n)$, i.e., $\widehat{f} = 0$. Thus $(\mathcal{F}L^p, \|\cdot\|_{\mathcal{F}L^p})$ is a normed space and its completeness follows from the completeness of $L^p(\mathbb{R}^n)$ for all $1 \le p \le \infty$. □

   Recall that for $1 \le p < \infty$, L^{p^*} is the dual space of L^p where $\frac{1}{p} + \frac{1}{p^*} = 1$, notably $\left(L^1\right)^* = L^\infty$ but not the inverse. The same things hold true for $\mathcal{F}L^p$ as well.

   Theorem 2.24. *For $1 \le p < \infty$, $\left(\mathcal{F}L^{p^*}, \|\cdot\|_{\mathcal{F}L^{p^*}}\right)$ is the dual space of $(\mathcal{F}L^p, \|\cdot\|_{\mathcal{F}L^p})$ where $\frac{1}{p} + \frac{1}{p^*} = 1$.*

Proof. Let $\Lambda : (\mathcal{F}L^p, \|\cdot\|_{\mathcal{F}L^p}) \longrightarrow \mathbb{C}$ be a linear bounded functional, we have the following diagram:



$$\Longrightarrow \qquad \Lambda \circ \mathcal{F} \in (L^p)^*.$$

Using the duality of L^p when $1 \le p < \infty$, there exists a unique $f_\Lambda \in L^{p^*}$ with $1 < p^* \le \infty$ such that

$$\Lambda(\widehat{\varphi}) = \int_{\mathbb{R}^n} \varphi(x)\overline{f_\Lambda(x)}\, dx \qquad \text{for every} \qquad \varphi \in L^p(\mathbb{R}^n)$$

and furthermore $\|f_\Lambda\|_{L^{p^*}} = \|\Lambda \circ \mathcal{F}\|_{(L^p)^*}$. By an application of Hahn-Banach theorem we have

$$\|\Lambda\|_{(\mathcal{F}L^p)^*} = \sup_{\|\widehat{\varphi}\|_{\mathcal{F}L^p}=1} \left|\Lambda\widehat{\varphi}\right| = \sup_{\|\varphi\|_{L^p}=1} \left|(\Lambda \circ \mathcal{F})\varphi\right| = \|\Lambda \circ \mathcal{F}\|_{(L^p)^*} = \|f_\Lambda\|_{L^{p^*}}.$$

Thus the mapping $\Phi : (\mathcal{F}L^p)^* \longrightarrow \mathcal{F}L^{p^*}$ maps $\Lambda \longmapsto \widehat{f_\Lambda}$ is a linear isometry. We only need to show that Φ is surjective, indeed for any $f \in L^{p^*}$, we can define $\Lambda : \mathcal{F}L^p \longrightarrow \mathbb{C}$ by

$$\Lambda(\widehat{\varphi}) = \int_{\mathbb{R}} \varphi(x)\overline{f(x)}\, dx$$

It is clear that Λ is linear and bounded since

$$|\Lambda(\widehat{\varphi})| \le \int_{\mathbb{R}} |\phi(x)\overline{f(x)}|\, dx \le \|f\|_{L^{p^*}}\|\varphi\|_{L^p} = \|\widehat{f}\|_{\mathcal{F}L^{p^*}}\|\widehat{\varphi}\|_{\mathcal{F}L^p} \quad \Longrightarrow \quad \|\Lambda\|_{(\mathcal{F}L^p)^*} \le \|\widehat{f}\|_{\mathcal{F}L^{p^*}}.$$

By the duality of $(L^p)^* = L^{p^*}$ (the uniqueness part) we obtain $\Phi(\Lambda) = \widehat{f}$. Thus Φ is a linear surjective isometry, thus it is an isomorphism between two spaces and the proof is complete. □

From this result, by pairing in $\mathcal{F}L^{p^*}$ and $\mathcal{F}L^p$ we means (for $1 \le p < \infty$)

$$\boxed{\left\langle \widehat{\varphi}, \widehat{f} \right\rangle_{(\mathcal{F}L^p, \mathcal{F}L^{p^*})} = \int_{\mathbb{R}^n} f(x)\overline{\varphi(x)}\, dx.}$$

5. (**Support of distributions**) Suppose $\Lambda \in \mathcal{D}'(\Omega)$, if $\mathcal{O}$ is an open subset of Ω and if $\Lambda\phi = 0$ for every $\phi \in \mathcal{D}(\mathcal{O})$, we say that Λ vanishes in $\mathcal{O}$. Let W be the union of all open sets $\mathcal{O} \subset \Omega$ in which Λ vanishes, we define the complement $\Omega\backslash W$ to be the support of Λ. In case $\Omega = \mathbb{R}^n$ and Λ is a tempered distribution, we can extend the notion of support in the same way:

Definition 2.25. *A tempered distribution $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ vanishes on an open set $\mathcal{O} \subset \mathbb{R}^n$, if $\langle \phi, \Lambda \rangle = 0$ for all $\phi \in \mathcal{S}(\mathbb{R}^n)$ with compact support contained in $\mathcal{O}$.*

Proposition 2.26. *If $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ vanishes on $\{\mathcal{O}\}_{\alpha\in A}$ where $\mathcal{O}_\alpha$ is open, then Λ vanishes on $\bigcup_{\alpha\in A} \mathcal{O}_\alpha$.*

Proof. Let $\Gamma = \{\mathcal{O}_\alpha\}_{\alpha\in A}$ where Λ vanishes in $\mathcal{O}_\alpha$. Let $\{\varphi_j\}_{j=1}^\infty$ be a partition of unity subordinate to $\{\mathcal{O}_\alpha\}_{\alpha\in A}$. If $f \in \mathcal{D}(W)$ then supp f only intersects with finitely many $\mathcal{O}_{\alpha_j}$, i.e., $f = \sum_{j=1}^\infty f\varphi_j$ where only finitely many terms of this sum are different from 0. Hence $\Lambda f = \sum_{\text{finite}} \Lambda(f\varphi_j) = 0$ since $\text{supp}(f\varphi_j) \subset \mathcal{O}_{\alpha_j}$. □

Thus we can define the support supp(Λ) of $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ is the complement of the largest open set $\mathcal{O} \subset \mathbb{R}^n$ on which Λ vanishes. In other words, $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ can be viewed as a distribution on $\mathcal{D}(\mathbb{R}^n)$ simply by its restriction to $\mathcal{D}(\mathbb{R}^n)$, where its support is already well-defined, and then take:

$$\text{supp}\,(\Lambda) := \text{supp}\left(\Lambda|_{\mathcal{D}(\mathbb{R}^n)}\right).$$

The definition of supp Λ implies that if $\phi \in \mathcal{S}(\mathbb{R}^n)$ with supp ϕ is compact and supp $\phi \cap$ supp $\Lambda = \emptyset$ then $\langle \phi, \Lambda \rangle = 0$. The same still true even if ϕ doesn't have compact support.

Proposition 2.27. *If* $\phi \in \mathcal{S}(\mathbb{R}^n)$ *has* $\operatorname{supp}\phi \cap \operatorname{supp}\Lambda = \emptyset$ *then* $\langle \Lambda, \phi \rangle = 0$.

Proof. Let $\eta \in C_c^\infty(\mathbb{R}^n)$ with $\operatorname{supp}\eta \subset B(0,1)$, $\|\eta\|_{L^1} = 1$ and $0 \le \eta \le 1$ on $\mathbb{R}^n$ with $\eta(0) = 1$. We claim that

$$\eta(\lambda x)\phi(x) \longrightarrow \phi(x) \qquad \text{in} \qquad \mathcal{S}(\mathbb{R}^n) \quad \text{as} \quad \lambda \longrightarrow 0.$$

In deed, for any $N \in \mathbb{N}$ and $\alpha \neq 0$ we have

$$\begin{aligned}
\|\eta(\lambda x)\phi(x) - \phi(x)\|_{N,\alpha} &= \sup_{x\in\mathbb{R}^n}(1+|x|)^N \left|\partial^\alpha\left(\phi(x)(\eta(\lambda x)-1)\right)\right| \\
&= \sup_{x\in\mathbb{R}^n}(1+|x|)^N \left|\sum_{\beta\le\alpha} c_{\alpha\beta} D^{\alpha-\beta}\phi(x) D^\beta\Big(\eta(\lambda x)-1\Big)\right| \\
&\le \sum_{\beta\le\alpha}\left(c_{\alpha\beta}\|\phi\|_{N,\alpha-\beta}\right)\lambda^{|\beta|}\left(\sup_{x\in\mathbb{R}^n}\left|\left(D^\beta\eta\right)(\lambda x)\right|\right) \\
&\le \sum_{\beta\le\alpha}\left(c_{\alpha\beta}\|\phi\|_{N,\alpha}\right)\lambda^{|\beta|}\|\eta\|_{0,\beta} \longrightarrow 0 \qquad \text{as} \qquad \lambda \longrightarrow 0.
\end{aligned}$$

The case $\alpha = 0$ we have

$$\|\eta(\lambda x)\phi(x) - \phi(x)\|_{N,0} = \sup_{x\in\mathbb{R}^n}(1+|x|)^N \left|(\phi(x)(\eta(\lambda x)-1))\right| \le \|\phi\|_{N,0}\left(\sup_{x\in\mathbb{R}^n}\left|\eta(\lambda x)-1\right|\right).$$

By the fundamental theorem of calculus, for any $x \in \mathbb{R}^n$, let $\gamma(t) = t(\lambda x)$, we have

$$\begin{aligned}
\eta(\lambda x) - 1 &= \eta(\lambda x) - \eta(0) \\
&= (\eta_\circ\gamma)(1) - (\eta_\circ\gamma)(0) = \int_0^1 \nabla\eta(\gamma(s))\cdot\gamma'(s)\,ds
\end{aligned}$$

and thus

$$\left|\eta(\lambda x) - 1\right| \le \int_0^1 \|\nabla\eta\|_{L^\infty}|\gamma'(s)|\,ds \le \frac{\lambda}{2}\|\nabla\eta\|_{L^\infty} \longrightarrow 0$$

uniformly in x as $\lambda \longrightarrow 0$. Thus $\eta(\lambda x)\phi(x) \longrightarrow \phi(x)$ in $\mathcal{S}(\mathbb{R}^n)$ as $\lambda \longrightarrow 0$. Clearly $\operatorname{supp}(\eta(\lambda\cdot)\phi) \subset \operatorname{supp}\phi \subset \mathbb{R}^n \backslash \operatorname{supp}(\Lambda)$, thus $\langle \Lambda, \eta(\lambda\cdot)\phi\rangle = 0$ for all $\lambda > 0$ and hence as a linear functional on $\mathcal{S}(\mathbb{R})$ we have $\langle \phi, \Lambda\rangle = \lim_{\lambda\longrightarrow\infty}\langle \Lambda, \eta(\lambda\cdot)\phi\rangle = 0$. ☐

Let's consider $n = 1$ from now on.

6. So far, we have studied that:
   (a) If Λ is (identified with) a summable function then $\widehat{\Lambda}$ is (identified with) a function in $C_0(\mathbb{R})$ by Riemann-Lebesgue lemma.
   (b) If If Λ is (identified with) a finite complex Radon measure in $M(\mathbb{R})$ then $\widehat{\Lambda}$ is (identified with) a uniformly continuous bounded function.
   (c) If Λ is (identified with) a $L^p(\mathbb{R}^n)$ function with $1 \le p \le 2$ then $\widehat{\Lambda}$ is (identified with) a function in $L^q(\widehat{\mathbb{R}})$ with $p^{-1} + q^{-1} = 1$.

   Let's pay attention to $\mathcal{F}L^\infty$ as the dual space of $\mathcal{F}L^1$. Every function $f \in L^1(\mathbb{R})$ corresponds to a tempered distribution $\widehat{f} \in \mathcal{S}'(\widehat{\mathbb{R}})$ acting by

$$\widehat{f} : \mathcal{S}(\widehat{\mathbb{R}}) \longrightarrow \mathbb{C} \qquad \text{maps} \qquad \widehat{\varphi} \longmapsto \left\langle \widehat{\varphi}, \widehat{f}\right\rangle = \langle \varphi, f\rangle = \int_{\mathbb{R}} \phi(x)\overline{f(x)}\,dx.$$

   On the other hand, $\widehat{f}$ can be seen as a linear function acting on $\mathcal{F}L^1$ by

$$\widehat{f} : \mathcal{F}L^1 \longrightarrow \mathbb{C} \qquad \text{maps} \qquad \widehat{\phi} \longmapsto \left\langle \widehat{\phi}, \widehat{f}\right\rangle_{\mathcal{F}L^1,\mathcal{F}L^\infty} = \langle \phi, f\rangle_{L^1,L^\infty} = \int_{\mathbb{R}} \phi(x)\overline{f(x)}\,dx.$$

   Their actions are identical on $\mathcal{S}(\widehat{\mathbb{R}})$, thus we have extended $\widehat{f}$ as a linear function on $\mathcal{S}(\widehat{\mathbb{R}})$ onto a larger space $\mathcal{F}L^1$.

7. We will extend the notion of support to linear functional $\mathcal{F}L^\infty$. Let's recall the standard construction of mollifiers, we pick a function $\eta \in C_c^\infty(\mathbb{R})$ with supp $\eta \subset [-1,1]$, $\|\eta\|_{L^1} = 1$ and $0 \leq \eta \leq 1$ on $\mathbb{R}$, we then define

$$\eta_n(x) = n\eta(nx) \quad \text{has} \quad \eta_n \in C_c^\infty(\mathbb{R}), \quad \text{supp}\, \eta_n \subset \left[-\frac{1}{n}, \frac{1}{n}\right], \quad \|\eta_n\|_{L^1} = 1.$$

Some facts about $\{\eta_n\}_{n=1}^\infty$:

- In the sense of (tempered or $\mathcal{D}$) distribution, we have $\eta_n \longrightarrow \delta_0$ in $\mathcal{S}'(\mathbb{R})$ or $\mathcal{D}'(\mathbb{R})$ as $n \longrightarrow \infty$.
- If $\psi \in \mathcal{D}(\mathbb{R}^n)$ then $\psi * \eta_n \longrightarrow \psi$ in $\mathcal{D}(\mathbb{R})$ as $n \longrightarrow \infty$.
- If $\Lambda \in \mathcal{D}'(\mathbb{R})$ then $\Lambda * \eta_n \longrightarrow \Lambda$ in $\mathcal{D}'(\mathbb{R})$ as $n \longrightarrow \infty$.

Since $\{\eta_n\}_{n=1}^\infty \subset \mathcal{D}(\mathbb{R}) \subset \mathcal{S}(\mathbb{R})$, we have $\{\widehat{\eta_n}\}_{n=1}^\infty \subset \mathcal{S}(\mathbb{R})$ as well, and it is obvious that $\widehat{\eta_n} \longrightarrow \widehat{\delta_0} = 1$ in $\mathcal{D}'(\mathbb{R})$. Furthermore

Proposition 2.28. *We have* $\widehat{\eta_n} \longrightarrow \widehat{\delta_0} = 1$ *a.e.*

Lemma 2.29. *Let* $\eta \in L^1(\mathbb{R}^n)$ *with* $c = \int_{\mathbb{R}^n} \eta(x)\, dx$, *then for any* $g \in L^1(\mathbb{R}^n)$. *For* $\lambda > 0$ *let's define* $\eta_\lambda(x) = \frac{1}{\lambda^n}\eta(\frac{x}{\lambda})$ *then* $g * \eta_\lambda \longrightarrow cg$ *in* $L^1(\mathbb{R}^n)$ *as* $\lambda \longrightarrow 0$.

Proof. Obviously $\int_{\mathbb{R}^n} \eta_\lambda(x)\, dx = \int_{\mathbb{R}^n} \eta(x)\, dx$. For each $x \in \mathbb{R}^n$ we have

$$(g * \eta_\lambda)(x) - cg(x) = \int_{\mathbb{R}^n} \Big(g(x-y) - g(x)\Big)\eta_\lambda(y)\, dy = \int_{\mathbb{R}^n} \Big(g(x-\lambda y) - g(x)\Big)\eta(y)\, dy$$

and thus by using Fubini's theorem

$$\begin{aligned}\int_{\mathbb{R}^n} \Big|(g * \eta_\lambda)(x) - cg(x)\Big|\, dx &\leq \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |g(x-\lambda y) - g(x)|\, dx\right)|\eta(y)|\, dy \\ &\leq \int_{\mathbb{R}^n} \|\tau_{\lambda y} g - g\|_{L^1} |\eta(y)|\, dy \longrightarrow 0\end{aligned}$$

as $\lambda \longrightarrow 0$ by dominated convergence theorem. □

Theorem 2.30. *If* $\varphi \in L^1(\mathbb{R})$ *with* $\text{supp}(\widehat{\varphi}) \cap \text{supp}(\widehat{f}) = \emptyset$ *then*

$$\left\langle \widehat{\varphi}, \widehat{f} \right\rangle_{\mathcal{F}L^1, \mathcal{F}L^\infty} = 0.$$

8. $\mathcal{S}(\mathbb{R})$ is an algebra under point-wise multiplication, therefore we can define the product $\varphi\nu$ of a function $\varphi \in \mathcal{S}(\mathbb{R})$ and a distribution $\nu \in \mathcal{S}'(\mathbb{R})$ by

$$\langle \phi, \varphi\nu \rangle = \langle \phi\overline{\varphi}, \nu \rangle \qquad \text{for all} \qquad \varphi \in \mathcal{S}(\mathbb{R}).$$

It is clear that supp $(\varphi\nu) \subset$ supp $\varphi \cap$ supp ν.

2.3 Pseudo-measures

1. If $\mu \in M(\widehat{\mathbb{R}})$, it can be identified with $\mathcal{F}L^\infty$ by setting $f(x) = \widehat{\mu}(-x) \in L^\infty(\mathbb{R})$ then by Parseval's formula and the fact that $\widehat{\widehat{\varphi}}(x) = 2\pi\varphi(-x)$ for $\varphi \in \mathcal{S}(\mathbb{R})$ we have

$$\langle \widehat{\varphi}, \mu \rangle = \int_{\widehat{\mathbb{R}}} \widehat{\varphi}(\xi)\overline{d\mu(\xi)} = \int_{\mathbb{R}} \varphi(-x)\overline{\widehat{\mu}(x)}\, dx = \int_{\mathbb{R}} \varphi(x)\overline{f(x)}\, dx$$

for any $\varphi \in \mathcal{S}(\mathbb{R})$, note that $\widehat{\varphi} \in C_0(\widehat{\mathbb{R}})$ by Riemann-Lebesgue lemma. Thus $\mu = \widehat{f} \in \mathcal{F}L^\infty$ as distributions (recall that $\mathcal{F} : \mathcal{S} \longrightarrow \mathcal{S}$ is an is isomorphism). Thus we can conclude that $\boxed{M(\widehat{\mathbb{R}}) \subset \mathcal{F}L^\infty.}$

The elements of $\mathcal{F}L^\infty$ are commonly referred to as pseudo-measures. Note that $M(\widehat{\mathbb{R}})$ is a relatively small part of $\mathcal{F}L^\infty$; for instance, if $\varphi \in L^\infty(\mathbb{R})$ is not uniformly continuous then $\widehat{\varphi}$ cannot be a measure.

2. (**Convolutions of pseudo-measures**) We take something which we have proved for measures, as a definition for the larger class of pseudo-measures.

Definition 2.31. *If $f,g \in L^\infty(\mathbb{R})$, then $\widehat{f},\widehat{g} \in \mathcal{F}L^\infty \subset \mathcal{S}'(\widehat{\mathbb{R}})$ as tempered distributions. We define the convolution $\widehat{f} * \widehat{g}$ of pseudo-measures $\widehat{f}$ and $\widehat{g}$ to be the pseudo-measure $\widehat{fg}$, namely*

$$\boxed{\widehat{f} * \widehat{g} = \widehat{fg}.}$$

Note that this definition is consistent, in the sense that if $\widehat{f},\widehat{g}$ happen to be measures, then $\widehat{f} * \widehat{g}$ is their (measure theoretic) convolution. Indeed if $f,g \in L^\infty(\mathbb{R}^n)$ such that $\widehat{f} = \mu$ and $\widehat{g} = \nu$, then $f(x) = \widehat{\mu}(-x)$ and $g(x) = \widehat{\nu}(-x)$. Recall that $\widehat{\mu * \nu}(x) = \widehat{\mu}(x)\widehat{\nu}(x)$, by Parseval's formula and $\mathcal{F}^2\varphi(x) = 2\pi\varphi(-x)$ for $\varphi \in \mathcal{S}(\mathbb{R})$ we have

$$\langle \widehat{\varphi}, \mu * \nu \rangle = \int_{\widehat{\mathbb{R}}} \widehat{\varphi}(\xi) \overline{d(\mu * \nu)} = \int_{\mathbb{R}} \varphi(-x) \overline{\widehat{\mu}(x)\widehat{\nu}(x)}\, dx = \int_{\mathbb{R}} \varphi(x) \overline{(fg)(x)}\, dx = \langle \varphi, fg \rangle = \langle \widehat{\varphi}, \widehat{fg} \rangle.$$

3. We will extend some results about supports of distribution to $\mathcal{F}L^\infty$. First of all, if $f \in L^\infty(\mathbb{R})$ and $g \in \mathcal{S}(\mathbb{R})$, then $\widehat{g} \in \mathcal{S}(\widehat{\mathbb{R}})$ and thus the convolution $\widehat{f} * \widehat{g}$ should be identical to the old definition of convolution between distribution and test function (as a function)

$$\left(\widehat{f} * \widehat{g}\right)(\xi) = \left\langle \widehat{f}, \tau_\xi \mathcal{R}\widehat{g} \right\rangle.$$

It follows from $(\Lambda * \phi)(x) = \langle \Lambda, \tau_x \tilde{\phi} \rangle$. We check that this actually agrees with our definition above.

Proposition 2.32. *Let $f \in L^\infty(\mathbb{R})$ and $g \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ then $fg \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$, thus we have $\widehat{f} * \widehat{g} = \mathcal{F}(fg)$ is a function in $C_0(\mathbb{R})$, thus it makes sense to talk about its value at one point, and*

$$\left(\widehat{f} * \widehat{g}\right)(\xi) = \left\langle \widehat{f}, \tau_\xi \mathcal{R}\widehat{g} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1}.$$

Proof. If $f \in \mathcal{S}(\mathbb{R})$ first, we can use the Fourier inversion formula for $f(x)$ to deduce that

$$\begin{aligned}
\widehat{fg}(\xi) &= \int_{\mathbb{R}} f(x) g(x) e^{-2\pi i \xi x}\, dx \\
&= \int_{\mathbb{R}} \left(\int_{\widehat{\mathbb{R}}} \widehat{f}(\eta) e^{2\pi i \eta x}\, d\eta \right) g(x) e^{-2\pi i \xi x}\, dx \\
&= \int_{\widehat{\mathbb{R}}} \widehat{f}(\eta) \left(\int_{\mathbb{R}} g(x) e^{-2\pi i (\xi - \eta) x}\, dx \right) d\eta = \int_{\widehat{\mathbb{R}}} \widehat{f}(\eta) \widehat{g}(\xi - \eta)\, d\eta = \left\langle \widehat{f}, \tau_\xi \mathcal{R}\widehat{g} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1}.
\end{aligned}$$

Now using theorem 2.22 we can extend the result to all $f \in L^\infty(\mathbb{R})$. More precisely, let $\{f_n\} \subset \mathcal{S}(\mathbb{R})$ such that $f_n \overset{*}{\rightharpoonup} f$ in $L^\infty(\mathbb{R})$ with $\|f_n\|_{L^\infty} \leq \|f\|_{L^\infty}$, and $f_n \longrightarrow f$ a.e. Since $f_n \in \mathcal{S}(\mathbb{R})$ we first have

$$\widehat{f_n g}(\xi) = \left\langle \widehat{f_n}, \tau_\xi \mathcal{R}\widehat{g} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1} \qquad \text{for all} \qquad n \in \mathbb{N}. \tag{28}$$

By weak* convergence of $f_n \overset{*}{\rightharpoonup} f$, we have

$$\lim_{n \longrightarrow \infty} \int_{\mathbb{R}} f_n(x) g(x) e^{-i\xi x}\, dx = \int_{\mathbb{R}} f(x) g(x) e^{-i\xi x}\, dx \quad \Longrightarrow \quad \lim_{n \longrightarrow \infty} \widehat{f_n g}(\xi) = \widehat{fg}(\xi) \tag{29}$$

for all $\xi \in \widehat{\mathbb{R}}$. On the other hand $\widehat{f_n} \overset{*}{\rightharpoonup} \widehat{f}$ in $\sigma(\mathcal{F}L^\infty, \mathcal{F}L^1)$, since for any $\phi \in L^1(\mathbb{R})$ then

$$\lim_{n \longrightarrow \infty} \left\langle \widehat{f_n}, \widehat{\phi} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1} = \lim_{n \longrightarrow} \int_{\mathbb{R}} (f_n)(x) \tilde{\phi}(x)\, dx = \int_{\mathbb{R}} (f_n)(x) \tilde{\phi}(x)\, dx = \left\langle \widehat{f}, \widehat{\phi} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1}. \tag{30}$$

Replace $\widehat{\phi}$ by $\tau_\xi \mathcal{R}\widehat{g}$ in (30) and using (28), (29) we obtain the result. ☐

Corollary 2.33. *If $f \in L^\infty(\mathbb{R})$ and $g \in L^1 \cap L^\infty(\mathbb{R})$ then in terms of support of distribution we have*

$$\operatorname{supp}\left(\widehat{f} * \widehat{g}\right) \subseteq \operatorname{supp}\left(\widehat{f}\right) + \operatorname{supp}\left(\widehat{g}\right).$$

Proof. We have

$$(\widehat{f} * \widehat{g})(\xi) = \left\langle \widehat{f}, \tau_\xi \mathcal{R}\widehat{g} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^\infty}.$$

Since $\widehat{g} \in C_0(\mathbb{R})$ it makes sense to talk about the classical meaning of the support of $\widehat{g}$, observe that

$$\eta \notin \xi - \operatorname{supp} \widehat{g} \quad \Longrightarrow \quad \xi - n \notin \operatorname{supp} \widehat{g} \quad \Longrightarrow \quad \widehat{g}(\xi - \eta) = 0 \quad \Longrightarrow \quad \tau_\xi \mathcal{R}\widehat{g}(\xi - \eta) = 0.$$

Thus $\operatorname{supp}\left(\tau_\xi \mathcal{R}\widehat{g}\right) \subset \xi - \operatorname{supp} \widehat{g}$. Consequently, if $\xi \notin \operatorname{supp} \widehat{g} + \operatorname{supp} \widehat{f}$ then $(\xi - \operatorname{supp} \widehat{g}) \cap \operatorname{supp} \widehat{f} = \emptyset$, thus

$$\operatorname{supp}\left(\tau_\xi \mathcal{R}\widehat{g}\right) \cap \operatorname{supp} \widehat{f} = \emptyset \quad \Longleftrightarrow \quad (\widehat{f} * \widehat{g})(\xi) = \left\langle \widehat{f}, \tau_\xi \mathcal{R}\widehat{g} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^\infty} = 0$$

by theorem 2.30. From that we obtain $\operatorname{supp}(\widehat{f} * \widehat{g}) \subset \operatorname{supp} \widehat{f} + \operatorname{supp} \widehat{g}$. □

Indeed it is still true in the general case $f, g \in L^\infty(\mathbb{R})$.

Theorem 2.34. *If $f, g \in L^\infty(\mathbb{R})$ then in terms of support of distribution we have*

$$\operatorname{supp}\left(\widehat{f} * \widehat{g}\right) \subseteq \operatorname{supp}\left(\widehat{f}\right) + \operatorname{supp}\left(\widehat{g}\right).$$

Proof. Let $\phi \in \mathcal{S}$ with $\widehat{\phi}$ has compact support disjoint from $\operatorname{supp} \widehat{f} + \operatorname{supp} \widehat{g}$, we will show that $\langle \widehat{fg}, \widehat{\phi} \rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1} = 0$. Observe that

$$\left\langle \widehat{fg}, \widehat{\phi} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1} = \int_{\mathbb{R}} f(x) g(x) \tilde{\phi}(x)\, dx = \left\langle \widehat{f}, \widehat{g\phi} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1}.$$

After identifying with distribution, $\widehat{g\phi}$ can be seen as the convolution of two pseudo-measures

$$\widehat{g\phi} = \widehat{g} * \widehat{\phi}$$

and thus corollary 2.33 can be applied to deduce that

$$\operatorname{supp}\left(\widehat{\tilde{g}} * \widehat{\phi}\right) \subset \operatorname{supp} \widehat{\tilde{g}} + \operatorname{supp} \widehat{\varphi} = -\operatorname{supp} \widehat{g} + \operatorname{supp} \widehat{\varphi}.$$

We claim that $\operatorname{supp}\left(\widehat{\tilde{g}} * \widehat{\phi}\right) \cap \operatorname{supp} \widehat{f} = \emptyset$, since otherwise there exists $\xi \in \operatorname{supp} \widehat{f}$, $\eta \in \operatorname{supp} \widehat{\varphi}$ and $\zeta \in \operatorname{supp} \widehat{g}$ such that

$$\xi = \eta - \zeta \quad \Longrightarrow \quad \eta = \xi + \zeta \in \operatorname{supp} \widehat{\varphi} \cap \left(\operatorname{supp} \widehat{f} + \operatorname{supp} \widehat{g}\right) = \emptyset$$

which is a contradiction. Thus $\widehat{g\phi} \in \mathcal{F}L^1$ with support disjoint from $\operatorname{supp} \widehat{f}$, which implies

$$\left\langle \widehat{fg}, \widehat{\phi} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1} = \left\langle \widehat{f}, \widehat{g\phi} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1} = 0$$

by theorem 2.30 and the proof is complete. □

4. We now show that a pseudo-measure with fintie support is a measure.

Theorem 2.35. *A pseudo-measure carried by one point is a Dirac measure.*

Proof. Let $f \in L^\infty(\mathbb{R})$ and assume $\operatorname{supp} \widehat{f} = \{0\}$.

Claim. If $\varphi_1, \varphi_2 \in A(\widehat{\mathbb{R}}) = \mathcal{F}(L^1(\mathbb{R}))$ and $\varphi_1(\xi) = \varphi_2(\xi)$ in a neighborhood of $\xi = 0$, then we have $\operatorname{supp}(\varphi_1 - \varphi_2)$ is compact and away from $\{0\} = \operatorname{supp} \widehat{f}$, thus

$$\left\langle \widehat{f}, \varphi_1 - \varphi_2 \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1} = 0 \quad \Longrightarrow \quad \left\langle \widehat{f}, \varphi_1 \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1} = \left\langle \widehat{f}, \varphi_2 \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1}.$$

Thus we can define $c = \langle \widehat{f}, \varphi \rangle$ where φ is any function in $\mathcal{F}L^1$ which $\varphi(\xi) = 1$ in a neighborhood of $\xi = 0$, and c will be well-defined independently to $\varphi \in \mathcal{F}L^1$. Let's recall the Fejer's kernel

$$\mathcal{K}(x) = \frac{1}{2\pi}\left(\frac{\sin x/2}{x/2}\right)^2 \qquad \text{has} \qquad \widehat{\mathcal{K}}(\xi) = \max\{1 - |\xi|, 0\}.$$

Since $\mathcal{K} \in L^1(\mathbb{R})$, we have

$$\left(\widehat{f} * \widehat{\mathcal{K}}\right)(\xi) = \widehat{f\mathcal{K}}(\xi) = \left\langle \widehat{f}, \tau_\xi \widehat{\mathcal{K}} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1} = \left\langle \widehat{f}, \widehat{\mathcal{K}}(\xi - \cdot) \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1}.$$

And

$$\operatorname{supp} \widehat{f\mathcal{K}} \subseteq \operatorname{supp} \widehat{\mathcal{K}} + \operatorname{supp} \widehat{f} = [-1, 1] + \{0\} = [-1, 1].$$

Thus if $|\xi| \geq 1$ we have $\widehat{f\mathcal{K}}(\xi) = 0$.

- If $-1 < \xi_1 < \xi_2 < 0$, there exists $\varepsilon > 0$ such that $\xi_i - \eta \in (-1, 0)$ for all $\eta \in (-\varepsilon, \varepsilon)$, then

  $$\widehat{\mathcal{K}}(\xi_2 - \eta) - \widehat{\mathcal{K}}(\xi_1 - \eta) = \xi_2 - \xi_1 \qquad \text{for all} \qquad \eta \in (-\varepsilon, \varepsilon).$$

  Let's define $\mathcal{J}(\eta) = (\xi_2 - \xi_1)\chi(\eta)$ where $\chi \in C_c^\infty$ with $\chi \equiv 1$ in $(-1, 1)$. It is clear that since $\chi \in C_c^\infty(\mathbb{R}) \subset \mathcal{S}(\mathbb{R})$, we have $\mathcal{J} \in \mathcal{F}L^1$. It follows that $\widehat{\mathcal{K}}(\xi_2 - \cdot) - \widehat{\mathcal{K}}(\xi_2 - \cdot) \equiv \mathcal{J}(\eta)$ where $\eta \in (-\varepsilon, \varepsilon)$, the claim implies that

  $$\left\langle \widehat{f}, \widehat{\mathcal{K}}(\xi_2 - \cdot) - \widehat{\mathcal{K}}(\xi_2 - \cdot) \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1} = \left\langle \widehat{f}, \mathcal{J} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1}$$

  and hence

  $$\widehat{f\mathcal{K}}(\xi_2) - \widehat{f\mathcal{K}}(\xi_1) = \left\langle \widehat{f}, \mathcal{J} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1} = c(\xi_2 - \xi_1).$$

  Since $\xi \longrightarrow \widehat{f\mathcal{K}}(\xi)$ is continuous, upon letting $\xi_1 \longrightarrow -1$ we obtain $\widehat{f\mathcal{K}}(\xi) = c(1 + \xi)$ for $-1 < \xi < 0$.

- If $0 < \xi_1 < \xi_2 < 1$, there exists $\varepsilon > 0$ such that $\xi_i - \eta \in (0, 1)$ for all $\eta \in (-\varepsilon, \varepsilon)$, then

  $$\widehat{\mathcal{K}}(\xi_2 - \eta) - \widehat{\mathcal{K}}(\xi_1 - \eta) = \xi_1 - \xi_2 \qquad \text{for all} \qquad \eta \in (-\varepsilon, \varepsilon).$$

  Let's define $\mathcal{J}(\eta) = (\xi_1 - \xi_2)\chi(\eta)$ where $\chi \in C_c^\infty$ with $\chi \equiv 1$ in $(-1, 1)$. It is clear that since $\chi \in C_c^\infty(\mathbb{R}) \subset \mathcal{S}(\mathbb{R})$, we have $\mathcal{J} \in \mathcal{F}L^1$. It follows that $\widehat{\mathcal{K}}(\xi_2 - \cdot) - \widehat{\mathcal{K}}(\xi_2 - \cdot) \equiv \mathcal{J}(\eta)$ where $\eta \in (-\varepsilon, \varepsilon)$, the claim implies that

  $$\left\langle \widehat{f}, \widehat{\mathcal{K}}(\xi_2 - \cdot) - \widehat{\mathcal{K}}(\xi_2 - \cdot) \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1} = \left\langle \widehat{f}, \mathcal{J} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1}$$

  and hence

  $$\widehat{f\mathcal{K}}(\xi_2) - \widehat{f\mathcal{K}}(\xi_1) = \left\langle \widehat{f}, \mathcal{J} \right\rangle_{\mathcal{F}L^\infty, \mathcal{F}L^1} = c(\xi_1 - \xi_2).$$

  Since $\xi \longrightarrow \widehat{f\mathcal{K}}(\xi)$ is continuous, upon letting $\xi_1 \longrightarrow 1$ we obtain $\widehat{f\mathcal{K}}(\xi) = c(1 - \xi)$ for $0 < \xi < 1$.

From that we have

$$\widehat{f\mathcal{K}}(\xi) = c \max\{1 - |\xi|, 0\} = c\widehat{\mathcal{K}}(\xi).$$

By the uniqueness of Fourier transform on $L^1(\mathbb{R})$, note that $f\mathcal{K} \in L^1(\mathbb{R}$, we have

$$\mathcal{F}(f\mathcal{K} - c\mathcal{K})(\xi) \equiv 0 \qquad \Longrightarrow \qquad f\mathcal{K} = c\mathcal{K} \qquad \Longrightarrow \qquad f \equiv c \quad \text{a.e}$$

and thus $\widehat{f} = c\delta_0$. □

5. Theorem 2.35 implies following approximation theorem.

Theorem 2.36. *Let $\xi \in \widehat{\mathbb{R}}$ and denote $I(\xi) = \{f \in \mathcal{F}L^1 : f(\xi) = 0\}$ and $I_0(\xi) = \{f \in \mathcal{S}(\widehat{\mathbb{R}}) : \xi \notin \operatorname{supp} f\}$, then $I_0(\xi)$ is dense in $I(\xi)$ in the topology of $(\mathcal{F}L^1, \|\cdot\|_{\mathcal{F}L^1})$.*

Proof. Note that $I_0(\xi)$ and $I(\xi)$ are linear sub-spaces of $\mathcal{F}L^1$. If the conclusion of the theorem is not true, then by Hahn-Banach theorem there exists $\widehat{f} \in (\mathcal{F}L^1)^* = \mathcal{F}L^\infty$ such that $\widehat{f}$ doesn't vanish on $I(\xi)$ and $\langle \widehat{\varphi}, \widehat{f} \rangle = 0$ for all $\varphi \in I_0(\xi)$. It is easy to see that $\operatorname{supp} \widehat{f} = \{0\}$, and thus by theorem 2.35 we have $\widehat{f} = c\delta_0$ for some $c \in \mathbb{R}$, then obviously $\langle \widehat{\varphi}, \widehat{f} \rangle = 0$ for all $\widehat{\varphi} \in I(\xi)$, which is a contradiction. □

3 Almost periodic functions on the real line

Let f be a complex-value function on $\mathbb{R}$ and let $\varepsilon > 0$. We call

- An ε-almost period of f is a number τ such that $\sup_{x\in\mathbb{R}} |f(x-\tau) - f(x)| < \varepsilon$.
- A set $\mathcal{F} \subset \mathbb{R}$ is called "relatively dense" in $\mathbb{R}$ if and only if there exists a constant $\Lambda = \Lambda(\mathcal{F})$ such that $(x, x+\Lambda) \cap \mathcal{F} \neq \emptyset$ for all $x \in \mathbb{R}$.
- A continuous function f is (uniformly) almost periodic, denoted by u.a.p if for every $\varepsilon > 0$, the set $\mathcal{F}_\varepsilon$ of all ε-almost period of f is relatively dense in $\mathbb{R}$.

We denote the space of all uniformly almost-periodic functions on $\mathbb{R}$ to be AP($\mathbb{R}$), and for $f \in$ AP($\mathbb{R}$) we denote by $\Lambda = \Lambda(\varepsilon, f)$ the length of the interval in the above definition. Some examples and properties of functions in AP($\mathbb{R}$) are:

- Continuous periodic functions are almost-periodic.
- $f(x) = \cos x + \cos(\sqrt{2}x)$ is almost-periodic but not periodic.
- If $f \in$ AP($\mathbb{R}$) then so are $|f|$, $\widehat{f}$ (if $\widehat{f}$ makes sense), af for any $a \in \mathbb{C}$ and $f(\lambda x)$ for any real λ.

3.1 Definition and basic properties

1. We have AP($\mathbb{R}$) $\subset$ BUC($\mathbb{R}$). Indeed, take $\Lambda = \Lambda(1, f)$ and for any $x \in \mathbb{R}$ let $\tau \in (x - \Lambda, x)$ be an 1-almost period then
$$|f(x)| \le 1 + |f(x-\tau)| \le 1 + \sup_{x\in[0,\Lambda]} |f(x)|.$$
For $\varepsilon > 0$, let $\Lambda\Lambda(\varepsilon/3, f)$ then since f is uniformly continuous on $[-2\Lambda, 2\Lambda]$ there exists $0 < \delta < \Lambda$ such that $|f(x) - f(y)| < \frac{\varepsilon}{3}$ whenever $x, y \in [-2\Lambda, 2\Lambda]$ with $|x-y| < \delta$. Let $\tau \in (x-\Lambda, x)$ be an $\frac{\varepsilon}{3}$-almost period of f, then for any $y \in \mathbb{R}$ with $|x-y| < \delta$ we have
$$|f(x) - f(y)| \le |f(x) - f(x-\tau)| + |f(x-\tau) - f(y-\tau)| + |f(y) - f(y-\tau)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$
where in the second term we used the uniform continuity of f on $[-2\Lambda, 2\Lambda]$.

   As a corollary, it is easy to see that if $f \in$ AP($\mathbb{R}$) then $f^2 \in$ AP($\mathbb{R}$).

2. For a function $f \in L^\infty(\mathbb{R})$ we denote we denote by $f_a(x) = \tau_a f(x) = f(x-a)$ the translation with $a \in \mathbb{R}$, and $W_0(f) = \{f_a(\cdot) : a \in \mathbb{R}\}$ set of all translations.

   Theorem 3.1. *A function $f \in L^\infty(\mathbb{R})$ is uniformly almost periodic if and only if $W_0(f)$ is precompact in metric space $L^\infty(\mathbb{R})$.*

3. The **translation convex hull** $W(f)$ of a function $f \in L^\infty(\mathbb{R})$ is the closed convex hull of $\bigcup_{|a|\le 1} W_0(af)$. In other words, it the the set of uniform limits (in L^∞) of functions of the form
$$\sum_{k=1}^m a_k \tau_{x_k} f \qquad \text{where} \qquad x_k \in \mathbb{R}, \qquad \sum_{k=1}^m |a_k| \le 1, \qquad m \in \mathbb{N}.$$
For $f \in L^\infty(\mathbb{R})$, we define
$$W'(f) = \text{closure}\Big\{\varphi * f : \|\varphi\|_{L^1(\mathbb{R})} \le 1\Big\}.$$
Here by closure we mean the closure in $L^\infty(\mathbb{R})$.

   Lemma 3.2. *If $f \in$ BUC($\mathbb{R}$) then $W'(f) \equiv W(f)$.*

   Lemma 3.3.

   (a) If $f \in L^\infty(\mathbb{R})$ then for every $\xi \in \mathbb{R}$ we have $W(e^{i\xi x} f) = \Big\{e^{i\xi x} g : g \in W(f)\Big\}$.

(b) *If* $f \in C_0(\mathbb{R})$ *then* $W(f) \subset C_0(\mathbb{R})$.

Theorem 3.4. *If* $f \in L^\infty(\mathbb{R})$, *we have* $W(f)$ *is convex and closed in* $L^\infty(\mathbb{R})$. *Furthermore* $W(f)$ *is compact iff* $W_0(f)$ *is precompact iff* $f \in \mathrm{AP}(\mathbb{R})$.

This characterization of $\mathrm{AP}(\mathbb{R})$ gives us a powerful tool to prove the following theorem.

Theorem 3.5. *The space of uniformly almost-periodic functions* $\mathrm{AP}(\mathbb{R})$ *is a closed sub-algebra of* $L^\infty(\mathbb{R})$, *i.e it is closed as a sub-space, also closed under multiplication and addition.*

Proof. Let $f,g \in \mathrm{AP}(\mathbb{R})$, we first show $f+g \in \mathrm{AP}(\mathbb{R})$. Observe that $W(f+g) \subset W(f)+W(g)$, and since $f,g \in \mathrm{AP}(\mathbb{R})$ we have $W(f), W(g)$ are both compact in $L^\infty(\mathbb{R})$, thus $W(f)+W(g)$ is compact, therefore $W(f+g)$ is a closed subset of $W(f)+W(g)$, thus $W(f+g)$ is compact and hence $f+g \in \mathrm{AP}(\mathbb{R})$. Since $f^2, g^2, (f+g)^2 \in \mathrm{AP}(\mathbb{R})$ we have

$$fg = \frac{1}{2}\left((f+g)^2 - f^2 - g^2\right) \in \mathrm{AP}(\mathbb{R}).$$

We have proved that $\mathrm{AP}(\mathbb{R})$ is a sub-algebra of $L^\infty(\mathbb{R})$. Now let's consider f in the closure of $\mathrm{AP}(\mathbb{R})$ in $L^\infty(\mathbb{R})$, it is clear that f is bounded uniformly continuous. Given $\varepsilon > 0$, we can find $g \in \mathrm{AP}(\mathbb{R})$ such that $\|f-g\|_{L^\infty} \le \frac{\varepsilon}{3}$. Now let τ is an $\frac{\varepsilon}{3}$-almost period of g, we then have

$$\sup_{x\in\mathbb{R}} |f(x-\tau) - f(x)| \le \sup_{x\in\mathbb{R}} |f(x-\tau) - g(x-\tau)| + \sup_{x\in\mathbb{R}} |g(x-\tau) - g(x)| + \sup_{x\in\mathbb{R}} |g(x) - f(x)| < \varepsilon.$$

Thus τ is an ε-almost period of f, and every interval of length $\Lambda\left(\frac{\varepsilon}{3}, g\right)$ contains an ε-almost period of f, so $f \in \mathrm{AP}(\mathbb{R})$. ☐

As a consequence, sum of any finite almost-periodic functions is again almost-periodic.

4. A trigonometric polynomial on $\mathbb{R}$ is a function of the form $f(x) = \sum_{k=1}^n a_k e^{i\xi_k x}$ where $a_k \in \mathbb{C}$ and $\xi_k \in \mathbb{R}$ are called the frequencies of f. Theorem 3.4 says that all trigonometric polynomials and its uniform limits are almost-periodic.

3.2 Mean value of almost periodic functions

1. The norm spectrum of a function $f \in L^\infty(\mathbb{R})$ is the set

$$\sigma(f) = \left\{\xi \in \mathbb{R} : ae^{i\xi x} \in W(f) \text{ for a complex number } a \neq 0\right\}.$$

Note that $\sigma(f)$ maybe empty, for example if $f \in C_0(\mathbb{R})$ then $W(f) \subset C_0(\mathbb{R})$ by lemma 3.3, but there is no $a \neq 0$ such that $ae^{i\xi x} \in C_0(\mathbb{R})$ for some ξ.

Lemma 3.6. *For* $f \in L^\infty(\mathbb{R})$ *and* $\xi \in \mathbb{R}$ *then* $\sigma\left(e^{i\xi(\cdot)} f\right) = \xi + \sigma(f) = \{\xi + \eta : \eta \in \sigma(f)\}$.

2. If $f \in L^\infty(\mathbb{R})$, $\widehat{f}$ can be seen as a tempered distribution on $\mathcal{S}(\mathbb{R})$, which can be extended to a linear functional on $\mathcal{F}L^1$ by

$$\left\langle \widehat{\phi}, \widehat{f} \right\rangle_{\mathcal{F}L^1, \mathcal{F}L^\infty} = \langle \phi, f \rangle_{L^1, L^\infty} = \int_{\mathbb{R}} \phi(x) \overline{f(x)}\, dx \qquad \text{for all} \qquad \phi \in L^1(\mathbb{R}).$$

The multiplication of a distribution with a function $g \in L^\infty$ function is defined by the action

$$\left\langle \widehat{\varphi}, g\widehat{f} \right\rangle_{\mathcal{F}L^1, \mathcal{F}L^\infty} = \left\langle \widehat{\varphi g}, \widehat{f} \right\rangle_{\mathcal{F}L^1, \mathcal{F}L^\infty}.$$

The notion of support of $\widehat{f}$ extends naturally and consistently onto the new test space $\mathcal{F}L^1$.

Proposition 3.7. *If* $f \in L^\infty(\mathbb{R})$ *then* $\sigma(f) \subset \operatorname{supp} \widehat{f}$.

Proof. As a distribution, for any $a \in \mathbb{R}$ we have

$$\widehat{\tau_a f}(\xi) = e^{-i\xi a}\widehat{f}(\xi) \qquad \Longrightarrow \qquad \operatorname{supp} \widehat{\tau_a(f)} = \operatorname{supp} \widehat{f}.$$

Consequently, any $\varphi \in W(f)$ satisfies $\operatorname{supp} \widehat{\varphi} \subset \operatorname{supp} \widehat{f}$. Indeed:

- Any finite convex combination of translations is clearly has support contained in $\operatorname{supp} \widehat{f}$.
- When $\varphi \in W(f)$ is the uniform limit of a sequence φ_n where φ_n are convex finite combination of translations, i.e. $\|\varphi_n - \varphi\|_{L^\infty} \longrightarrow 0$, then since $\operatorname{supp} \widehat{\varphi_n} \subset \operatorname{supp} \widehat{f}$, for any $\phi \in L^1(\mathbb{R})$ with $\operatorname{supp} \widehat{\phi} \cap \operatorname{supp} \widehat{f} = \emptyset$, we have $\operatorname{supp} \widehat{\phi} \cap \operatorname{supp} \widehat{\varphi_n} = \emptyset$, thus

$$\left\langle \widehat{\phi}, \widehat{\varphi} \right\rangle_{\mathcal{F}L^1, \mathcal{F}L^\infty} = \lim_{n \longrightarrow \infty} \left\langle \widehat{\phi}, \widehat{\varphi_n} \right\rangle_{\mathcal{F}L^1, \mathcal{F}L^\infty} = 0$$

  since $\widehat{\varphi_n} \longrightarrow \widehat{\varphi}$ in $\mathcal{F}L^\infty$. Hence as a distribution, $\widehat{\varphi}$ vanish on $\mathbb{R}^n \backslash \operatorname{supp} \widehat{f}$, i.e., $\operatorname{supp} \widehat{\varphi} \subset \operatorname{supp} \widehat{f}$.

If $\varphi(\cdot) = a e^{i\xi(\cdot)} \in W(f)$ then $\widehat{\varphi}(\cdot) = 2\pi a \delta_0(\cdot - \xi)$ must satisfy $\operatorname{supp} \widehat{\varphi} = \{\xi\} \subset \operatorname{supp} \widehat{f}$. Thus we have proved that $\xi \in \sigma(f)$ implies $\xi \in \operatorname{supp} \widehat{f}$. ☐

3. The Fejér kernel $\mathcal{K}_\lambda$ is a approximation of identity as $\lambda \longrightarrow \infty$. However as $\lambda \longrightarrow 0$ we have the following result.

Proposition 3.8. *Let $f \in \mathrm{BUC}(\mathbb{R})$, assume that $\mathcal{K}_\lambda * f$ converges uniformly in L^∞ norm as $\lambda \longrightarrow 0$ to a limit which is not identically zero, then $0 \in \sigma(f)$.*

Proof. Assume $\mathcal{K}_\lambda * f \longrightarrow u$ in $L^\infty(\mathbb{R})$ as $\lambda \longrightarrow 0$, let's define $g_\lambda = \mathcal{K}_\lambda * f \in L^\infty(\mathbb{R})$ for $\lambda > 0$ then as distributions we have

$$\widehat{g_\lambda} = \widehat{\mathcal{K}}\left(\frac{\cdot}{\lambda}\right)\widehat{f} = \in \mathcal{F}L^\infty \qquad \Longrightarrow \qquad \operatorname{supp} \widehat{g_\lambda} \subset \operatorname{supp} \widehat{\mathcal{K}}\left(\frac{\cdot}{\lambda}\right) \cap \operatorname{supp} \widehat{f} \subset [\lambda, \lambda].$$

Since $\widehat{g_\lambda} \longrightarrow \widehat{u}$ in $\mathcal{F}L^\infty$ as $\lambda \longrightarrow 0$, we have $\operatorname{supp} \widehat{u} = \{0\}$. Indeed, if $\varphi \in L^1(\mathbb{R})$ with $\operatorname{supp} \widehat{\varphi}$ is compactly supported away from $\{0\}$, there exists $\lambda_0 > 0$ so that $\operatorname{supp} \widehat{\varphi} \cap [\lambda, \lambda] = \emptyset$ for all $\lambda < \lambda_0$, thus $\langle \widehat{\varphi}, \widehat{g_\lambda} \rangle_{\mathcal{F}L^1, \mathcal{F}L^\infty} = 0$ for all $\lambda < \lambda_0$ which implies that $\langle \widehat{\varphi}, \widehat{u} \rangle_{\mathcal{F}L^1, \mathcal{F}L^\infty} = 0$, and since $u \not\equiv 0$, we have $\operatorname{supp} \widehat{u} = \{0\}$. Since $\operatorname{supp} \widehat{u} = \{0\}$ we have $\widehat{u} \equiv c\delta_0$ for some $c \neq 0$, which implies $u \equiv c$. By lemma 3.2, we have

$$g_\lambda = \mathcal{K}_\lambda * f \in W(f) \text{ for all } \lambda \qquad \Longrightarrow \qquad u = \lim_{\lambda \longrightarrow 0} g_\lambda \in W(f)$$

as the limit is taken in L^∞ norm. Thus we have shown that $0 \neq c \in W(f)$, which mean $0 \in \sigma(f)$. ☐

Note that within our notation:

- <u>**As a distribution**</u> u we have

$$\widehat{\delta_0} = \frac{1}{2\pi}, \qquad \widehat{1} = \delta_0, \qquad \widehat{\widehat{u}} = \frac{1}{2\pi}\tilde{u}$$

  and if $\varphi \in \mathcal{S}(\mathbb{R})$ is a test function then

$$(\varphi^\vee)^\vee = \frac{1}{2\pi}\tilde{\varphi}.$$

- <u>**As a measure**</u> we have

$$\widehat{\delta_0}(\xi) = \int_{\mathbb{R}} e^{-i\xi x}\, d\delta_0(x) = 1 \qquad \Longrightarrow \qquad \widehat{\delta_0} = 1.$$

4. If $\mu \in M(\widehat{\mathbb{R}})$, it can be identified with $\mathcal{F}L^\infty$ by setting $f(x) = \widehat{\mu}(-x) \in L^\infty(\mathbb{R})$ then by Parseval's formula and $\widehat{\widehat{f}}(x) = 2\pi f(-x)$ we have

$$\langle \widehat{\varphi}, \mu \rangle_{C_0(\widehat{\mathbb{R}}), M(\widehat{\mathbb{R}})} = \int_{\widehat{\mathbb{R}}} \widehat{\varphi}(\xi) \overline{d\mu(\xi)} = \int_{\mathbb{R}} \varphi(-x) \overline{\widehat{\mu}(x)}\, dx = \int_{\mathbb{R}} \varphi(x) \overline{f(x)}\, dx = \left\langle \widehat{\varphi}, \widehat{f} \right\rangle_{\mathcal{F}L^1, \mathcal{F}L^\infty}$$

for any $\varphi \in L^1(\mathbb{R})$, note that $\widehat{\varphi} \in C_0(\widehat{\mathbb{R}})$ by Riemann-Lebesgue lemma. Thus $\mu = \widehat{f} \in \mathcal{F}L^\infty$, note that there is no factor 2π in this case, based on our notations.

Corollary 3.9. *Let $\mu \in M(\widehat{\mathbb{R}})$ and assume $\mu(\{0\}) \neq 0$. Let $f(x) = \widehat{\mu}(-x)$ then $0 \in \sigma(f)$. In fact we have*

$$\lim_{\lambda \longrightarrow 0} \left\| \mathcal{K}_\lambda * f - \mu(\{0\}) \right\|_{L^\infty} = 0.$$

Proof. Let $g_\lambda = \mathcal{K}_\lambda * f$ then $\widehat{g_\lambda} = \widehat{\mathcal{K}_\lambda}\, \mu \in M(\widehat{\mathbb{R}})$. For any test function $\varphi \in L^\infty(\widehat{\mathbb{R}})$ we have

$$\begin{aligned}\int_{\widehat{\mathbb{R}}} \varphi(\xi)\, d\left(\widehat{g_\lambda}(\xi)\right) &= \int_{\widehat{\mathbb{R}}} \left(1 - \frac{|\xi|}{\lambda}\right) \varphi(\xi) \chi_{(-\lambda,\lambda)}(\xi)\, d\mu(\xi) \\ &= \varphi(0)\mu(\{0\}) + \int_{\widehat{\mathbb{R}}} \left(1 - \frac{|\xi|}{\lambda}\right) \varphi(\xi) \chi_{(-\lambda,\lambda)\backslash\{0\}}(\xi)\, d\mu(\xi) \longrightarrow \varphi(0)\mu(\{0\})\end{aligned}$$

as $\lambda \longrightarrow 0$ by the dominated convergence theorem, since $\chi_{(-\lambda,\lambda)\backslash\{0\}} \longrightarrow 0$ point-wise everywhere and $(1 - |\xi/\lambda)\varphi(\xi)$ is bounded by $\|\varphi\|_{L^\infty}$. If we restrict the space of test functions to $C_0(\widehat{\mathbb{R}})$ then it gives $\widehat{g_\lambda} \overset{*}{\rightharpoonup} \mu(\{0\})\delta_0$ in the weak* topology of $M(\widehat{\mathbb{R}})$, or in the distribution sense. In deed from the above estimate we get

$$\left| \int_{\widehat{\mathbb{R}}} \varphi\, d\left(\widehat{g_\lambda} - \mu(\{0\})\delta_0\right) \right| \leq \|\varphi\|_{L^\infty} \int_{(-\lambda,\lambda)\backslash\{0\}} \left(1 - \frac{|\xi|}{\lambda}\right) d|\mu|(\xi) \leq \|\varphi\|_{L^\infty} |\mu|\left((\lambda,\lambda)\backslash\{0\}\right)$$

which gives us, by Rieze representation theorem

$$\left\| \widehat{g_\lambda} - \mu(\{0\})\delta_0 \right\|_{M(\widehat{\mathbb{R}})} = \sup_{\|\varphi\|_{L^\infty} \leq 1} \left| \int_{\widehat{\mathbb{R}}} \varphi\, d\left(\widehat{g_\lambda} - \mu(\{0\})\delta_0\right) \right| \leq |\mu|\left((\lambda,\lambda)\backslash\{0\}\right) \longrightarrow 0$$

thus $g_\lambda \longrightarrow \mu(\{0\})$ uniformly in $L^\infty(\mathbb{R})$ and the result follows from proposition 3.8. □

5. In fact proposition 3.8 is true for any general summability kernel.

Proposition 3.10. *Let $f \in \mathrm{BUC}(\mathbb{R})$ and $F \in L^1(\mathbb{R})$. Define $F_\lambda(x) = \lambda F(\lambda x)$ for $\lambda > 0$. Assume that $F_\lambda * f$ converges uniformly in L^∞ norm as $\lambda \longrightarrow 0$ to a limit which is not identically zero, then $0 \in \sigma(f)$.*

Proof. Assume $F_\lambda * f \longrightarrow u$ in $L^\infty(\mathbb{R})$ as $\lambda \longrightarrow 0$ and $u \not\equiv 0$. Let $\mathcal{K}_n(x) = n\mathcal{K}(nx)$ where $\mathcal{K}$ is the Fejer's kernel as usual and $n \in \mathbb{N}$, let $G_n = F * \mathcal{K}_n$ then $G_n \longrightarrow F$ in $L^1(\mathbb{R})$ as $n \longrightarrow \infty$, then $H_n = F - G_n \in L^1(\mathbb{R})$ with $\|H_n\|_{L^1} \longrightarrow 0$ as $n \longrightarrow \infty$. We have the decomposition

$$F = G_n + H_n, \qquad \lim_{n \longrightarrow \infty} \|H_n\|_{L^1} = 0, \qquad \widehat{G_n} = \widehat{F\mathcal{K}_n} \text{ is compactly supported in } [-n, n].$$

For each $\lambda > 0$, let

$$\begin{cases} G_{n,\lambda}(x) = \lambda G_n(\lambda x) \\ H_{n,\lambda}(x) = \lambda H_n(\lambda x) \end{cases} \implies \quad F_\lambda(x) = G_{n,\lambda}(x) + H_{n,\lambda}(x) \quad \text{for} \quad n \in \mathbb{N}, \lambda > 0. \tag{31}$$

Since $\widehat{G_{n,\lambda}}(\xi) = \widehat{G_n}\left(\frac{\xi}{\lambda}\right)$, for each $\lambda > 0$ and $n \in \mathbb{N}$ we have

$$\operatorname{supp} \widehat{G_{n,\lambda} * f} = \operatorname{supp} \left(\widehat{G_{n,\lambda}} \cdot \widehat{f}\right) \subseteq \operatorname{supp} \widehat{G_{n,\lambda}} \subset [-\lambda n, \lambda n]. \tag{32}$$

Let $\lambda = \frac{1}{n^2}$, from (31) we have

$$\widehat{F_{n^{-2}} * f} = \widehat{G_{n,n^{-2}} * f} + \widehat{H_{n,n^{-2}} * f} \qquad \text{as elements of} \quad \mathcal{F}L^\infty. \tag{33}$$

Since $f \in L^\infty(\mathbb{R})$, we have

$$\|H_{n,n^{-2}} * f\|_{L^\infty} \le \|f\|_{L^\infty}\|H_{n,n^{-2}}\|_{L^1} = \|f\|_{L^\infty}\|H_n\|_{L^1} \longrightarrow 0 \qquad \text{as} \qquad n \longrightarrow \infty.$$

From that and (31) we obtain $G_{n,n^{-2}} * f \longrightarrow u$ in $L^\infty(\mathbb{R})$, which means $\widehat{G_{n,n^{-2}} * f} \longrightarrow \widehat{u}$ in $\mathcal{F}L^\infty$, and together with (33) we obtain $\operatorname{supp} \widehat{u} = \{0\}$ by an argument similar to the proof of proposition 3.8. Therefore $u \equiv c$ for some $c \neq 0$ and by lemma 3.2, we have

$$F_\lambda * f \in W(f) \text{ for all } \lambda \qquad \Longrightarrow \qquad u = \lim_{\lambda \longrightarrow 0} F_\lambda * f \in W(f)$$

in L^∞ norm. Thus we have shown that $0 \neq c \in W(f)$, which mean $0 \in \sigma(f)$. □

6. The condition of existence of a uniform limit of $F_\lambda * f$ as $\lambda \longrightarrow 0$ can clearly be replaced by the less stringent condition of the existence of a nonvanishing limit point.

Proposition 3.11. *Let $f \in \mathrm{AP}(\mathbb{R})$ and assume $0 \notin \sigma(f)$, then for any $F \in L^1(\mathbb{R})$ we have*

$$\lim_{\lambda \longrightarrow 0} \|F_\lambda * f\|_{L^\infty(\mathbb{R})} = 0$$

where $F_\lambda(x) = \lambda F(\lambda x)$.

Proof. Since $f \in \mathrm{AP}(\mathbb{R})$, $W(f)$ is compact in $L^\infty(\mathbb{R})$ and $\{F_\lambda * f\}_{\lambda>0}$ is a sequence in $L^\infty(\mathbb{R})$, if the claim is not true then there exists a sequence $\lambda_n \longrightarrow 0$ such that $F_{\lambda_n} * f \longrightarrow u$ in $L^\infty(\mathbb{R})$ as $n \longrightarrow \infty$ where $u \in L^\infty(\mathbb{R})$ so that $u \not\equiv 0$. Proposition 3.10 implies that $0 \in \sigma(f)$, which is a contradiction. □

Conversely, we have the following:

Proposition 3.12. *Let $f \in \mathrm{AP}(\mathbb{R})$, $F \in L^1(\mathbb{R})$ and $\int_{\mathbb{R}} F(x)\,dx \neq 0$. If for some sequence $\lambda_n \longrightarrow 0$ we have*

$$\lim_{\lambda_n \longrightarrow \infty} \left\|F_{\lambda_n} * f\right\|_{L^\infty(\mathbb{R})} = 0$$

where $F_\lambda(x) = \lambda F(\lambda x)$, then $0 \notin \sigma(f)$.

Proof. It is easy to see that for any translation $\varphi = \tau_a(f)$ we have $\|F_{\lambda_n} * \varphi\|_{L^\infty} = \|F_{\lambda_n} * f\|_{L^\infty}$, thus $\lim_{\lambda_n \longrightarrow \infty} \left\|F_{\lambda_n} * \varphi\right\|_{L^\infty(\mathbb{R})} = 0$. Consequently this limit holds true for all φ whose are finite convex combinations of translations, which in turn implies $\lim_{\lambda_n \longrightarrow \infty} \left\|F_{\lambda_n} * \varphi\right\|_{L^\infty(\mathbb{R})} = 0$ for all $\varphi \in W(f)$. To be precise, let $\{\varphi_k\}$ be a sequence in $L^\infty(\mathbb{R})$ where each φ_k is a finite convex combination of translation of f, and $\|\varphi_k - \varphi\|_{L^\infty} \longrightarrow 0$, we then have $\lim_{n \longrightarrow \infty} \|F_{\lambda_n} * \varphi_k\|_{L^\infty} = 0$ for all $k \in \mathbb{N}$, and

$$\left\|F_{\lambda_n} * \varphi\right\|_{L^\infty} \le \left\|F_{\lambda_n} * \varphi_k\right\|_{L^\infty} + \left\|F_{\lambda_n} * (\varphi - \varphi_k)\right\|_{L^\infty} \le \left\|F_{\lambda_n} * \varphi_k\right\|_{L^\infty} + \|F_{\lambda_n}\|_{L^1}\|\varphi_k - \varphi\|_{L^\infty}.$$

Since $\|F_{\lambda_n}\|_{L^1} = \|F\|_{L^1}$ for all $n \in \mathbb{N}$, let $n \longrightarrow \infty$ we obtain

$$\limsup_{n \longrightarrow \infty} \left\|F_{\lambda_n} * \varphi\right\|_{L^\infty} \le \|F\|_{L^1}\|\varphi_k - \varphi\|_{L^\infty}.$$

Now let $k \longrightarrow \infty$ we obtain the limit is zero. Now we observe that:

$$0 \in \sigma(f) \qquad \Longleftrightarrow \qquad \text{there exists a constant } C \neq 0 \text{ such that} \qquad C \in W(f).$$

Thus if $0 \in \sigma(f)$, with the constant C as above we must have

$$\lim_{\lambda_n \longrightarrow 0} \|F_{\lambda_n} * C\|_{L^\infty} = C\left(\int_{\mathbb{R}} F(x)\,dx\right) = 0$$

which is a contradiction. □

7. We are now ready to prove one of the most important property of almost periodic functions.

Theorem 3.13 (Mean-value theorem)**.** *To every $f \in \mathrm{AP}(\mathbb{R})$ there corresponds a unique number $M(f)$, called the mean value of f, having the property that $0 \notin \sigma\big(f - M(f)\big)$. Furthermore, for any $F \in L^1(\mathbb{R})$ if we set $F_\lambda(x) = \lambda F(\lambda x)$ then*

$$\lim_{\lambda \longrightarrow 0} \left\| F_\lambda * f - \widehat{F}(0) M(f) \right\|_{L^\infty(\mathbb{R})} = 0.$$

Proof. Recall that for $\mathcal{K}_\lambda(x) = \lambda \mathcal{K}(\lambda x)$ we then have $\{\mathcal{K}_\lambda * f\}_{\lambda > 0}$ is a subset of the compact set $W(f) \subset L^\infty(\mathbb{R})$, thus to any sequence $\lambda_n \longrightarrow 0$ there corresponds a subsequence and a limit point $\mathcal{K}_{\lambda_{n_k}} * f \longrightarrow u$ in $L^\infty(\mathbb{R})$ as $\lambda_{n_k} \longrightarrow 0$, and $u \equiv C$ as a constant by proposition 3.8. Let α be such a limit point of a sequence $\mathcal{K}_{\lambda_{n_k}} * f$ as $\lambda_{n_k} \longrightarrow 0$ like that, we have

$$\lim_{\lambda_{n_k} \longrightarrow 0} \left\| \mathcal{K}_{\lambda_n} * f - \alpha \right\|_{L^\infty(\mathbb{R})} = \lim_{\lambda_{n_k} \longrightarrow 0} \left\| \mathcal{K}_{\lambda_{n_k}} * (f - \alpha) \right\|_{L^\infty(\mathbb{R})} = 0.$$

By proposition 3.12 we obtain $0 \notin \sigma(f - \alpha)$, consequently proposition 3.11 implies that the limit above holds for the full sequence $\lambda \longrightarrow 0$. If β is another constant such that $0 \notin \sigma(f - \beta)$ then by proposition 3.11 we must have

$$\lim_{\lambda \longrightarrow 0} \left\| \mathcal{K}_\lambda * (f - \beta) \right\|_{L^\infty(\mathbb{R})} = 0$$

and hence

$$|\alpha - \beta| = \lim_{\lambda \longrightarrow 0} \left\| \mathcal{K}_\lambda * (\alpha - \beta) \right\|_{L^\infty(\mathbb{R})} = 0 \qquad \Longrightarrow \qquad \alpha = \beta.$$

Thus the property $0 \notin \sigma(f - \alpha)$ determines α uniquely and we set $M(f) = \alpha$. Finally by replacing the Fejer's kernel $\mathcal{K}$ by $F \in L^1(\mathbb{R})$ and using proposition 3.11 with f being replaced by $f - \widehat{F}(0)M(f)$ we obtain the second limit. □

In particular, by taking some specific $F \in L^1(\mathbb{R})$ we obtain the following:

Corollary 3.14. *Let $f \in \mathrm{AP}(\mathbb{R})$ then*

$$M(f) = \lim_{\lambda \longrightarrow 0} \frac{\lambda}{2} \int_{-\frac{1}{\lambda}}^{\frac{1}{\lambda}} f(x)\,dx = \lim_{T \longrightarrow \infty} \frac{1}{2T} \int_{-T}^{T} f(x)\,dx = \lim_{T \longrightarrow \infty} \frac{1}{T} \int_{-T}^{0} f(x)\,dx = \lim_{T \longrightarrow \infty} \frac{1}{T} \int_{0}^{T} f(x)\,dx.$$

8. Using the mean value we can determine the norm spectrum of $f \in \mathrm{AP}(\mathbb{R})$ completely. From lemma 3.6 it is clear that

$$\xi \in \sigma(f) \qquad \Longleftrightarrow \qquad 0 \in \sigma(f e^{-i\xi x}) \qquad \Longleftrightarrow \qquad M(f e^{-i\xi x}) \neq 0. \tag{34}$$

Recall that every measure $\mu \in M(\mathbb{R})$ can be decomposed to a sum $\mu = \mu_c + \mu_d$ where μ_c is continuous and μ_d is discrete. If $\widehat{f}$ is a measure then corollary 3.9 implies that $\widehat{f}(\{0\}) = M(f)$ and similarly $\widehat{f}(\{\xi\}) = M(f e^{-i\xi x})$ and thus we can recover the discrete part of f. We shall soon see that $\widehat{f}$ has no continuous part when $f \in \mathrm{AP}(\mathbb{R})$.

9. We summarize some basic properties of the mean value of $f \in \mathrm{AP}(\mathbb{R})$ in the following theorem.

Theorem 3.15 (Basic properties of the mean value)**.** *For $f, g \in \mathrm{AP}(\mathbb{R})$ and $a \in \mathbb{R}$ we have*

(a) $M(f + g) = M(f) + M(g)$, $M(af) = aM(f)$ and $M\big(\tau_a f\big) = M(f)$.

(b) If $f(x) \geq 0$ and $f \not\equiv 0$, then $M(f) > 0$.

Proof.

(a) Using the mean value theorem, let $F \in L^1(\mathbb{R})$ with $\int_{\mathbb{R}} F(x)\,dx = 1$, and $F_\lambda(x) = \lambda F(x)$, we then have

$$F_\lambda * (af + g) = aF_\lambda * f + F_\lambda * g \qquad \Longrightarrow \qquad M(af + g) = aM(f) + M(g)$$

by letting $\lambda \longrightarrow 0$. For translation, we have

$$(F_\lambda * (\tau_a f)(x) = (F_\lambda * f)(x - a) \qquad \Longrightarrow \qquad M(\tau_a f) = M(f)$$

by letting $\lambda \longrightarrow 0$ since the limit is uniform in $L^\infty(\mathbb{R})$.

(b) Since $f \not\equiv 0$ and $M(\tau_a f) = M(f)$ for all $a \in \mathbb{R}$, we can assume $f(0) > 0$. Since f is continuous, there exists $\varepsilon > 0$ such that $f(x) \geq \varepsilon$ on $[-\varepsilon, \varepsilon]$. Let $\Lambda = \Lambda\left(\frac{\varepsilon}{2}, f\right)$, for any interval of length Λ, say $[a, a+\Lambda]$ we can pick $\tau \in [a, a+\Lambda]$ as a $\frac{\varepsilon}{2}$-almost period of f, then

$$\sup_{x\in\mathbb{R}} |f(x+\tau) - f(x)| < \frac{\varepsilon}{2} \qquad \Longrightarrow \qquad f(x) \geq \frac{\varepsilon}{2} \qquad \text{for all} \qquad x \in [\tau - \varepsilon, \tau + \varepsilon].$$

Since $\tau \in [a, a+\Lambda]$, we have $I_a = [a, a+\Lambda] \cap [\tau - \varepsilon, \tau + \varepsilon]$ is non-empty and is an interval of length ε, consequently

$$\frac{1}{\Lambda} \int_a^{a+\Lambda} f(x)\,dx \geq \frac{1}{\Lambda} \int_{I_a} f(x)\,dx \geq \frac{\varepsilon^2}{\Lambda}.$$

It holds true for all $a \in \mathbb{R}$, let $a = n\Lambda$ where $n = 1, 2, \ldots$ and by corollary 3.14 we obtain

$$M(f) = \lim_{n\longrightarrow\infty} \frac{1}{n\Lambda} \int_0^{n\Lambda} f(x)\,dx = \lim_{n\longrightarrow\infty} \frac{1}{n} \left(\sum_{j=0}^{n-1} \frac{1}{\Lambda} \int_{j\Lambda}^{(j+1)\Lambda} f(x)\,dx \right) \geq \frac{\varepsilon^2}{\Lambda} > 0.$$

□

3.3 Pre-Hilbert space structure on AP($\mathbb{R}$)

Let $\mathcal{H}$ be a complex vector space, an inner product (or scalar product) on $\mathcal{H}$ is a map $(x, y) \longmapsto \langle x, y\rangle$ from $\mathcal{H} \times \mathcal{H} \longmapsto \mathbb{C}$ such that:

1. $\langle ax + by, z\rangle = a\langle x, z\rangle + b\langle y, z\rangle$ for all $x, y, z \in \mathcal{H}$ and $a, b \in \mathbb{C}$.
2. $\langle y, x\rangle = \overline{\langle x, y\rangle}$ for all $x, y \in \mathcal{H}$.
3. $\langle x, x\rangle \in (0, \infty)$ for all nonzero $x \in \mathcal{H}$.

A complex vector space equipped with an inner product is called a pre-Hilbert space. If $\mathcal{H}$ is a pre-Hilbert space, for $x \in \mathcal{H}$ we define $\|x\| = \sqrt{\langle x, x\rangle}$. By Schwartz inequality $\left|\langle x, y\rangle\right| \leq \|x\|\|y\|$ we deduce that $x \longmapsto \|x\|$ is a norm on $\mathcal{H}$. If $x, y \in \mathcal{H}$ and $\langle x, y\rangle = 0$, we say x is orthogonal to y and write $x \perp y$. If $E \subset \mathcal{H}$ then $E^\perp = \{x \in \mathcal{H} : \langle x, y\rangle = 0 \text{ for all } y \in E\}$ is a closed subspace of $\mathcal{H}$.

Proposition 3.16. *If $x_n \longrightarrow x$ and $y_n \longrightarrow y$ then $\langle x_n, y_n\rangle \longrightarrow \langle x, y\rangle$.*

Proposition 3.17 (The Parallelogram Law)**.** *For $x, y \in \mathcal{H}$ then $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$.*

Proposition 3.18 (The Pythagorean theorem)**.** *If $x_1, \ldots, x_n \in \mathcal{H}$ and $x_j \perp x_k$ for $j \neq k$ then*

$$\left\| \sum_{j=1}^{n} x_j \right\|^2 = \sum_{j=1}^{n} \|x_j\|^2.$$

A subset $\{u_\alpha\}_{\alpha\in A}$ of $\mathcal{H}$ is called othonormal if $\|u_\alpha\| = 1$ for all $\alpha \in A$ and $u_\alpha \perp u_\beta$ whenever $\alpha \perp \beta$.

Proposition 3.19 (Bessel's inequality)**.** *If $\{u_\alpha\}_{\alpha\in A}$ is an orthonomal set in $\mathcal{H}$, then for any $x \in \mathcal{H}$* [1]

$$\sum_{\alpha\in A} |\langle x, u_\alpha\rangle|^2 \leq \|x\|^2.$$

In particular, $\{\alpha \in A : \langle u_x, \alpha\rangle \neq 0\}$ is countable.

Theorem 3.20. *If $\{u_\alpha\}_{\alpha\in A}$ is an orthonormal set in $\mathcal{H}$ then the following are equivalent:*

(a) (Completeness) If $\langle x, u_\alpha\rangle = 0$ for all $\alpha \in A$ then $x \equiv 0$.

[1] Where $\sum_{x\in S} f(x)$ is the supremum of $\sum_{x\in E} f(x)$ over all finite subset $E \subset S$.

(b) (Parseval's Identity) $\|x\|^2 = \sum_{\alpha\in A} |\langle x, u_\alpha\rangle|^2$ *for all* $x \in \mathcal{H}$.

(c) For each $x \in \mathcal{H}$, $x = \sum_{\alpha\in A}\sum_{x\in\mathcal{H}} \langle x, u_\alpha\rangle u_\alpha$ *where the sum on the right has only countably many nonzero terms and converges in the norm topology no matter how these terms are ordered.*

1. ($\mathrm{AP}(\mathbb{R})$ **as a Pre-Hilbert space.**) Recall that $\mathrm{AP}(\mathbb{R})$ is a sub-algebra of $L^\infty(\mathbb{R})$, we now define the inner product on $\mathrm{AP}(\mathbb{R})$ by

$$\langle f, g\rangle_M := M(f\overline{g}) = \lim_{T\longrightarrow+\infty} \frac{1}{2T}\int_{-T}^{T} f(x)\overline{g(x)}\,dx \qquad \text{for all} \qquad f, g \in \mathrm{AP}(\mathbb{R}).$$

It is clear that this inner product is well-defined and hence $(\mathrm{AP}(\mathbb{R}), \langle\cdot,\cdot\rangle_M)$ is a pre-Hilbert space.

Proposition 3.21. *In the pre-Hilbert space* $\mathrm{AP}(\mathbb{R})$ *defined above, the exponentials* $\{e^{i\xi x}\}_{\xi\in\mathbb{R}}$ *form an orthonomal family.*

Proof. We have

$$\left\langle e^{i\xi x}, e^{i\eta x}\right\rangle_M = \lim_{T\longrightarrow\infty}\frac{1}{2T}\int_{-T}^{T} e^{i(\xi-\eta)x}\,dx = \begin{cases} 1 & \text{if } \xi = \eta \\ 0 & \text{if } \xi \neq \eta. \end{cases}$$

□

Let's introduce the notation[2]

$$\widehat{f}(\{\xi\}) := \left\langle f, e^{i\xi x}\right\rangle_M = M\left(f e^{-i\xi(\cdot)}\right).$$

Recall (34) we have

$$\xi \in \sigma(f) \qquad \Longleftrightarrow \qquad \widehat{f}(\{\xi\}) \neq 0. \tag{35}$$

In other words, $\widehat{f}(\{\xi\})$ are the Fourier coefficients of f relative to the orthonormal family $\{e^{i\xi x}\}_{\xi\in\widehat{\mathbb{R}}}$. The Bessel's inequality reads

$$\sum_{\xi\in\widehat{\mathbb{R}}} \left|\widehat{f}(\{\xi\})\right|^2 \leq \langle f, f\rangle_M = M\left(|f|^2\right).$$

It follows that $\left\{\xi \in \widehat{\mathbb{R}} : \widehat{f}(\{\xi\}) \neq 0\right\}$ is a countable set, thus together with 35 we have $\sigma(f)$ is countable for all $f \in \mathrm{AP}(\mathbb{R})$.

2. (**Convolution in** $\mathrm{AP}(\mathbb{R})$) We now introduce the mean convolution $f *_M g$ of two functions $f, g \in \mathrm{AP}(\mathbb{R})$ as following:

$$(f *_M g)(x) = M\left(\left(\tau_x\tilde{f}\right)g\right) = \lim_{T\longrightarrow\infty}\frac{1}{2T}\int_{-T}^{T} f(x-y)g(y)\,dy.$$

Since $\mathrm{AP}(\mathbb{R})$ is a sub-algebra of $L^\infty(\mathbb{R})$, $\left(\tau_x\tilde{f}\right)g \in \mathrm{AP}(\mathbb{R})$ provided $f, g \in \mathrm{AP}(\mathbb{R})$, thus the above definition is well-defined.

Proposition 3.22. *If* $f, g \in \mathrm{AP}(\mathbb{R})$ *then* $f *_M g \in \mathrm{AP}(\mathbb{R})$. *If* $M(|g|) \leq 1$ *then* $f *_M g \in W(f)$.

Proof. Let's consider $g \in \mathrm{AP}(\mathbb{R})$ with $M(|g|) < 1$. For each $n \in \mathbb{N}$ large, we define

$$g_n = \frac{1}{2n}\chi_{[-n,n]}g \in L^1(\mathbb{R}) \qquad \text{has} \qquad \|g_n\|_{L^1(\mathbb{R})} = \frac{1}{2n}\int_{-n}^{n}|g(x)|\,dx. \tag{36}$$

By the mean-value theorem, we have $\|g_n\|_{L^1} \longrightarrow M(|g|) < 1$ as $n \longrightarrow \infty$, thus for n large enough we have $\|g_n\|_{L^1} < 1$. Theorem 3.2 implies that

$$x \longmapsto \frac{1}{2n}\int_{-n}^{n} f(x-y)g(y)\,dy = f * g_n(x) \in W(f) \qquad \text{for all} \qquad n \text{ large}.$$

[2]By abuse of language we refer to $\widehat{f}(\{\xi\})$ for $f \in \mathrm{AP}(\mathbb{R})$ as the mass of the pseudo-measure $\widehat{f}$ at ξ.

The compactness of $W(f)$ implies that there exists a sub-sequence $\{n_k\} \subset \mathbb{N}$ such that $f * g_{n_k} \longrightarrow h$ uniformly in $L^\infty(\mathbb{R})$. On the other hand, for each $x \in \mathbb{R}$ we have

$$\left(f * g_{n_k}\right)(x) = \frac{1}{2n_k} \int_{-n_k}^{n_k} f(x-y)g(y)\, dy \longrightarrow (f *_M g)(x)$$

as $n_k \longrightarrow \infty$. Thus $f *_M g \equiv h$ point-wise, hence

$$\lim_{n \longrightarrow \infty} \|f * g_n - f *_M g\|_{L^\infty} = 0.$$

Thus $f *_M g \in W(f)$, hence it is uniformly almost periodic. The case $M(|g|) = 1$ and the general case follow from scaling and linearity of the mean value. □

Proposition 3.23. *For $f, g \in \mathrm{AP}(\mathbb{R})$ and $\xi \in \widehat{\mathbb{R}}$ then*

(a)

$$\widehat{(f *_M g)}(\{\xi\}) = \widehat{f}(\{\xi\})\widehat{g}(\{\xi\}).$$

(b)

$$\left(f *_M e^{i\xi(\cdot)}\right)(x) = \widehat{f}(\{\xi\})e^{i\xi x}.$$

As a consequence, if $g(x) = \sum_{j=1}^m \alpha_j e^{i\xi x}$ then

$$\left(f *_M e^{i\xi(\cdot)}\right)(x) = \sum_{j=1}^{n} \alpha_j \widehat{f}(\{\xi\})e^{i\xi x}.$$

Proof. By definition we have

$$\widehat{(f *_M g)}(\{\xi\}) = M\left((f *_M g)\, e^{-i\xi x}\right) = M_x\left[M_y\left(f(x-y)g(y)\right)e^{-i\xi x}\right].$$

Given $\varepsilon > 0$, by the property of mean value, there exists $n_1 \in \mathbb{N}$ such that

$$\left|\frac{1}{2n}\int_{-n}^{n} M_y\left(f(x-y)g(y)\right)e^{i\xi x}\, dx - M_x\left[M_y\left(f(x-y)g(y)\right)e^{-i\xi x}\right]\right| < \frac{\varepsilon}{2} \qquad \text{for} \qquad n \geq n_1.$$

Similarly, there exists $n_2 \in \mathbb{N}$ such that

$$\left|\frac{1}{2k}\int_{-k}^{k} f(x-y)g(y)\, dy - M_y\left(f(x-y)g(y)\right)\right| < \frac{\varepsilon}{2} \qquad \text{for} \qquad k \geq n_2$$

which implies that

$$\left|\frac{1}{2n}\int_{-n}^{n} M_y\left(f(x-y)g(y)\right)e^{-i\xi x}\, dx - \frac{1}{2n}\int_{-n}^{n}\left(\frac{1}{2k}\int_{-k}^{k} f(x-y)g(y)\, dy\right)e^{-i\xi x}\, dx\right| < \frac{\varepsilon}{2} \qquad \text{for} \qquad k \geq k_2.$$

From these facts we obtain

$$\left|\left(\widehat{f *_M g}\right)(\{\xi\}) - \frac{1}{2n}\int_{-n}^{n}\left(\frac{1}{2k}\int_{-k}^{k} f(x-y)g(y)\, dy\right)e^{-i\xi x}\, dx\right| < \varepsilon \tag{37}$$

for $n \geq n_1, k \geq n_2$. On the other hand by the mean value theorem there exist n_3, n_4 such that

$$\left|\widehat{f}(\{\xi\}) - \frac{1}{2n}\int_{-n}^{n} f(z)e^{-i\xi z}\, dz\right| < \frac{\varepsilon}{\|g\|_{L^\infty}} \qquad \text{for} \qquad n \geq n_3$$

and

$$\left|\widehat{g}(\{\xi\}) - \frac{1}{2k}\int_{-k}^{k} g(z)e^{-i\xi z}\, dz\right| < \frac{\varepsilon}{|\widehat{f}(\{\xi\})| + 1} \qquad \text{for} \qquad k \geq n_4.$$

By Fubini's theorem, the latter term in (37) can be written as

$$\frac{1}{2n}\int_{-n}^{n}\left(\frac{1}{2k}\int_{-k}^{k} f(x-y)g(y)\,dy\right)e^{-i\xi x}\,dx = \frac{1}{2k}\int_{-k}^{k} g(y)e^{-i\xi y}\left(\frac{1}{2n}\int_{-n}^{n} f(x-y)e^{-i\xi(x-y)}\right)dy$$

and thus

$$\left|\frac{1}{2n}\int_{-n}^{n}\left(\frac{1}{2k}\int_{-k}^{k} f(x-y)g(y)\,dy\right)e^{-i\xi x}\,dx - \left(\frac{1}{2k}\int_{-k}^{k} g(y)e^{-i\xi y}\,dy\right)\widehat{f}(\{\xi\})\right| \le \varepsilon$$

for all $n \ge n_3$, therefore

$$\left|\frac{1}{2n}\int_{-n}^{n}\left(\frac{1}{2k}\int_{-k}^{k} f(x-y)g(y)\,dy\right)e^{-i\xi x}\,dx - \widehat{g}(\{\xi\})\widehat{f}(\{\xi\})\right| \le \varepsilon \tag{38}$$

for $n \ge n_3, k \ge n_4$. From (37) and (38) we deduce that

$$\left|\left(\widehat{f *_M g}\right)(\{\xi\}) - \widehat{f}(\{\xi\})\widehat{g}(\{\xi\})\right| < 2\varepsilon$$

for all $\varepsilon > 0$ and hence the result follows. For the second part we have

$$\begin{aligned}
\left(f *_M e^{i\xi(\cdot)}\right)(x) &= \lim_{T\longrightarrow\infty}\frac{1}{2T}\int_{-T}^{T} f(x-y)e^{i\xi y}\,dy \\
(z = x-y) \qquad &= \left(\lim_{T\longrightarrow\infty}\frac{1}{2T}\int_{-T}^{T} f(z)e^{-i\xi z}\,dy\right)e^{i\xi x} = M\left(f e^{-i\xi(\cdot)}\right)e^{i\xi x} = \widehat{f}(\{\xi\})e^{i\xi x}.
\end{aligned}$$

The latter result follows from linearity of the mean value of functions in AP($\mathbb{R}$). □

3. Now for $f \in \mathrm{AP}(\mathbb{R})$, let's define $f^* = \overline{\tilde{f}}$, i.e., $f^*(x) = \overline{f(-x)}$ and

$$h = f *_M f^*, \qquad \text{i.e.,} \qquad h(x) = M_y\left(f(x+y)\overline{f(y)}\right).$$

By definition we have

$$\widehat{f^*}(\{\xi\}) = M\left(f^*(x)e^{-i\xi x}\right) = \lim_{T\longrightarrow\infty}\frac{1}{2T}\int_{-T}^{T}\overline{f(-x)}e^{-i\xi x}\,dx = \overline{\lim_{T\longrightarrow\infty}\frac{1}{2T}\int_{-T}^{T} f(-x)e^{i\xi x}\,dx} = \overline{\widehat{f}(\{\xi\})}$$

and thus by proposition 3.23

$$\widehat{h}(\{\xi\}) = \widehat{f}(\{\xi\})\widehat{f^*}(\{\xi\}) = \left|\widehat{f}(\{\xi\})\right|^2.$$

Also it is easy to see that if $\|f\|_{L^\infty} \le 1$ then $h \in W(f)$ by proposition 3.22.

Lemma 3.24. *The function $h = f^* *_M f$ is positive definite, in the sense that for every choice of $\xi_1,\ldots,\xi_m \in \mathbb{R}$ and complex numbers $\lambda_1,\ldots,\lambda_m$ we have*

$$\sum_{j,k=1}^{m} h(\xi_j-\xi_k)\lambda_j\overline{\lambda_k} \ge 0.$$

Proof. We have

$$\begin{aligned}
\sum_{j,k=1}^{m} h(\xi_j-\xi_k)\lambda_j\overline{\lambda_k} &= \lim_{T\longrightarrow\infty}\int_{-T}^{T}\left(\frac{1}{2T}\sum_{j,k=1}^{m} f(\xi_j+x)\overline{f(\xi_k+x)}\lambda_j\overline{\lambda_k}\right)dx \\
&= \lim_{T\longrightarrow\infty}\int_{-T}^{T}\left(\frac{1}{2T}\left|\sum_{j,k=1}^{m} f(\xi_j+x)\lambda_j\right|^2\right)dx \ge 0.
\end{aligned}$$

□

For $f \in \mathrm{AP}(\mathbb{R})$ then $h = f *_M f^* \in \mathrm{AP}(\mathbb{R})$ is continuous and positive definite. Bochner's theorem says that h is the Fourier transform of a positive measure or equivalently, $\widehat{h}$ is a positive measure.

4. ($\mathrm{AP}(\mathbb{R})$ **is a pre-Hilbert space**.)

Theorem 3.25 (Fourier inversion formula)**.** *If $f \in \mathrm{AP}(\mathbb{R})$ and $\widehat{f} \in M(\widehat{\mathbb{R}})$, then*

$$\widehat{f} = \sum_{\xi \in \widehat{\mathbb{R}}} \widehat{f}(\{\xi\})\delta_\xi \qquad \text{with} \qquad \|\widehat{f}\|_{M(\widehat{\mathbb{R}})} = \sum_{\xi \in \widehat{\mathbb{R}}} |\widehat{f}(\{\xi\})|$$

and we have an inversion formula

$$f(x) = \sum_{\xi \in \widehat{\mathbb{R}}} \widehat{f}(\{\xi\}) e^{i\xi x}.$$

Proof. Let's write $\widehat{f} = \mu + \nu$ then the discrete part of $\widehat{f}$ is $\nu = \sum_{\xi \in \widehat{\mathbb{R}}} \widehat{f}(\{\xi\})\delta_\xi$, where the series converges absolutely in the normed space $M(\widehat{\mathbb{R}})$. We claim the continuous part $\mu \equiv 0$. We have

$$\mu = \widehat{f} - \sum_{\xi \in \widehat{\mathbb{R}}} \widehat{f}(\{\xi\})\delta_\xi$$

Here we understand $\delta_\xi(x) = \delta_0(x - \xi)$ as measure in $M(\widehat{\mathbb{R}})$. Let $g(x) = \widehat{\mu}(-x)$ then $\widehat{g} \equiv \mu$ as distributions. As measure, we have $\widehat{\delta_\xi}(x) = e^{-i\xi x}$, and thus (as Fourier transform of measure)

$$g(x) = \widehat{\mu}(-x) = f(x) - \sum_{\xi \in \widehat{\mathbb{R}}} \widehat{f}(\{\xi\}) e^{i\xi x}$$

where the sum converges uniformly in $L^\infty(\mathbb{R})$, as $\mathcal{F} : M(\widehat{\mathbb{R}}) \longrightarrow \mathrm{BUC}(\mathbb{R})$ with $\|\widehat{\nu}\|_{L^\infty} \leq \|\nu\|_{M(\widehat{\mathbb{R}})}$ for $\nu \in M(\widehat{\mathbb{R}})$. Since $\mathrm{AP}(\mathbb{R})$ is an algebra and the sum converges uniformly in L^∞, we obtain $g \in \mathrm{AP}(\mathbb{R})$. Now by Wiener's theorem for measures in $M(\widehat{\mathbb{R}})$ we obtain

$$0 = \sum_{\xi \in \widehat{\mathbb{R}}} |\mu(\{\xi\})|^2 = \lim_{\lambda \longrightarrow \infty} \frac{1}{2\lambda} \int_{-\lambda}^{\lambda} |\widehat{\mu}(x)|^2 \, dx = \lim_{\lambda \longrightarrow \infty} \int_{-\lambda}^{\lambda} |g(x)|^2 \, dx = M(|g|^2).$$

Since $|g|^2 \in \mathrm{AP}(\mathbb{R})$ and $|g|^2 \geq 0$, theorem 3.15 concludes that $g \equiv 0$, thus $\mu \equiv \widehat{g} \equiv 0$ as a distribution, and thus as a measure. Thus we have proved that $\widehat{f} = \sum_{\xi \in \widehat{\mathbb{R}}} \widehat{f}(\{\xi\})\delta_\xi$ and hence $\|\widehat{f}\|_{\widehat{\mathbb{R}}} = \sum_{\xi \in \widehat{\mathbb{R}}} |\widehat{f}(\xi)|$ follows. The inversion formula follows from the fact that $g \equiv 0$. □

This theorem enables us to show the following version of Parseval's formula:

Theorem 3.26 (Parseval's formula)**.** *Let $f \in \mathrm{AP}(\mathbb{R})$ then*

$$\sum_{\xi \in \widehat{\mathbb{R}}} |\widehat{f}(\{\xi\})|^2 = M(|f|^2).$$

Proof. Let $h = f *_M f^*$ then $\widehat{h} \in M(\widehat{\mathbb{R}})$ and thus theorem 3.25 can be applied to deduce that

$$h(0) = \sum_{\xi \in \widehat{\mathbb{R}}} \widehat{h}(\xi) = \sum_{\xi \in \widehat{\mathbb{R}}} |\widehat{f}(\xi)|^2 = M(|f|^2).$$

□

From that we can conclude that $\mathrm{AP}(\mathbb{R})$ with the pre-Hilbert and $\{e^{i\xi(\cdot)}\}_{\xi \in \widehat{\mathbb{R}}}$ is a complete orthonormal basic for $\mathrm{AP}(\mathbb{R})$. Note that it is not a Hilbert space with the norm induced by the mean value. The uniqueness reads

Corollary 3.27 (Uniqueness)**.** *If $f \in \mathrm{AP}(\mathbb{R})$ and $f \not\equiv 0$ then $\sigma(f) \neq \emptyset$.*

For $f \in \mathrm{AP}(\mathbb{R})$, the series $\sum_{\xi\in\widehat{\mathbb{R}}}\widehat{f}(\{\xi\})e^{i\xi x}$, to which we refer as the Fourier series of f converges to f in the norm induced by the (mean valued) inner product $\langle\cdot,\cdot\rangle_M$.

5. We will show that the Fourier series of $f \in \mathrm{AP}(\mathbb{R})$ is indeed summable to f in the uniform norm.

Lemma 3.28. *Given a finite number of points $\xi_1,\dots,\xi_m \in \widehat{\mathbb{R}}$ and an $\varepsilon > 0$, there exists a trigonometric polynomial $\mathbf{B}$ having the following properties: $\mathbf{B}(x) \geq 0$, $M(\mathbf{B}) = 1$ and $\widehat{\mathbf{B}}(\{\xi_j\}) > 1-\varepsilon$ for $j = 1,2,\dots,m$.*

Proof. Let $\lambda_1,\dots,\lambda_q$ be a basis for $\xi_1,\dots,\xi_m$, that is, $\lambda_1,\dots,\lambda_q$ are linearly independent over $\mathbb{Q}$ and every ξ_j can be written in the form $\xi_j = \sum_{k=1}^q c_{j,k}\lambda_k$ for $j = 1,\dots,m$ where $c_{j,k}$ are integers. Let $\delta > 0$ such that $(1-\delta)^q > 1-\varepsilon$ and let

$$N > \delta^{-1}\max_{j,k}|c_{j,k}|.$$

Using the discrete Fejér kernel $\mathcal{K}_m(x) = \sum_{-m}^{m}\left(1-\frac{|j|}{m+1}\right)e^{ijx}$, let's define

$$\begin{aligned}\mathbf{B}(x) = \prod_{k=1}^{q}\mathcal{K}_N(\lambda_k x) &= \prod_{k=1}^{q}\left(\sum_{l_k=-N}^{N}\left(1-\frac{|l_k|}{N+1}\right)e^{il_k(\lambda_k x)}\right)\\ &= \sum_{|l_k|\leq N}\left(1-\frac{|l_1|}{N+1}\right)\dots\left(1-\frac{|l_q|}{N+1}\right)e^{i(l_1\lambda_1+\dots+l_q\lambda_q)x}.\end{aligned}$$

It is clear that $\mathbf{B}(x) \geq 0$ since it is the product of non-negative functions $\mathcal{K}_N(\lambda_k x)$. Since $\mathbf{B}(x)$ is a polynomial (which is quasi-periodic), a simple argument showing that it mean value is the constant term in it representation, which is the term corresponding to $(l_1,\dots,l_q) = (0,\dots,0)$, thus $M(\mathbf{B}) = \widehat{\mathbf{B}}(0) = 1$. Finally for each $j = 1,\dots,m$ we have

$$\widehat{\mathbf{B}}(\{\xi_j\}) = \widehat{\mathbf{B}}\left(\left\{\sum_{k=1}^{q}c_{j,k}\lambda_k\right\}\right) = M\left(\mathbf{B}e^{-i(c_{j,1}\lambda_1+\dots+c_{j,k}\lambda_k)x}\right) = \prod_{k=1}^{q}\left(1-\frac{|c_{j,k}|}{N+1}\right) > (1-\delta)^q > 1-\varepsilon$$

where we have used the fact that $\widehat{\mathbf{B}}(\{\xi_j\})$ is the constant in $\mathbf{B}(x)$ which corresponds to the case $(l_1,\dots,l_q) = (c_{j,1},\dots,c_{j,q})$, thus the proof is complete. □

Theorem 3.29. *Let $f \in \mathrm{AP}(\mathbb{R})$, then f can be approximated uniformly by trigonometric polynomials $P_n \in W(f)$.*

Proof. Since $\sigma(f)$ is countable, we can write it as $\{\xi_j\}_{j=1}^{\infty}$. For each $n \in \mathbb{N}$, let $\mathbf{B}_n$ be the polynomial described in the lemma 3.28 for $\xi_1,\dots,\xi_n$ and $\varepsilon = \frac{1}{n}$. For each $n \in \mathbb{N}$ we have $P_n = f *_M \mathbf{B}_n \in W(f)$ by proposition 3.22 and for each $\xi_j \in \sigma(f)$ then

$$\widehat{P_n}(\{\xi_j\}) = \widehat{f}(\{\xi_j\})\widehat{\mathbf{B}}_n(\{\xi_j\}) \longrightarrow \widehat{f}(\{\xi\}) \qquad \text{as} \qquad n \longrightarrow \infty.$$

If $\xi \notin \sigma(f)$ then $\widehat{P_n}(\{\xi\}) = \widehat{f}(\{\xi\}) = 0$ for all $n \in \mathbb{N}$. Now $\{P_n\}_{n\in\mathbb{N}}$ is a sequence in the compact space $W(f)$, thus it has limit points. Assume $P_{n_k} \longrightarrow u$ uniformly as $n_k \longrightarrow \infty$ where $u \in W(f)$, then

$$\left|\widehat{P}_{n_k}(\{\xi\}) - \widehat{u}(\{\xi\})\right| = \left|M\left(\left(P_{n_k}(x)-u(x)\right)e^{-i\xi x}\right)\right| \leq \lim_{T\longrightarrow\infty}\frac{1}{2T}\int_{-T}^{T}\left|P_{n_k}(x)-u(x)\right|\,dx \leq \|P_{n_k}-u\|_{L^\infty} \longrightarrow 0$$

as $n_k \longrightarrow \infty$. Thus $\widehat{u}(\{\xi\}) = \widehat{f}(\{\xi\})$ for all $\xi \in \mathbb{R}$, which implies $u \equiv f$ by the uniqueness. Thus there only one limit point and the convergence holds in the full sequence $P_n \longrightarrow f$ uniformly in $L^\infty(\mathbb{R})$. Note that P_n is a trigonometric polynomial by proposition 3.23. □

6. We have a simple lemma:

Lemma 3.30. *If f_n be a sequence in $\mathrm{AP}(\mathbb{R})$ and $f_n \longrightarrow f$ uniformly on $\mathbb{R}$, then $M(f_n) \longrightarrow M(f)$ as $n \longrightarrow \infty$.*

Proof. For any $T > 0$ we have

$$\left|\frac{1}{2T}\int_{-T}^{T}\Big(f_n(t)-f(t)\Big)\,dt\right| \le \frac{1}{2T}\int_{-T}^{T}|f_n(t)-f(t)|\,dt \le \|f_n-f\|_{L^\infty}$$

for all $n\in\mathbb{N}$. Given $\varepsilon>0$, we can choose n large such that $\|f_n-f\|_{L^\infty}<\varepsilon$, then in the limit as $T\longrightarrow\infty$ we obtain $\limsup|M(f_n)-M(f)|<\varepsilon$ and hence $M(f_n)\longrightarrow M(f)$ since ε is chosen arbitrary. □

The trigonometric polynomial that approximates $f\in\mathrm{AP}(\mathbb{R})$ in theorem 3.29 has the property

$$M(P_n)=\widehat{P_n}(\{0\})=\widehat{f *_M \mathbf{B}_n}(\{0\})=\widehat{f}(\{0\})\widehat{\mathbf{B}_n}(\{0\})=\widehat{f}(0)=M(f)$$

by proposition 3.23.

Theorem 3.31. *If $f\in\mathrm{AP}(\mathbb{R})$ with $M(f)=0$ then its anti-derivative*

$$F(x)=\int_0^x f(t)\,dt$$

satisfies

$$\lim_{|x|\longrightarrow\infty}\frac{F(x)}{x}=0.$$

Proof. We have

$$\lim_{|x|\longrightarrow\infty}\frac{F(x)}{x}=\lim_{|x|\longrightarrow\infty}\frac{1}{x}\int_0^x f(t)\,dt=M(f)=0.$$

□

An analog of theorem 1.1 is not true generally for AP($\mathbb{R}$). Anti-derivative of a almost periodic function with mean zero is not necessary be bounded. Indeed, since AP($\mathbb{R}$) is a Banach space under the uniform norm, if for all $f\in\mathrm{AP}(\mathbb{R})$ has mean zero then its anti-derivative is bounded, then the following operator is well-defined:

$$T:(\mathrm{AP}(\mathbb{R}),\|\cdot\|_{L^\infty})\longrightarrow L^\infty(\mathbb{R})$$

$$f\longmapsto Tf \qquad \text{where} \qquad Tf(x)=\int_0^x f(t)\,dt.$$

It is clear that T is linear. Let's $\mathrm{AP}_0(\mathbb{R})$ be the set of functions in AP($\mathbb{R}$) with mean value 0, we first see that $\mathrm{AP}_0(\mathbb{R})$ is a closed subspace of AP($\mathbb{R}$) under the uniform norm by lemma 3.30. We will use the Closed Graph Theorem to show that T is bounded. Indeed, if $f_n\longrightarrow f$ in AP($\mathbb{R}$) under the uniform norm and $Tf_n\longrightarrow g$ in $L^\infty(\mathbb{R})$, then for each $x\in\mathbb{R}$ we have

$$|Tf(x)-g(x)|\le\|Tf-Tf_n\|_{L^\infty}.|x|+\|Tf_n-g\|_{L^\infty}\longrightarrow 0$$

as $n\longrightarrow\infty$. Hence $Tf=g$ and by Closed Graph Theorem T is bounded, i.e. there exists a uniform constant C such that

$$\|Tf\|_{L^\infty}\le C\|f\|_{L^\infty}$$

for all $f\in\mathrm{AP}_0(\mathbb{R})$. This is absurd, since for example let us take $f(x)=e^{i\lambda x}$ then $Tf(x)=\frac{1}{i\lambda}e^{i\lambda x}$, hence

$$\|Tf\|_{L^\infty}\le C\|f\|_{L^\infty} \qquad \Longrightarrow \qquad \left|\frac{1}{\lambda}\right|\le C$$

for all $\lambda\neq 0$, which is clearly false.

A constructive example is as following:

Theorem 3.32. *Let us consider*

$$f(x)=\sum_{n=1}^{\infty}\frac{1}{n^2}e^{in^{-2}x}\in\mathrm{AP}_0(\mathbb{R}).$$

It is clear that the series converges absolutely in $L^\infty(\mathbb{R})$. The anti-derivative of f is well-defined but it is unbounded.

Proof. Let

$$F(x) = \int_0^x f(t)\,dt \qquad \text{and} \qquad F_m(x) = \int_0^x \sum_{n=1}^{m} \frac{1}{n^2} e^{in^{-2}t}\,dt = (-i)\sum_{n=1}^{m}\left(e^{in^{-2}x} - 1\right)$$

for $m \in \mathbb{N}$. Observe that by the Euler's formula we have

$$e^{in^2x} - 1 = 2i\sin\left(\frac{x}{2n^2}\right)\left[\sin\left(\frac{x}{2n^2}\right) + i\cos\left(\frac{x}{2n^2}\right)\right] = 2i\sin\left(\frac{x}{2n^2}\right)e^{-i\frac{x}{2n^2}}$$

which implies that

$$\left|e^{in^2x} - 1\right| \le 2\left|\sin\left(\frac{x}{2n^2}\right)\right| \le \frac{|x|}{n^2}$$

when $\frac{|x|}{2n^2} \le 1$. Thus $F_m(x) \longrightarrow F(x)$ as $m \longrightarrow \infty$ by the dominated convergence theorem, and the anti-derivative of f is given by

$$F(x) = \int_0^x f(t)\,dt = (-i)\sum_{n=1}^{\infty}\left(e^{in^{-2}x} - 1\right).$$

In order to see that F is unbounded, we only need to look at the modulus of the imaginary part of $F(x)$, i.e.

$$\operatorname{Im} F(x) = \sum_{n=1}^{\infty}\left(\cos\left(\frac{x}{n^2}\right) - 1\right) = -2\sum_{n=1}^{\infty}\sin^2\left(\frac{x}{2n^2}\right).$$

Thus it is enough show that that $\sum_{n=1}^{\infty}\sin^2\left(\frac{x}{2n^2}\right)$ is an unbounded function of x. Utilizing the inequality

$$|\sin x| \ge \frac{2|x|}{\pi} \qquad \text{for} \qquad |x| \le \frac{\pi}{2}$$

we deduce that for each x fixed, then for n large enough such that $\frac{|x|}{2n^2} \le \frac{\pi}{2}$ we have

$$\begin{aligned}\sum_{n=1}^{\infty}\sin^2\left(\frac{x}{2n^2}\right) &= \sum_{n=1}^{\left[\sqrt{|x|/\pi}\right]}\sin^2\left(\frac{x}{2n^2}\right) + \sum_{n=\left[\sqrt{|x|/\pi}\right]}^{\infty}\sin^2\left(\frac{x}{2n^2}\right) \\ &\ge \sum_{n=\left[\sqrt{|x|/\pi}\right]}^{\infty}\frac{|x|^2}{n^4\pi^2} = \frac{|x|^2}{\pi^2}\left(\sum_{n=\left[\sqrt{|x|/\pi}\right]}^{\infty}\frac{1}{n^4}\right).\end{aligned}$$

Although we know $\sum_{n=1}^{\infty}\frac{1}{n^4} = \frac{\pi^4}{90}$, it is still unclear why the sum above is unbounded in term of x since when $|x|$ gets larger, the sum gets smaller. Let's consider $x = k^2\pi$ with $k \in \mathbb{N}$ then

$$\frac{|x|^2}{\pi^2}\left(\sum_{n=\left[\sqrt{|x|/\pi}\right]}^{\infty}\frac{1}{n^4}\right) = k^4\sum_{n=k}\frac{1}{n^4} \ge k^4\int_k^{\infty}\frac{dx}{x^4} = \frac{k^4}{3k^3} = \frac{k}{3} \longrightarrow \infty$$

as $k \longrightarrow \infty$. Thus $F(x)$ is unbounded. □

3.4 Some sufficient conditions for functions to be almost-periodic

1. (**Bohr theorem**) We have the following criterion for function to be almost-periodic.

   Theorem 3.33 (Bohr)**.** *Let $f \in L^\infty(\mathbb{R})$ with its classical derivative $f' \in \mathrm{AP}(\mathbb{R})$, then $f \in \mathrm{AP}(\mathbb{R})$.*

   Proof. Since $\mathrm{AP}(\mathbb{R})$ is a sub-algebra, without loss of generality we can assume f is real-valued. Note that f is continuous already, thus all we need to do is to show that for given $\varepsilon > 0$, there exists

$\Lambda = \Lambda(\varepsilon, f)$ such that any interval of length Λ contains an ε-almost-period of f. Let $\alpha = \sup_{x\in\mathbb{R}} f(x)$ and $\beta = \inf_{x\in\mathbb{R}} f(x)$, let x_α, x_β be real numbers such that

$$f(x_\alpha) > \alpha - \frac{\varepsilon}{8}, \qquad f(x_\beta) < \beta + \frac{\varepsilon}{8}, \qquad \text{and} \qquad \delta = \frac{\varepsilon}{4|x_\alpha - x_\beta|}.$$

We claim that if τ is an δ-almost-period of f' then $f(x_\beta - \tau) \in \left[\beta, \beta + \frac{\varepsilon}{2}\right)$. Indeed we have

$$\begin{aligned} f(x_\alpha - \tau) - f(x_\beta - \tau) &= \int_{x_\beta}^{x_\alpha} f'(x-\tau)\,dx = \int_{x_\beta}^{x_\alpha} f'(x)\,dx + \int_{x_\beta}^{x_\alpha} \big(f'(x-\tau) - f'(x)\big)\,dx \\ &= f(x_\alpha) - f(x_\beta) + \int_{x_\beta}^{x_\alpha} \big(f'(x-\tau) - f'(x)\big)\,dx. \end{aligned}$$

Since τ is an δ-almost-period of f', the last term is bounded by $\frac{\varepsilon}{4}$, thus we deduce that

$$f(x_\alpha - \tau) - f(x_\beta - \tau) > \alpha - \beta - \frac{\varepsilon}{2} \quad \Longrightarrow \quad f(x_\beta - \tau) < m + \frac{\varepsilon}{2}. \tag{39}$$

Let $L = \Lambda\left(\frac{\delta}{2}, f'\right)$. For $x \in \mathbb{R}$, let $\rho \in (x - x_\beta, x - x_\beta + L)$ be an $\frac{\delta}{2}$-almost-period of f'. We have

$$\begin{aligned} f(x-\tau) - f(x) &= \big(f(x-\tau) - f(x_\beta - \tau - \rho)\big) + \big(f(x_\beta - \tau - \rho) - f(x_\beta - \rho)\big) + \big(f(x_\beta - \rho) - f(x)\big) \\ &= f(x_\beta - \tau - \rho) - f(x_\beta - \rho) + \int_x^{x_\beta - \rho} \big(f'(y) - f'(y-\tau)\big)\,dy \end{aligned}$$

for any $\tau \in \mathbb{R}$. Observe that if we choose $\eta = \min\left\{\frac{\varepsilon}{2L}, \frac{\delta}{2}\right\}$ and τ to be an η-almost-period of f', which is independent to x and ρ, then $\tau + \rho$ and ρ are δ-almost-periods of f', thus from (39) we have

$$f(x_\beta - \tau - \rho), f(x_\beta - \rho) \in \left[\beta, \beta + \frac{\varepsilon}{2}\right) \quad \Longrightarrow \quad \left|f(x_\beta - \tau - \rho) - f(x_\beta - \rho)\right| < \frac{\varepsilon}{2} \tag{40}$$

and

$$\left|\int_x^{x_\beta - \rho} \big(f'(y) - f'(y-\tau)\big)\,dy\right| < \eta(x_\beta - x - \rho) < \frac{\varepsilon}{2L} \times L = \frac{\varepsilon}{2} \tag{41}$$

From (40) and (41) we obtain $f|(x-\tau) - f(x)| < \varepsilon$ and hence every interval of length $\Lambda = \Lambda(\eta, f')$ contains an ε-almost-period of f. □

2. For $f \in AP(\mathbb{R})$ we say that $\widehat{f}$ is an *"almost-periodic pseudo-measure"*. We call a pseudo-measure $\nu \in \mathcal{F}L^\infty$ is *"almost-periodic"* at a point $\xi_0 \in \widehat{\mathbb{R}}$ if there exists a function $\varphi \in \mathcal{F}L^1$ with $\varphi \equiv 1$ in a neighborhood of ξ_0 such that $\varphi\nu$ is an almost-periodic pseudo-measure. The definition clearly implies that the set of points where ν is almost-periodic is an open set of $\widehat{\mathbb{R}}$.

Lemma 3.34. *$\nu \in \mathcal{F}L^\infty$ is almost-periodic at ξ_0 if and only if $\psi\nu$ is almost-periodic for all $\psi \in \mathcal{F}L^1$ with* supp ψ *is sufficiently close to ξ_0.*

Proof. If ν is almost-periodic at ξ_0, let $\varphi \in \mathcal{F}L^1$ with $\varphi \equiv 1$ on $(\xi_0 - \varepsilon, \xi_0 + \varepsilon)$ such that $\varphi\nu = \widehat{g}$ for some $g \in AP(\mathbb{R})$. Now for $\psi = \widehat{f}$ with $f \in L^1$ and supp $\psi \subset (\xi_0 - \varepsilon, \xi_0 + \varepsilon)$ we have $\varphi\psi = \psi$ on $(\xi_0 - \varepsilon, \xi_0 + \varepsilon)$ and thus $\psi\nu = \varphi(\psi\nu) = \psi(\varphi\nu) = \widehat{f * g}$ which is almost-periodic since $f * g \in AP(\mathbb{R})$ due to $f \in L^1(\mathbb{R})$ and $g \in AP(\mathbb{R})$. The inverse is obvious. □

Corollary 3.35. *If $\nu \in \mathcal{F}L^\infty$ then it is is almost-periodic at every $\xi \notin$* supp ν.

Proof. If $\xi \notin$ supp ν then there exists $\varepsilon > 0$ such that $(\xi - \varepsilon, \xi + \varepsilon) \cap \text{supp}\,\nu = \emptyset$. Let $\varphi \in C_c^\infty(\mathbb{R})$ be a function such that $\varphi = 1$ in a neighborhood of ξ and supp $\varphi \subset (\xi - \varepsilon, \xi + \varepsilon)$ then $\varphi\nu \equiv 0$ as a distribution, which belongs to $\mathcal{F}(AP(\mathbb{R}))$. □

Lemma 3.36. *Let $\nu \in \mathcal{F}L^\infty$ with* supp ν *is compact and ν is almost-periodic at every point of* supp ν, *then ν is almost-periodic.*

Proof. For each $\xi \in \operatorname{supp} \nu$ let $\varphi_\xi \in \mathcal{F}L^1$ such that $\varphi_\xi = 1$ on $(\xi - \varepsilon_\xi, \xi + \varepsilon_\xi)$ and $\varphi_\xi \nu$ is almost-periodic. Since $\operatorname{supp} \nu$ is compact, there is finite $\xi_1, \dots, \xi_m$ and $\varepsilon > 0$ such that $\operatorname{supp} \nu$ is covered by $\bigcup_{j=1}^m (\xi_j - \varepsilon, \xi_j + \varepsilon)$. Let $\varphi_j \in C_c^\infty(\mathbb{R})$ such that $\operatorname{supp} \varphi_j \subset (\xi_j - \varepsilon, \xi_j + \varepsilon)$ and $\sum_{j=1}^m \varphi_j = 1$. By lemma 3.34 $\varphi_j \nu = \widehat{f_j}$ for $f_j \in \mathrm{AP}(\mathbb{R})$ and thus $\nu = \left(\sum_{j=1}^m \varphi_j\right)\nu = \sum_{j=1}^m \varphi_j \nu = \sum_{j=1}^m \widehat{f_j}$ which is almost periodic. □

We now claim that Bohr theorem 3.33 implies the following criterion.

Theorem 3.37. *Let $f \in L^\infty(\mathbb{R})$ such that $\operatorname{supp} \widehat{f}$ is compact and $\widehat{f}$ is almost-periodic at every $\xi \in \widehat{\mathbb{R}}$ except, possibly, at $\xi = 0$ then $f \in \mathrm{AP}(\mathbb{R})$.*

Proof. First we claim that if $\operatorname{supp} \widehat{f}$ is compact then $f \in C^1(\mathbb{R})$ and $\widehat{f'} = i\xi\widehat{f}$. Let $\varphi \in \mathcal{S}(\mathbb{R})$ be a test function such that $\widehat{\varphi}(\xi) = 1$ in a neighborhood of $\operatorname{supp} \widehat{f}$, then $\widehat{f} = \widehat{\varphi}\widehat{f}$ and thus $f = \varphi * f$. Since f is bounded and $\varphi \in \mathcal{S}(\mathbb{R})$, we have $f = \varphi * f$ is smooth and in particular $f' = \varphi' * f$, therefore $\widehat{f'}(\xi) = \widehat{\varphi'}(\xi)\widehat{f}(\xi) = i\xi\widehat{\varphi}(\xi)\widehat{f}(\xi) \equiv i\xi\widehat{f}(\xi)$ since $\widehat{\varphi} = 1$ in a neighborhood of $\operatorname{supp} \widehat{f}$.

Since $\widehat{\varphi'}(0) = 0$, and $\{\psi \in \mathcal{S}(\widehat{\mathbb{R}})) : 0 \notin \operatorname{supp} \psi\}$ is dense in $\{\psi \in \mathcal{F}L^1(\widehat{\mathbb{R}}) : \psi(0) = 0\}$, there exists $\{\psi_n\} \subset \mathcal{S}(\mathbb{R})$ such that $\widehat{\varphi_n} = 0$ in a neighborhood of 0 and $\|\widehat{\psi_n} - \widehat{\varphi'}\|_{\mathcal{F}L^1} \longrightarrow 0$ as $n \longrightarrow \infty$, therefore

$$\|\psi_n * f - f'\|_{L^\infty(\mathbb{R})} = \left\|\widehat{\psi_n}\widehat{f} - \widehat{\varphi'}\widehat{f}\right\|_{\mathcal{F}L^\infty} \leq \|\widehat{f}\|_{\mathcal{F}L^\infty}\|\widehat{\psi_n} - \widehat{\varphi'}\|_{\mathcal{F}L^1} \longrightarrow 0$$

as $n \longrightarrow \infty$. Now we claim that $\mu_n = \psi_n * f$ is almost-periodic, or equivalently, $\widehat{\mu_n} = \widehat{\psi_n}\widehat{f}$ is almost-periodic as a pseudo-measure. Since $\operatorname{supp} \widehat{\psi_n}$ is compact and is supported away from 0, by lemma 3.36 we only need to show that $\widehat{\mu_n}$ is almost-periodic at every point $\xi \neq 0$ in the support of $\widehat{f}$. Let $\xi \in \operatorname{supp} \widehat{f}$ with $\xi \neq 0$, since $\widehat{f}$ is almost-periodic at ξ, there exists $\widehat{\chi} \in \mathcal{F}L^1$ with $\widehat{\chi} = 1$ in a neighborhood of ξ such that $\widehat{\chi}\widehat{f} = \widehat{g}$ for some $g \in \mathrm{AP}(\mathbb{R})$. Since

$$\widehat{\chi}\widehat{\mu_n} = \widehat{\chi}\widehat{\psi_n}\widehat{f} = \widehat{\psi * g}$$

therefore by definition we have $\widehat{\mu_n}$ is almost-periodic. Thus the pseudo-measure $\mu_n \in \mathrm{AP}(\mathbb{R})$ for all $n \in \mathbb{N}$, hence $f' \in \mathrm{AP}(\mathbb{R})$ since $f' = \lim \mu_n$ in $L^\infty(\mathbb{R})$. Finally Bohr theorem 3.33 concludes that $f \in \mathrm{AP}(\mathbb{R})$. □

The point $\xi = 0$ in the theorem above plays no specific role. In fact by the same argument we can show the following.

Lemma 3.38. *If $\nu \in \mathcal{F}L^\infty$ is almost-periodic for all $\xi \in (\xi_0 - \varepsilon, \xi_0 + \varepsilon)\backslash\{\xi_0\}$ then ν is also almost-periodic at ξ_0. In other word, the set of points where μ is not almost-periodic has no isolated point.*

Proof. Let $\widehat{\psi} \in C_c^\infty(\widehat{\mathbb{R}})$ with $\operatorname{supp} \widehat{\psi} \subset (\xi_0 - \varepsilon, \xi_0 + \varepsilon)$ and $\widehat{\psi} = 1$ in a neighborhood of ξ_0, we only need to show that $\widehat{\psi}\nu$ is almost-periodic as a pseudo-measure.

- The pseudo-measure $\widehat{\psi}\nu$ has supp $\subset \operatorname{supp} \widehat{\psi} \subset (\xi_0 - \varepsilon, \xi_0 + \varepsilon)$, thus $\widehat{\psi}\nu$ has compact support and it is almost-periodic for all $\xi \notin \widehat{\mathbb{R}}\backslash\operatorname{supp} \widehat{\psi}$.
- For $\xi \in \operatorname{supp} \widehat{\psi} \subset (\xi_0 - \varepsilon, \xi_0 + \varepsilon)$ and $\xi \neq \xi_0$, since ν is almost-periodic at ξ, there exists $\widehat{\varphi} \in \mathcal{F}L^1$ such that $\widehat{\varphi} = 1$ in a neighborhood of ξ and $\widehat{\varphi}\nu = \widehat{g}$ for some $g \in \mathrm{AP}(\mathbb{R})$, then $\widehat{\varphi}(\widehat{\psi}\nu) = \widehat{\psi}(\widehat{\varphi}\nu) = \widehat{\psi * g}$ is almost-periodic, thus $\widehat{\psi}\nu$ is almost-periodic at ξ by definition.

Thus $\widehat{\psi}\nu$ has compact support and is almost-periodic for all $\xi \in \widehat{\mathbb{R}}\backslash\{\xi_0\}$. By theorem 3.37 we conclude that $\widehat{\psi}\nu$ is almost-periodic as a pseudo-measure. □

Theorem 3.39. *Let $f \in L^\infty(\mathbb{R})$ such that $\operatorname{supp} \widehat{f}$ is compact and countable, then $f \in \mathrm{AP}(\mathbb{R})$.*

Proof. From lemma 3.38 we see that $\mathcal{A} = \{\xi \in \widehat{\mathbb{R}} : \widehat{f}$ is not almost-periodic at $\xi\}$ then clearly $\mathcal{A} \subset \operatorname{supp} \widehat{f}$ has no isolated point and is countable. It is closed since its complement, the set of points ξ such that ν is almost-periodic at ξ is open. Thus $\mathcal{A}$ is a countable perfect set, which has to be empty. Thus theorem 3.37 concludes that ν is almost-periodic as a pseudo-measure. □

3. Theorem 3.39 can be improved by replacing the condition supp $\widehat{f}$ is compact by a weaker condition $f \in \mathrm{BUC}(\mathbb{R})$.

   Theorem 3.40. *If* $f \in \mathrm{BUC}(\mathbb{R})$ *such that* supp $\widehat{f}$ *is countable then* $f \in \mathrm{AP}(\mathbb{R})$.

   Proof. Let $\mathcal{K}_\lambda$ be the Fejér kernel as usual, then for each $\lambda > 0$ the $L^\infty(\mathbb{R})$ function $g_\lambda = \mathcal{K}_\lambda * f$ satisfies $\widehat{g_\lambda} = \widehat{\mathcal{K}_\lambda}\widehat{f}$ and thus supp $\widehat{g_\lambda} \subset$ supp $\widehat{\mathcal{K}_\lambda} \cap$ supp $\widehat{f}$ which is compact and countable. By theorem 3.39 we have $\mathcal{K}_\lambda * f \in \mathrm{AP}(\mathbb{R})$ for all $\lambda > 0$. Since $\mathrm{AP}(\mathbb{R})$ is closed in $L^\infty(\mathbb{R})$, there exists a sub-sequence $\lambda_n \longrightarrow \infty$ such that $\mathcal{K}_{\lambda_n} * f \longrightarrow g$ as $\lambda_n \longrightarrow \infty$ for some $g \in \mathrm{AP}(\mathbb{R})$. It is clear that $\mathcal{K}_\lambda * f \longrightarrow f$ point-wise in $\mathbb{R}$ as $\lambda \longrightarrow \infty$, which implies $g = f$ and hence $f \in \mathrm{AP}(\mathbb{R})$. Indeed it is easy to see that $\|\mathcal{K}_\lambda * f - f\|_{L^\infty(\mathbb{R})} \longrightarrow 0$ as $\lambda \longrightarrow \infty$ for the full sequence. □

4. In fact theorem 3.37 and Bohr theorem 3.33 are equivalent. Indeed, let $f \in L^\infty(\mathbb{R})$ with its classical derivative $f' \in \mathrm{AP}(\mathbb{R})$ we will show that $f \in \mathrm{AP}(\mathbb{R})$.

   - Assume that $\widehat{f}$ has compact support first.
     - It is clear that $\widehat{f'} = i\xi\widehat{f}$ as pseudo-measures. In fact if $\varphi \in \mathcal{S}(\mathbb{R})$ is a function such that $\widehat{\varphi} = 1$ in a neighborhood of supp $\widehat{f}$, then $\widehat{f} = \widehat{\varphi}\widehat{f}$ and hence $f = \varphi * f$ is smooth and $f' = \varphi' * f$, therefore $\widehat{f'} = \widehat{\varphi'}\widehat{f} = i\xi\widehat{f}$.
     - Now we will show that $\widehat{f}$ is almost-periodic for all $\xi \neq 0$ in the support of $\widehat{f}$.
     - Let ξ_0 be such a point and $0 \notin (\xi_0 - \varepsilon, \xi_0 + \varepsilon)$ for some $\varepsilon > 0$, we can choose $\eta \in \mathcal{S}(\mathbb{R})$ such that supp $\widehat{\eta} \subset (\xi_0 - \varepsilon, \xi_0 + \varepsilon)$ and $\widehat{\eta} = 1$ on $\left(\xi_0 - \frac{\varepsilon}{2}, \xi_0 + \frac{\varepsilon}{2}\right)$.
     - Clearly we have $\xi \longmapsto \frac{1}{i\xi}\widehat{\eta}(\xi)$ belongs to $C_c^\infty(\widehat{\mathbb{R}})$, thus in turns we can find $\chi \in \mathcal{S}(\mathbb{R})$ such that $\widehat{\chi}(\xi) = \frac{1}{i\xi}\widehat{\eta}(\xi)$.
     - We have
$$\widehat{\chi}\widehat{f'} = \frac{1}{i\xi}\widehat{\eta}\left(i\xi\widehat{f}\right) = \widehat{\eta}\widehat{f} \qquad \Longrightarrow \qquad \widehat{\eta}\widehat{f} = \widehat{\chi * f'}$$
       is almost-periodic as a pseudo-measure.

     Thus we have proved that $\widehat{f}$ is almost periodic at every ξ in the compact of $\widehat{f}$ except 0. By theorem 3.37 we obtain $f \in \mathrm{AP}(\mathbb{R})$.
   - Now in the general case, for each $\lambda > 0$ the $L^\infty(\mathbb{R})$ function $g_\lambda = \mathcal{K}_\lambda * f$ is smooth, bounded with $g'_\lambda = \mathcal{K}_\lambda * f' \in \mathrm{AP}(\mathbb{R})$. Also $\widehat{g_\lambda}$ has compact support. Thus by the result above $g_\lambda \in \mathrm{AP}(\mathbb{R})$, and thus since $g_\lambda \longrightarrow f$ everywhere as $\lambda \longrightarrow 0$ and $\mathrm{AP}(\mathbb{R})$ is closed in $L^\infty(\mathbb{R})$ we deduce that the uniform limit $f = \lim_{\lambda \longrightarrow \infty} g_\lambda$ also belongs to $\mathrm{AP}(\mathbb{R})$.

4 Kronecker's theorem

We first establish the equivalence between the following two theorems:

Theorem 4.1 (Kronecker's theorem)**.** *Let* $\lambda_1, \ldots, \lambda_n$ *be real numbers, independent over the rationals. Let* $\alpha_1, \ldots, \alpha_n$ *be real numbers and* $\varepsilon > 0$. *Then there exists a real number* x *such that*

$$\left|e^{i\lambda_j x} - e^{i\alpha_j}\right| < \varepsilon \qquad \textit{for all} \qquad j = 1, 2, \ldots, n.$$

Theorem 4.2. *Let* $\lambda_1, \ldots, \lambda_n$ *be real numbers, independent over rationals,* $\lambda_0 = 0$, *and let* $a_0, a_1, \ldots, a_n$ *be any complex numbers. Then*

$$\sup_{x \in \mathbb{R}} \left|\sum_{j=0}^{n} a_j e^{i\lambda_j x}\right| = \sum_{j=0}^{n} |a_j|.$$

Proof of (4.1) *implies* (4.2). It is obvious that $\sup_{x\in\mathbb{R}}\left|\sum_{j=0}^n a_j e^{i\lambda_j x}\right| \le \sum_{j=0}^n |a_j|$. For the converse, given $\varepsilon > 0$, if we write $a_j = |a_j|e^{i\xi_j}$ for $j = 0,\ldots,n$ then by theorem (4.1) there exists $x \in \mathbb{R}$ such that

$$\left|e^{i\lambda_j x} - e^{i(\xi_0-\xi_j)}\right| < \varepsilon \qquad \text{for} \qquad j = 1,\ldots,n$$

which implies that

$$\begin{aligned}
\left|\left|\sum_{j=0}^n a_j e^{i\lambda_j x}\right| - \sum_{j=0}^n |a_j|\right| &\le \left|\sum_{j=0}^n a_j e^{i\lambda_j x} - e^{i\xi_0}\sum_{j=0}^n |a_j|\right| \\
&= \left|\sum_{j=0}^n |a_j| e^{i(\lambda_j x+\xi_j)} - e^{i\xi_0}\sum_{j=0}^n |a_j|\right| \\
&\le \sum_{j=0}^n |a_j|.\left|e^{i(\lambda_j x+\xi_j)} - e^{i\xi_0}\right| = \sum_{j=0}^n |a_j|.\left|e^{i\lambda_j x} - e^{i(\xi_0-\xi_j)}\right| < \varepsilon\left(\sum_{j=0}^n |a_j|\right)
\end{aligned}$$

and thus the proof is complete since ε is arbitrary. □

Proof of (4.2) *implies* (4.1). Consider the polynomial $1 + \sum_1^n e^{-i\xi_j} e^{i\lambda_j x}$, by theorem 4.2 for any $\varepsilon > 0$ we can choose $x \in \mathbb{R}$ such that

$$n+1 \ge \left|1 + \sum_{j=1}^n e^{i(\xi_j+\lambda_j x)}\right| > n+1-\frac{\varepsilon}{2}.$$

By a simple argument, each component must be closed to 1, therefore theorem 4.1 follows. □

Now we prove theorem 4.2, first for a simple case.

Theorem 4.3. *Let $\lambda_1,\ldots,\lambda_n$ be real numbers having the following properties:*

(a) $\sum_{j=1}^n c_j\lambda_j = 0$, $c_j \in \{-1,0,1\}$ implies $c_j = 0$ for all $j = 1,2,\ldots,n$.

(b) $\sum_{j=1}^n c_j\lambda_j = \lambda_k$, $c_j \in \{-1,0,1\}$ implies $c_j = 0$ for all $j \ne k$.

Then for any complex numbers $a_1,\ldots,a_n$ we have

$$\sup_{x\in\mathbb{R}}\left|\sum_{j=1}^n a_j e^{i\lambda_j x}\right| \ge \frac{1}{2}\sum_{j=1}^n |a_j|.$$

Proof. Let $a_j = r_j e^{i\xi_j}$ where $r_j = |a_j|$ and

$$g(x) = \prod_{j=1}^n \left(1+\cos\left(\lambda_j x + \xi_j\right)\right), \qquad \text{and} \qquad f(x) = \sum_{j=1}^n a_j e^{i\lambda_j x} = \sum_{j=1}^n r_j e^{i(\lambda_j x+\xi_j)}.$$

We have g is a non-negative trigonometric polynomial whose frequencies all have the form $\sum_{j\in S}\left(c_j\lambda_j x\right)$, i.e.,

$$g(x) = 1 + \sum_{S_k}\left(A_k e^{i\left(\sum_{j\in S_k} c_j\lambda_j\right)x}\right) \qquad \text{where} \qquad S_k \text{ is some subset of } \{1,2,\ldots,n\}, \quad A_k = \frac{1}{2^{|S_k|}} e^{i\left(\sum_{j\in S_k} c_j\xi_j\right)}.$$

It is easy to see that if $\xi \ne 0$, then the mean value of $e^{i\xi x}$ is

$$M\left(e^{i\xi x}\right) = \lim_{T\longrightarrow\infty}\frac{1}{T}\int_0^T e^{i\xi x}\,dx = 0.$$

Therefore, all the terms which contribute to the mean value of $g(x)$ is the one whose frequency is zero, that is

$$\sum_{j=1}^{n} c_j\lambda_j = 0, \qquad c_j \in \{-1,0,1\} \qquad \Longrightarrow \qquad c_j = 0 \quad \text{for all} \quad j = 1,2,\dots,n.$$

by (a), which is cannot the case. Thus $M(g) = M(|g|) = 1$. By the same argument, we have

$$fg(x) = \left(\sum_{j=1}^{n} a_j e^{i\lambda_j x}\right)\left[1 + \sum_{k=1}^{n}\left(A_k e^{i\left(\sum_{j\in S_k} c_j\lambda_j\right)x}\right)\right].$$

This is a trigonometric, and its frequency is zero if and only if

$$\sum_{j=1}^{n} c_j\lambda_j = \lambda_k, \qquad c_j \in \{-1,0,1\} \qquad \Longrightarrow \qquad c_j = 0 \quad \text{for all} \quad j \neq k.$$

Therefore the mean value, which is the constant term in fg is $\frac{1}{2}\sum_{j=1}^{n} r_j = \frac{1}{2}\sum_{j=1}^{n}|a_j|$. Finally, we have

$$M(fg) \le \|f\|_{L^\infty} M(|g|) = \|f\|_{L^\infty} M(g) \qquad \Longrightarrow \qquad \frac{1}{2}\sum_{j=1}^{n}|a_j| \le \sup_{x\in\mathbb{R}}\left|\sum_{j=1}^{n} a_j e^{i\lambda_j x}\right|$$

and the proof is complete. The inequality we used above is obvious since

$$M(fg) = \lim_{T\longrightarrow\infty}\frac{1}{T}\int_0^T f(x)g(x)\,dx \le \|f\|_{L^\infty}\left(\lim_{T\longrightarrow\infty}\frac{1}{T}\int_0^T |g(x)|\,dx\right) = \|f\|_{L^\infty} M(|g|).$$

□

This theorem is a special case of the following theorem.

Theorem 4.4. *Let $\lambda_1,\dots,\lambda_n$ be real numbers having the following properties:*

(a) $\sum_{j=1}^{n} c_j\lambda_j = 0$, $c_j \in \mathbb{Z}, |c_j| \le N$ *implies* $c_j = 0$ *for all* $j = 1,2,\dots,n$.

(b) $\sum_{j=1}^{n} c_j\lambda_j = \lambda_k$, $c_j \in \mathbb{Z}, |c_j| \le N$ *implies* $c_j = 0$ *for all* $j \neq k$.

Then for any complex numbers $a_1,\dots,a_n$ we have

$$\sup_{x\in\mathbb{R}}\left|\sum_{j=1}^{n} a_j e^{i\lambda_j x}\right| \ge \left(1 - \frac{1}{N+1}\right)\sum_{j=1}^{n}|a_j|.$$

Proof. It is similar to the proof of the last one, let $a_j = r_j e^{i\xi_j}$, we define

$$g(x) = \sum_{j=1}^{n} \mathcal{K}_N\left(\lambda_j x + \xi_j\right) \qquad \text{where} \qquad \mathcal{K}_N(x) = \sum_{k=-N}^{N}\left(1 - \frac{|k|}{N+1}e^{ikx}\right)$$

is the discrete Fejer's kernel and the proof follows in the same manner. □

Finally, in theorem 4.2, since $\lambda_1,\dots,\lambda_n$ are linearly independent over rationals, the conditions in our last theorem are satisfied for all $N \in \mathbb{N}$, thus

$$\sup_{x\in\mathbb{R}}\left|\sum_{j=1}^{n} a_j e^{i\lambda_j x}\right| \ge \left(1 - \frac{1}{N+1}\right)\sum_{j=1}^{n}|a_j|$$

for all $N \in \mathbb{N}$. Letting $N \longrightarrow \infty$ we obtain theorem 4.2 and hence 4.1.

Some problems

1.5. Let $f \in L^1(\mathbb{T})$ and m be a positive integer and write $f_{(m)}(t) = f(mt)$, we have

$$\widehat{f_{(m)}}(n) = \begin{cases} \widehat{f}\left(\frac{n}{m}\right) & \text{if} \quad m|n, \\ 0 & \text{if} \quad m \not| n. \end{cases}$$

Proof. We have

$$\begin{aligned}
\widehat{f_{(m)}}(n) &= \frac{1}{2\pi}\int_0^{2\pi} f(mt)e^{-int}\,dt \\
&= \frac{1}{2m\pi}\int_0^{2m\pi} f(s)e^{-i\frac{n}{m}s}\,ds \\
&= \frac{1}{2m\pi}\sum_{k=0}^{m-1}\int_{k2\pi}^{(k+1)2\pi} f(s)e^{-i\frac{n}{m}s}\,ds \\
&= \frac{1}{2m\pi}\sum_{k=0}^{m-1}\int_0^{2\pi} f(u)e^{-i\frac{n}{m}(u+k2\pi)}\,du = \frac{1}{m}\left(\sum_{k=0}^{m-1} e^{-i\frac{n}{m}2k\pi}\right)\left(\frac{1}{2\pi}\int_0^{2\pi} f(u)e^{-i\frac{n}{m}u}\,du\right).
\end{aligned}$$

Note that

$$\frac{1}{m}\sum_{k=0}^{m-1} e^{-i\frac{n}{m}2k\pi} = \frac{1-e^{-in2\pi}}{1-e^{-i2\pi\frac{n}{m}}} = 0 \ \text{ if } \quad m \not| n, \qquad \text{and} \qquad \frac{1}{m}\sum_{k=0}^{m-1} e^{-i\frac{n}{m}2k\pi} = 1 \ \text{ if } \quad m|n.$$

The proof is complete from this formula. □

2.8. (Fejer's lemma) If $f \in L^1(\mathbb{T})$ and $g \in L^\infty(\mathbb{T})$ then

$$\lim_{n\longrightarrow\infty} \frac{1}{2\pi}\int_0^{2\pi} f(t)g(nt)\,dt = \widehat{f}(0)\widehat{g}(0).$$

Proof. Let $P = \sum_{k=-m}^{m} a_k e^{-ikt}$ be a trigonometric polynomial with $\|f-P\|_{L^1(\mathbb{T})} < \varepsilon$, we have

$$\frac{1}{2\pi}\int_0^{2\pi} P(t)g_{(n)}(t)\,dt = \sum_{k=-m}^{m} a_k\left(\frac{1}{2\pi}\int_0^{2\pi} g_{(n)}e^{-int}\,dt\right) = \sum_{k=-m}^{m} a_k\widehat{g_{(n)}}(k)$$

where $g_{(n)}(t) = g(nt)$. Recall that for $n \neq 0$ then $\widehat{g_{(n)}}(k) = \widehat{g}\left(\frac{k}{n}\right)$ if $n|k$ and $\widehat{g_{(n)}}(k) = 0$ otherwise, if we choose $|n| > m$ then $\widehat{g_{(n)}}(k) = 0$ for all $|k| \leq m$, hence

$$\frac{1}{2\pi}\int_0^{2\pi} P(t)g_{(n)}(t)\,dt = a_0\widehat{g}(0).$$

Now since $|\widehat{f}(0) - \widehat{P}(0)| = |\widehat{f}(0) - a_0| \leq \|f-P\|_{L^1(\mathbb{T})} < \varepsilon$, we obtain

$$\left|\frac{1}{2\pi}\int_0^{2\pi} f(t)g(nt)\,dt - \widehat{f}(0)\widehat{g}(0)\right| \leq \frac{1}{2\pi}\int_0^{2\pi} |f(t)-P(t)||g(nt)\,dt + |a_0 - \widehat{f}(0)|.|\widehat{g}(0)| \leq \|g\|_{L^\infty}\varepsilon + \|g\|_{L^1(\mathbb{T})}\varepsilon.$$

Here we used the fact that $L^\infty(\mathbb{T}) \subset L^1(\mathbb{T})$. □

References

[1] Yitzhak Katznelson, An Introduction to Harmonic Analysis (Cambridge Mathematical Library) 3rd Edition, 1968.

[2] Rudin, Walter, Real and Complex Analysis (Higher Mathematics Series) 3rd Edition. McGraw-Hill, 1966.

[3] Brezis, Haim, Functional Analysis, Sobolev Spaces and Partial Differential Equations, Springer 2011.

[4] Gerald B. Folland, Real Analysis: Modern Techniques and Their Applications 2nd Edition, Wiley, 1999.