Gaussian measures and Bochner's theorem

Jordan Bell
jordan.bell@gmail.com
Department of Mathematics, University of Toronto

April 30, 2015

1 Fourier transforms of measures

Let m_n be *normalized Lebesgue measure* on $\mathbb{R}^n$: $dm_n(x) = (2\pi)^{-n/2}dx$. If μ is a finite positive Borel measure on $\mathbb{R}^n$, the *Fourier transform of* μ is the function $\hat{\mu} : \mathbb{R}^n \to \mathbb{C}$ defined by

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^n} e^{-i\xi\cdot x} d\mu(x), \qquad \xi \in \mathbb{R}^n.$$

One proves using the dominated convergence theorem that $\hat{\mu}$ is continuous. If $f \in L^1(\mathbb{R}^n)$, the Fourier transform of f is the function $\hat{f} : \mathbb{R}^n \to \mathbb{C}$ defined by

$$\hat{f}(\xi) = \int_{\mathbb{R}^n} e^{-i\xi\cdot x} f(x) dm_n(x), \qquad \xi \in \mathbb{R}^n.$$

Likewise, using the dominated convergence theorem, $\hat{f}$ is continuous. One proves that if $f \in L^1(\mathbb{R}^n)$ and $\hat{f} \in L^1(\mathbb{R}^n)$ then, for almost all $x \in \mathbb{R}^n$,

$$f(x) = \int_{\mathbb{R}^n} e^{ix\cdot\xi} \hat{f}(\xi) dm_n(\xi).$$

As

$$\hat{\mu}(0) = \int_{\mathbb{R}^n} d\mu(x) = \mu(\mathbb{R}^n),$$

μ is a probability measure if and only if $\hat{\mu}(0) = 1$. (By a probability measure we mean a positive measure with mass 1.)

If $\phi \in L^1(\mathbb{R}^n)$ and $\hat{\phi} \in L^1(\mathbb{R}^n)$, then, inverting the Fourier transform,

$$\begin{aligned}
\langle \phi, \mu \rangle &= \int_{\mathbb{R}^n} \phi(x) d\mu(x) \\
&= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \hat{\phi}(\xi) e^{ix\cdot\xi} dm_n(\xi) \right) d\mu(x) \\
&= \int_{\mathbb{R}^n} \hat{\phi}(\xi) \int_{\mathbb{R}^n} e^{i\xi\cdot x} d\mu(x) dm_n(\xi) \\
&= \int_{\mathbb{R}^n} \hat{\phi}(\xi) \hat{\mu}(-\xi) dm_n(\xi) \\
&= \int_{\mathbb{R}^n} \hat{\phi}(-\xi) \hat{\mu}(\xi) dm_n(\xi).
\end{aligned}$$

Theorem 1. *If μ and ν are finite Borel measures on $\mathbb{R}^n$ and $\hat{\mu} = \hat{\nu}$, then $\mu = \nu$.*

Proof. To prove that $\mu = \nu$ it suffices to prove that for any ball B in $\mathbb{R}^n$ we have $\mu(B) = \nu(B)$. Let $\phi_n \in C_c^\infty(\mathbb{R}^n) \to \chi_B$ pointwise. On the one hand, by the dominated convergence theorem, $\langle \phi_n, \mu \rangle \to \mu(B)$ and $\langle \phi_n, \nu \rangle \to \nu(B)$ as $n \to \infty$. On the other hand, because $\hat{\mu} = \hat{\nu}$ we have

$$\langle \phi_n, \mu \rangle = \int_{\mathbb{R}^n} \hat{\phi}_n(-\xi) \hat{\mu}(\xi) dm_n(\xi) = \int_{\mathbb{R}^n} \hat{\phi}_n(-\xi) \hat{\nu}(\xi) dm_n(\xi) = \langle \phi_n, \nu \rangle .$$

Therefore $\mu(B) = \nu(B)$, and it follows that $\mu = \nu$. □

2 Gaussian measures

Let $\lambda_1, \dots, \lambda_n > 0$, and let $\Lambda : \mathbb{R}^n \to \mathbb{R}^n$ be the linear map defined by $\Lambda e_i = \lambda_i e_i$. Define

$$d\mu(x) = \sqrt{\det \Lambda} \exp\left(-\frac{1}{2} x \cdot \Lambda x \right) dm_n(x),$$

called a *Gaussian measure*.

Theorem 2.

$$\hat{\mu}(\xi) = \exp\left(-\frac{1}{2} \xi \cdot \Lambda^{-1} \xi \right), \qquad \xi \in \mathbb{R}^n.$$

Proof. We have

$$\begin{aligned}
\hat{\mu}(\xi) &= \int_{\mathbb{R}^n} e^{-i\xi\cdot x} \sqrt{\det \Lambda} \exp\left(-\frac{1}{2} x \cdot \Lambda x \right) dm_n(x) \\
&= \int_{\mathbb{R}^d} e^{-i\xi_1 x_1 - \cdots - i\xi_n x_n} \sqrt{\lambda_1 \cdots \lambda_n} \exp\left(-\frac{1}{2} \lambda_1 x_1^2 - \cdots - \frac{1}{2} \lambda_n x_n^2 \right) dm_n(x) \\
&= \prod_{j=1}^n I_j,
\end{aligned}$$

where

$$I_j = \int_{\mathbb{R}} e^{-i\xi_j x_j} \sqrt{\lambda_j} \exp\left(-\frac{1}{2}\lambda_j x_j^2\right) dm_1(x_j).$$

Using

$$-i\xi_j x_j - \frac{1}{2}\lambda_j x_j^2 = -\frac{\lambda_j}{2}\left(\left(x_j + \frac{i\xi_j}{\lambda_j}\right)^2 + \frac{\xi_j^2}{\lambda_j^2}\right) = -\frac{\lambda_j}{2}\left(x_j + \frac{i\xi_j}{\lambda_j}\right)^2 - \frac{\xi_j^2}{2\lambda_j},$$

we get, doing contour integration,

$$\begin{aligned}
I_j &= \int_{\mathbb{R}} \sqrt{\lambda_j} \exp\left(-\frac{\lambda_j}{2}\left(x_j + \frac{i\xi_j}{\lambda_j}\right)^2\right) \exp\left(-\frac{\xi_j^2}{2\lambda_j}\right) dm_1(x_j) \\
&= \int_{\mathbb{R}} \sqrt{\lambda_j} \exp\left(-\frac{\lambda_j x_j^2}{2}\right) \exp\left(-\frac{\xi_j^2}{2\lambda_j}\right) dm_1(x_j) \\
&= \int_{\mathbb{R}} \sqrt{\lambda_j} \exp(-y_j^2) \exp\left(-\frac{\xi_j^2}{2\lambda_j}\right) \sqrt{\frac{2}{\lambda_j}} dm_1(y_j) \\
&= \exp\left(-\frac{\xi_j^2}{2\lambda_j}\right) \int_{\mathbb{R}} \sqrt{2} \exp(-y_j^2) dm_1(y_j) \\
&= \exp\left(-\frac{\xi_j^2}{2\lambda_j}\right) \int_{\mathbb{R}} \frac{1}{\sqrt{\pi}} \exp(-y_j^2) dy_j \\
&= \exp\left(-\frac{\xi_j^2}{2\lambda_j}\right).
\end{aligned}$$

Therefore, as $\Lambda^{-1}\xi = \sum_{j=1}^n \frac{\xi_j}{\lambda_j} e_j$ and $\xi \cdot \Lambda^{-1}\xi = \sum_{j=1}^n \frac{\xi_j^2}{\lambda_j}$,

$$\begin{aligned}
\hat{\mu}(\xi) &= \prod_{j=1}^n \exp\left(-\frac{\xi_j^2}{2\lambda_j}\right) \\
&= \exp\left(-\frac{1}{2}\sum_{j=1}^n \frac{\xi_j^2}{\lambda_j}\right) \\
&= \exp\left(-\frac{1}{2}\xi \cdot \Lambda^{-1}\xi\right).
\end{aligned}$$

□

From the above theorem we get

$$\hat{\mu}(0) = 1,$$

and hence a Gaussian measure is a probability measure.

For $h \in \mathbb{R}^n$, define $T_h : \mathbb{R}^n \to \mathbb{R}^n$ by $T_h(x) = x - h$. If E is a Borel subset of $\mathbb{R}^n$, because $\chi_{T_{-h}(E)} = \chi_E \circ T_h$,

$$((T_h)_*\mu)(E) = \mu(T_h^{-1}(E)) = \mu(T_{-h}(E)) = \int_{\mathbb{R}^n} \chi_{T_{-h}(E)} d\mu = \int_{\mathbb{R}^n} \chi_E \circ T_h d\mu.$$

Then, because $T_h \circ T_{-h} = \mathrm{id}_{\mathbb{R}^n}$,

$$\begin{aligned}
\int_{\mathbb{R}^n} \chi_E \circ T_h d\mu &= \int_{\mathbb{R}^n} \chi_E \circ T_h(x) \sqrt{\det \Lambda} \exp\left(-\frac{1}{2} x \cdot \Lambda x\right) dm_n(x) \\
&= \int_{\mathbb{R}^n} \chi_E(x) \sqrt{\det \Lambda} \exp\left(-\frac{1}{2}(T_{-h}x) \cdot (\Lambda T_{-h}x)\right) d((T_{-h})_* m_n)(x) \\
&= \int_{\mathbb{R}^n} \chi_E(x) \sqrt{\det \Lambda} \exp\left(-\frac{1}{2}(T_{-h}x) \cdot (\Lambda T_{-h}x)\right) dm_n(x).
\end{aligned}$$

As Λ is self-adjoint $\Lambda x \cdot h = x \cdot \Lambda h$,

$$\begin{aligned}
(T_{-h}x) \cdot (\Lambda T_{-h}x) &= (x+h) \cdot (\Lambda(x+h)) \\
&= (x+h) \cdot (\Lambda x + \Lambda h) \\
&= x \cdot \Lambda x + x \cdot \Lambda h + h \cdot \Lambda x + h \cdot \Lambda h \\
&= x \cdot \Lambda x + 2x \cdot \Lambda h + h \cdot \Lambda h.
\end{aligned}$$

Therefore,

$$\begin{aligned}
((T_h)_*\mu)(E) &= \int_{\mathbb{R}^n} \chi_E(x) \exp\left(-\frac{1}{2}\left(2x \cdot \Lambda h + h \cdot \Lambda h\right)\right) d\mu(x) \\
&= \int_{\mathbb{R}^n} \chi_E(x) \exp\left(-x \cdot \Lambda h - \frac{1}{2} h \cdot \Lambda h\right) d\mu(x).
\end{aligned}$$

This shows that the Radon-Nikodym derivative of $(T_h)_*\mu$ with respect to μ is

$$\frac{d(T_h)_*\mu}{d\mu}(x) = \exp\left(-x \cdot \Lambda h - \frac{1}{2} h \cdot \Lambda h\right).$$

3 Positive-definite functions

We say that a function $\phi : \mathbb{R}^n \to \mathbb{C}$ is *positive-definite* if $x_1, \ldots, x_r \in \mathbb{R}^n$ and $c_1, \ldots, c_r \in \mathbb{C}$ imply that

$$\sum_{i,j=1}^{r} c_i \overline{c_j} \phi(x_i - x_j) \geq 0;$$

in particular, the left-hand side is real.

Using $r = 1$, $c_1 = 1$, we have for any $x_1 \in \mathbb{R}^n$ that $\phi(x_1 - x_1) \geq 0$, i.e. $\phi(0) \geq 0$. For $x \in \mathbb{R}^n$, using $r = 2$, $x_1 = x, x_2 = 0$ and choosing fitting $c_1, c_2 \in \mathbb{C}$ gives

$$\phi(-x) = \overline{\phi(x)},$$

and using this with $c_2 = 1$ and for appropriate c_1 gives

$$|\phi(x)| \leq \phi(0).$$

For $f, g \in L^1(\mathbb{R}^n)$, the *convolution of f and g* is the function $f * g : \mathbb{R}^n \to \mathbb{C}$ defined by

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x-y)dm_n(y), \qquad x \in \mathbb{R}^n,$$

and $\|f * g\|_{L^1} \leq \|f\|_{L^1} \|g\|_{L^1}$, a case of *Young's inequality*. For $f : \mathbb{R}^n \to \mathbb{C}$, we denote by $\operatorname{supp} f$ the *essential support* of f; if f is continuous, then $\operatorname{supp} f$ is the closure of the set $\{x \in \mathbb{R}^n : f(x) \neq 0\}$. A fact that we will use later is[1]

$$\operatorname{supp}(f * g) \subseteq \overline{\operatorname{supp} f + \operatorname{supp} g}.$$

We denote by f^* the function defined by $f^*(x) = \overline{f(-x)}$.

$C_c(\mathbb{R}^n)$ is the set of all $f \in C(\mathbb{R}^n)$ for which $\operatorname{supp} f$ is a compact set. The set $C_c(\mathbb{R}^n)$ is dense in the Banach space $C_0(\mathbb{R}^n)$ and also in the Banach space $L^1(\mathbb{R}^n)$; $C_c(\mathbb{R}^n)$ is not a Banach space or even a Fréchet space, and thus does not have a robust structure itself, but is used because it is easier to prove things for it which one then extends in some way to spaces in which the set is dense. The proof of the following theorem follows Folland.[2]

Theorem 3. *If $\phi : \mathbb{R}^n \to \mathbb{C}$ is positive-definite and continuous and $f \in C_c(\mathbb{R}^n)$, then*

$$\int (f^* * f)\phi \geq 0.$$

Proof. Write $K = \operatorname{supp} f$, and define $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ by

$$F(x, y) = f(x)\overline{f(y)}\phi(x - y).$$

F is continuous, and $\operatorname{supp} F \subseteq K \times K$, hence $\operatorname{supp} F$ is compact. Thus $F \in C_c(\mathbb{R}^n \times \mathbb{R}^n)$; in particular F is uniformly continuous on $K \times K$, and it follows that for each $\epsilon > 0$ there is some $\delta > 0$ such that if $|x-a| < \delta$ and $|y-b| < \delta$ then $|F(x,y) - F(a,b)| < \epsilon$. The collection $\{B_\delta(x) : x \in K\}$ covers K and hence there are finitely many distinct $x_i \in K$ such that the collection $\{B_\delta(x_i) : i\}$ covers K. Then $\{B_\delta(x_i) \times B_\delta(x_j) : i, j\}$ covers $K \times K$. Let E_i be pairwise disjoint, measurable, and satisfy $x_i \in E_i \subseteq B_\delta(x_i)$. The collection $\{E_i : i,\}$ covers K, so the collection $\{E_i \times E_j : i, j\}$ covers $K \times K$.

Define

$$R = \sum_{i,j} \int_{E_i \times E_j} (F(x, y) - F(x_i, x_j))dm_n(x)dm_n(y).$$

[1] Gerald B. Folland, *Real Analysis: Modern Techniques and their Applications*, second ed., p. 240, Proposition 8.6.

[2] Gerald B. Folland, *A Course in Abstract Harmonic Analysis*, p. 85, Proposition 3.35.

R satisfies

$$\begin{aligned}
|R| &\le \sum_{i,j} \int_{E_i \times E_j} |F(x,y) - F(x_i,x_j)| dm_n(x) dm_n(y) \\
&\le \sum_{i,j} \int_{E_i \times E_j} \epsilon dm_n(x) dm_n(y) \\
&= \epsilon \sum_{i,j} m_n(E_i) m_n(E_j) \\
&= \epsilon m_n(K)^2.
\end{aligned}$$

We obtain

$$\begin{aligned}
\int_{K \times K} F(x,y) dm_n(x) dm_n(y) &= \sum_{i,j} \int_{E_i \times E_j} F(x,y) dm_n(x) dm_n(y) \\
&= \sum_{i,j} F(x_i,x_j) m_n(E_i) m_n(E_j) + R \\
&= \sum_{i,j} f(x_i) \overline{f(x_j)} \phi(x_i - x_j) m_n(E_i) m_n(E_j) + R.
\end{aligned}$$

Using $c_i = f(x_i) m_n(E_i)$, the fact that ϕ is positive-definite means that the sum is ≥ 0. Therefore

$$\int_{K \times K} F(x,y) dm_n(x) dm_n(y) \ge -|R| \ge -\epsilon m_n(K)^2.$$

This is true for all $\epsilon > 0$, hence

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x) \overline{f(y)} \phi(x-y) dm_n(x) dm_n(y) = \int_{K \times K} F(x,y) dm_n(x) dm_n(y) \ge 0.$$

But

$$\begin{aligned}
\int_{\mathbb{R}^n} (f^* * f)(x) \phi(x) dm_n(x) &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} f^*(y) f(x-y) dm_n(y) \right) \phi(x) dm_n(x) \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \overline{f(-y)} f(x-y) \phi(x) dm_n(x) dm_n(y) \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \overline{f(-y)} f(x) \phi(x+y) dm_n(x) dm_n(y) \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \overline{f(y)} f(x) \phi(x-y) dm_n(x) dm_n(y).
\end{aligned}$$

□

Corollary 4. *If $\phi : \mathbb{R}^n \to \mathbb{C}$ is positive-definite and continuous and $f \in L^1(\mathbb{R}^n)$, then*

$$\int (f^* * f) \phi \ge 0.$$

Proof. Let $f_n \in C_c(\mathbb{R}^n)$ converge to f in $L^1(\mathbb{R}^n)$ as $n \to \infty$; that there is such a sequence is given to us by the fact that $C_c(\mathbb{R}^n)$ is a dense subset of $L^1(\mathbb{R}^n)$. Using

$$\begin{aligned} f_n^* * f_n - f^* * f &= f_n^* * f_n - f_n^* * f + f_n^* * f - f^* * f \\ &= f_n^* * (f_n - f) + (f_n^* - f^*) * f \\ &= f_n^* * (f_n - f) + (f_n - f)^* * f, \end{aligned}$$

and $\|g^*\|_{L^1} = \|g\|_{L^1}$, we get

$$\begin{aligned} \|f_n^* * f_n - f^* * f\|_{L^1} &\leq \|f_n^* * (f_n - f)\|_{L^1} + \|(f_n - f)^* * f\|_{L^1} \\ &\leq \|f_n^*\|_{L^1} \|f_n - f\|_{L^1} + \|(f_n - f)^*\|_{L^1} \|f\|_{L^1} \\ &= \|f_n\|_{L^1} \|f_n - f\|_{L^1} + \|f_n - f\|_{L^1} \|f\|_{L^1}, \end{aligned}$$

which converges to 0 because $\|f_n - f\|_{L^1} \to 0$. Therefore, because ϕ is bounded,

$$\int_{\mathbb{R}^n} (f_n^* * f_n)\phi dm_n \to \int_{\mathbb{R}^n} (f^* * f)\phi dm_n.$$

As $\int_{\mathbb{R}^n} (f_n^* * f_n)\phi dm_n \geq 0$ for each n, this implies that $\int_{\mathbb{R}^n} (f^* * f)\phi dm_n \geq 0$. □

It is straightforward to prove that the Fourier transform of a finite positive Borel measure is a positive-definite function; one ends up with the expression

$$\int_{\mathbb{R}^n} \left| \sum_{j=1}^n c_j e^{i\xi_j \cdot x} \right|^2 d\mu(x),$$

which is finite and nonnegative because μ is finite and positive respectively. We have established already that the Fourier transform of a finite positive Borel measure μ on $\mathbb{R}^n$ is continuous and satisfies $\hat{\mu}(0) = 1$. *Bochner's theorem* is the statement that a function with these three properties is indeed the Fourier transform of a finite positive Borel measure. Our proof of the following theorem follows Folland.[3]

Theorem 5 (Bochner)**.** *If $\phi : \mathbb{R}^n \to \mathbb{C}$ is positive-definite, continuous, and satisfies $\phi(0) = 1$, then there is some Borel probability measure μ on $\mathbb{R}^n$ such that $\phi = \hat{\mu}$.*

Proof. Let $\{\psi_U\}$ be an approximate identity. That is, for each neighborhood U of 0, ψ_U is a function such that $\operatorname{supp}\psi_U$ is compact and contained in U, $\psi \geq 0$, $\psi_U(-x) = \psi_U(x)$, and $\int_{\mathbb{R}^n} \psi_U dm_n = 1$. For every $f \in L^1(\mathbb{R}^n)$, an approximate identity satisfies $\|f * \psi_U - f\|_{L^1} \to 0$ as $U \to \{0\}$.[4]

We have $\psi_U^* = \psi_{-U}$, so

$$\operatorname{supp}(\psi_U^* * \psi_U) \subseteq \overline{\operatorname{supp}\psi_{-U} + \operatorname{supp}\psi_U} = \operatorname{supp}\psi_{-U} + \operatorname{supp}\psi_U \subseteq -U + U,$$

[3]Gerald B. Folland, *A Course in Abstract Harmonic Analysis*, p. 95, Theorem 4.18.
[4]Gerald B. Folland, *A Course in Abstract Harmonic Analysis*, p. 53, Proposition 2.42.

and as always, $\int_{\mathbb{R}^n} f * g dm_n = \int_{\mathbb{R}^n} f dm_n \int_{\mathbb{R}^n} g dm_n$. Therefore $\{\psi_U^* * \psi_U\}$ is an approximate identity:

For $f, g \in L^1(\mathbb{R}^n)$, define

$$\langle f, g \rangle_\phi = \int_{\mathbb{R}^n} (g^* * f)\phi dm_n.$$

One checks that this is a positive Hermitian form; *positive* means that $\langle f, f \rangle_\phi \geq 0$ for all $f \in L^1(\mathbb{R}^n)$, and this is given to us by Corollary 4. Using the Cauchy-Schwarz inequality,[5]

$$|\langle f, g \rangle_\phi|^2 \leq \langle f, f \rangle_\phi \langle g, g \rangle_\phi .$$

We have laid out the tools that we will use. Let $f \in L^1(\mathbb{R}^n)$. $\psi_U * f \to f$ in L^1 as $U \to \{0\}$, and as ϕ is bounded this gives $\int_{\mathbb{R}^n} (\psi_U^* * f)\phi dm_n \to \int_{\mathbb{R}^n} f\phi dm_n$ as $U \to \{0\}$. Because $\{\psi_U^* * \psi_U\}$ is an approximate identity, $\int_{\mathbb{R}^n} (\psi_U^* * \psi_U)\phi dm_n \to \phi(0)$ as $U \to \{0\}$. That is, we have $\langle f, \psi_U \rangle_\phi \to \int_{\mathbb{R}^n} f\phi dm_n$ and $\langle \psi_U, \psi_U \rangle_\phi \to \phi(0$ as $U \to \{0\}$, and as $\phi(0) = 1$, the above statemtn of the Cauchy-Schwarz inequality produces

$$\left| \int_{\mathbb{R}^n} f\phi dm_n \right|^2 \leq \int_{\mathbb{R}^n} (f^* * f)\phi dm_n. \tag{1}$$

With $h = f^* * f$, the inequality (1) reads

$$\left| \int_{\mathbb{R}^n} f\phi dm_n \right|^2 \leq \int_{\mathbb{R}^n} h\phi dm_n.$$

Defining $h^{(1)} = h$, $h^{(2)} = h * h$, $h^{(3)} = h * h * h$, etc., applying (1) to h gives, because $h^* = h$,

$$\left| \int_{\mathbb{R}^n} h\phi dm_n \right|^2 \leq \int_{\mathbb{R}^n} h^{(2)}\phi dm_n.$$

Then applying (1) to $h^{(2)}$, which satisfies $(h^{(2)})^* = h^{(2)}$,

$$\left| \int_{\mathbb{R}^n} h^{(2)}\phi dm_n \right|^2 \leq \int_{\mathbb{R}^n} h^{(4)}\phi dm_n.$$

Thus, for any $m \geq 0$ we have

$$\begin{aligned} \left| \int_{\mathbb{R}^n} f\phi dm_n \right| &\leq \left| \int_{\mathbb{R}^n} h^{(2^m)}\phi dm_n \right|^{2^{-(m+1)}} \\ &\leq \left\| h^{(2^m)} \right\|_{L^1}^{2^{-(m+1)}} \\ &= \left(\left\| h^{(2^m)} \right\|_{L^1}^{2^{-m}} \right)^{1/2}, \end{aligned}$$

[5] Jean Dieudonne, *Foundations of Modern Analysis*, 1969, p. 117, Theorem 6.2.1.

since $\|\phi\|_\infty = \phi(0) = 1$.

With convolution as multiplication, $L^1(\mathbb{R}^n)$ is a commutative Banach algebra, and the Gelfand transform is an algebra homomorphism $L^1(\mathbb{R}^n) \to C_0(\mathbb{R}^n)$ that satisfies[6]

$$\|\hat{g}\|_\infty = \lim_{k\to\infty} \left\|g^{(k)}\right\|_{L^1}^{1/k}, \qquad g \in L^1(\mathbb{R}^n);$$

for $L^1(\mathbb{R}^n)$, the Gelfand transform is the Fourier transform. Write the Fourier transform as $\mathscr{F} : L^1(\mathbb{R}^n) \to C_0(\mathbb{R}^n)$. Stating that the Gelfand transform is a homomorphism means that $\mathscr{F}(g_1 * g_2) = \mathscr{F}(g_1)\mathscr{F}(g_2)$, because multiplication in the Banach algebra $C_0(\mathbb{R}^n)$ is pointwise multiplication. Then, since a subsequence of a convergent sequence converges to the same limit,

$$\lim_{m\to\infty} \left(\left\|h^{(2^m)}\right\|_{L^1}^{2^{-m}} \right)^{1/2} = \left(\left\|\hat{h}\right\|_\infty \right)^{1/2}.$$

But

$$\hat{h} = \mathscr{F}(f^* * f) = \mathscr{F}(f^*)\mathscr{F}(f) = \overline{\mathscr{F}(f)}\mathscr{F}(f) = |\mathscr{F}(f)|^2,$$

so

$$\left(\left\|\hat{h}\right\|_\infty \right)^{1/2} = \left(\left\||\hat{f}|^2\right\|_\infty \right)^{1/2} = \left\|\hat{f}\right\|_\infty.$$

Putting things together, we have that for any $f \in L^1(\mathbb{R}^n)$,

$$\left| \int_{\mathbb{R}^n} f\phi dm_n \right| \leq \left\|\hat{f}\right\|_\infty.$$

Therefore $\hat{f} \mapsto \int_{\mathbb{R}^n} f\phi dm_n$ is a bounded linear functional $\mathscr{F}(L^1(\mathbb{R}^n)) \to \mathbb{C}$, of norm ≤ 1. Using $\phi(0) = 1$, one proves that this functional has norm 1. (If we could apply this inequality to $\mathscr{F}(\delta)$ the two sides would be equal, thus to prove that the operator norm is 1, one applies the inequality to a sequence of functions that converge weakly to δ.) We take as known that $\mathscr{F}(L^1(\mathbb{R}^n))$ is dense in the Banach space $C_0(\mathbb{R}^n)$, so there is a bounded linear functional $\Phi : C_0(\mathbb{R}^n) \to \mathbb{C}$ whose restriction to $\mathscr{F}(L^1(\mathbb{R}^n))$ is equal to $\hat{f} \mapsto \int_{\mathbb{R}^n} f\phi dm_n$, and $\|\Phi\| = 1$.

Using the Riesz-Markov theorem,[7] there is a regular complex Borel measure μ on $\mathbb{R}^n$ such that

$$\Phi(g) = \int_{\mathbb{R}^n} g d\mu, \qquad g \in C_0(\mathbb{R}^n),$$

and $\|\mu\| = \|\Phi\|$; $\|\mu\|$ is the *total variation norm* of μ, $\|\mu\| = |\mu|(\mathbb{R}^n)$. Then for

[6] Gerald B. Folland, *A Course in Abstract Harmonic Analysis*, p. 15, Theorem 1.30. Namely, this is the *Gelfand-Naimark theorem.*

[7] Walter Rudin, *Real and Complex Analysis*, third ed., p. 130, Theorem 6.19.

$f \in L^1(\mathbb{R}^n)$ we have

$$\begin{aligned}
\int_{\mathbb{R}^n} f\phi dm_n &= \Phi(\hat{f}) \\
&= \int_{\mathbb{R}^n} \hat{f} d\mu \\
&= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{-i\xi\cdot x} f(x) dm_n(x) \right) d\mu(\xi) \\
&= \int_{\mathbb{R}^n} f(x) \left(\int_{\mathbb{R}^n} e^{-ix\cdot\xi} d\mu(\xi) \right) dm_n(x) \\
&= \int_{\mathbb{R}^n} f(x)\hat{\mu}(x) dm_n(x).
\end{aligned}$$

That this is true for all $f \in L^1(\mathbb{R}^n)$ implies that $\phi = \hat{\mu}$. As $\mu(\mathbb{R}^n) = \hat{\mu}(0) = \phi(0) = 1$ and $\|\mu\| = \|\Phi\| = 1$ we have $\mu(\mathbb{R}^n) = \|\mu\|$, and this implies that μ is positive measure, hence, as $\mu(\mathbb{R}^n) = 1$, a probability measure.

□