

11. APPROXIMATION THEOREMS AND CONVOLUTIONS

Let (X, \mathcal{M}, μ) be a measure space, $\mathcal{A} \subset \mathcal{M}$ an algebra.

Notation 11.1. Let $\mathbb{S}_f(\mathcal{A}, \mu)$ denote those simple functions $\phi : X \rightarrow \mathbb{C}$ such that $\phi^{-1}(\{\lambda\}) \in \mathcal{A}$ for all $\lambda \in \mathbb{C}$ and $\mu(\phi \neq 0) < \infty$.

For $\phi \in \mathbb{S}_f(\mathcal{A}, \mu)$ and $p \in [1, \infty)$, $|\phi|^p = \sum_{z \neq 0} |z|^p 1_{\{\phi=z\}}$ and hence

$$\int |\phi|^p d\mu = \sum_{z \neq 0} |z|^p \mu(\phi = z) < \infty$$

so that $\mathbb{S}_f(\mathcal{A}, \mu) \subset L^p(\mu)$.

Lemma 11.2 (Simple Functions are Dense). *The simple functions, $\mathbb{S}_f(\mathcal{M}, \mu)$, form a dense subspace of $L^p(\mu)$ for all $1 \leq p < \infty$.*

Proof. Let $\{\phi_n\}_{n=1}^\infty$ be the simple functions in the approximation Theorem 7.12. Since $|\phi_n| \leq |f|$ for all n , $\phi_n \in \mathbb{S}_f(\mathcal{M}, \mu)$ (verify!) and

$$|f - \phi_n|^p \leq (|f| + |\phi_n|)^p \leq 2^p |f|^p \in L^1.$$

Therefore, by the dominated convergence theorem,

$$\lim_{n \rightarrow \infty} \int |f - \phi_n|^p d\mu = \int \lim_{n \rightarrow \infty} |f - \phi_n|^p d\mu = 0.$$

■

Theorem 11.3 (Separable Algebras implies Separability of L^p - Spaces). *Suppose $1 \leq p < \infty$ and $\mathcal{A} \subset \mathcal{M}$ is an algebra such that $\sigma(\mathcal{A}) = \mathcal{M}$ and μ is σ -finite on \mathcal{A} . Then $\mathbb{S}_f(\mathcal{A}, \mu)$ is dense in $L^p(\mu)$. Moreover, if \mathcal{A} is countable, then $L^p(\mu)$ is separable and*

$$\mathbb{D} = \left\{ \sum a_j 1_{A_j} : a_j \in \mathbb{Q} + i\mathbb{Q}, A_j \in \mathcal{A} \text{ with } \mu(A_j) < \infty \right\}$$

is a countable dense subset.

Proof. First Proof. Let $X_k \in \mathcal{A}$ be sets such that $\mu(X_k) < \infty$ and $X_k \uparrow X$ as $k \rightarrow \infty$. For $k \in \mathbb{N}$ let \mathcal{H}_k denote those bounded \mathcal{M} - measurable functions, f , on X such that $1_{X_k} f \in \overline{\mathbb{S}_f(\mathcal{A}, \mu)}^{L^p(\mu)}$. It is easily seen that \mathcal{H}_k is a vector space closed under bounded convergence and this subspace contains 1_A for all $A \in \mathcal{A}$. Therefore by Theorem 8.12, \mathcal{H}_k is the set of all bounded \mathcal{M} - measurable functions on X .

For $f \in L^p(\mu)$, the dominated convergence theorem implies $1_{X_k \cap \{|f| \leq k\}} f \rightarrow f$ in $L^p(\mu)$ as $k \rightarrow \infty$. We have just proved $1_{X_k \cap \{|f| \leq k\}} f \in \overline{\mathbb{S}_f(\mathcal{A}, \mu)}^{L^p(\mu)}$ for all k and hence it follows that $f \in \overline{\mathbb{S}_f(\mathcal{A}, \mu)}^{L^p(\mu)}$. The last assertion of the theorem is a consequence of the easily verified fact that \mathbb{D} is dense in $\mathbb{S}_f(\mathcal{A}, \mu)$ relative to the $L^p(\mu)$ - norm.

Second Proof. Given $\epsilon > 0$, by Corollary 8.42, for all $E \in \mathcal{M}$ such that $\mu(E) < \infty$, there exists $A \in \mathcal{A}$ such that $\mu(E \Delta A) < \epsilon$. Therefore

$$(11.1) \quad \int |1_E - 1_A|^p d\mu = \mu(E \Delta A) < \epsilon.$$

This equation shows that any simple function in $\mathbb{S}_f(\mathcal{M}, \mu)$ may be approximated arbitrary well by an element from \mathbb{D} and hence \mathbb{D} is also dense in $L^p(\mu)$. ■

Corollary 11.4 (Riemann Lebesgue Lemma). *Suppose that $f \in L^1(\mathbb{R}, m)$, then*

$$\lim_{\lambda \rightarrow \pm\infty} \int_{\mathbb{R}} f(x) e^{i\lambda x} dm(x) = 0.$$

Proof. Let \mathcal{A} denote the algebra on \mathbb{R} generated by the half open intervals, i.e. \mathcal{A} consists of sets of the form

$$\prod_{k=1}^n (a_k, b_k] \cap \mathbb{R}$$

where $a_k, b_k \in \bar{\mathbb{R}}$. By Theorem 11.3 given $\epsilon > 0$ there exists $\phi = \sum_{k=1}^n c_k 1_{(a_k, b_k]}$ with $a_k, b_k \in \mathbb{R}$ such that

$$\int_{\mathbb{R}} |f - \phi| dm < \epsilon.$$

Notice that

$$\begin{aligned} \int_{\mathbb{R}} \phi(x) e^{i\lambda x} dm(x) &= \int_{\mathbb{R}} \sum_{k=1}^n c_k 1_{(a_k, b_k]}(x) e^{i\lambda x} dm(x) \\ &= \sum_{k=1}^n c_k \int_{a_k}^{b_k} e^{i\lambda x} dm(x) = \sum_{k=1}^n c_k \lambda^{-1} e^{i\lambda x} \Big|_{a_k}^{b_k} \\ &= \lambda^{-1} \sum_{k=1}^n c_k (e^{i\lambda b_k} - e^{i\lambda a_k}) \rightarrow 0 \text{ as } |\lambda| \rightarrow \infty. \end{aligned}$$

Combining these two equations with

$$\begin{aligned} \left| \int_{\mathbb{R}} f(x) e^{i\lambda x} dm(x) \right| &\leq \left| \int_{\mathbb{R}} (f(x) - \phi(x)) e^{i\lambda x} dm(x) \right| + \left| \int_{\mathbb{R}} \phi(x) e^{i\lambda x} dm(x) \right| \\ &\leq \int_{\mathbb{R}} |f - \phi| dm + \left| \int_{\mathbb{R}} \phi(x) e^{i\lambda x} dm(x) \right| \\ &\leq \epsilon + \left| \int_{\mathbb{R}} \phi(x) e^{i\lambda x} dm(x) \right| \end{aligned}$$

we learn that

$$\limsup_{|\lambda| \rightarrow \infty} \left| \int_{\mathbb{R}} f(x) e^{i\lambda x} dm(x) \right| \leq \epsilon + \limsup_{|\lambda| \rightarrow \infty} \left| \int_{\mathbb{R}} \phi(x) e^{i\lambda x} dm(x) \right| = \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have proven the lemma. ■

Theorem 11.5 (Continuous Functions are Dense). *Let (X, d) be a metric space, τ_d be the topology on X generated by d and $\mathcal{B}_X = \sigma(\tau_d)$ be the Borel σ -algebra. Suppose $\mu : \mathcal{B}_X \rightarrow [0, \infty]$ is a measure which is σ -finite on τ_d and let $BC_f(X)$ denote the bounded continuous functions on X such that $\mu(f \neq 0) < \infty$. Then $BC_f(X)$ is a dense subspace of $L^p(\mu)$ for any $p \in [1, \infty)$.*

Proof. First Proof. Let $X_k \in \tau_d$ be open sets such that $X_k \uparrow X$ and $\mu(X_k) < \infty$. Let k and n be positive integers and set

$$\psi_{n,k}(x) = \min(1, n \cdot d_{X_k^c}(x)) = \phi_n(d_{X_k^c}(x)),$$

and notice that $\psi_{n,k} \rightarrow 1_{d_{X_k^c} > 0} = 1_{X_k}$ as $n \rightarrow \infty$, see Figure 25 below.

Then $\psi_{n,k} \in BC_f(X)$ and $\{\psi_{n,k} \neq 0\} \subset X_k$. Let \mathcal{H} denote those bounded \mathcal{M} -measurable functions, $f : X \rightarrow \mathbb{R}$, such that $\psi_{n,k} f \in \overline{BC_f(X)}^{L^p(\mu)}$. It is easily seen that \mathcal{H} is a vector space closed under bounded convergence and this

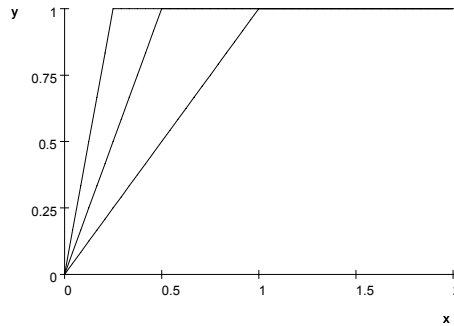


FIGURE 25. The plot of ϕ_n for $n = 1, 2,$ and 4 . Notice that $\phi_n \rightarrow 1_{(0,\infty)}$.

subspace contains $BC(X, \mathbb{R})$. By Corollary 8.13, \mathcal{H} is the set of all bounded real valued \mathcal{M} -measurable functions on X , i.e. $\psi_{n,k}f \in \overline{BC_f(X)}^{L^p(\mu)}$ for all bounded measurable f and $n, k \in \mathbb{N}$. Let f be a bounded measurable function, by the dominated convergence theorem, $\psi_{n,k}f \rightarrow 1_{X_k}f$ in $L^p(\mu)$ as $n \rightarrow \infty$, therefore $1_{X_k}f \in \overline{BC_f(X)}^{L^p(\mu)}$. It now follows as in the first proof of Theorem 11.3 that $\overline{BC_f(X)}^{L^p(\mu)} = L^p(\mu)$.

Second Proof. Since $\mathbb{S}_f(\mathcal{M}, \mu)$ is dense in $L^p(\mu)$ it suffices to show any $\phi \in \mathbb{S}_f(\mathcal{M}, \mu)$ may be well approximated by $f \in BC_f(X)$. Moreover, to prove this it suffices to show for $A \in \mathcal{M}$ with $\mu(A) < \infty$ that 1_A may be well approximated by an $f \in BC_f(X)$. By Exercises 8.4 and 8.5, for any $\epsilon > 0$ there exists a closed set F and an open set V such that $F \subset A \subset V$ and $\mu(V \setminus F) < \epsilon$. (Notice that $\mu(V) < \mu(A) + \epsilon < \infty$.) Let f be as in Eq. (10.1), then $f \in BC_f(X)$ and since $|1_A - f| \leq 1_{V \setminus F}$,

$$(11.2) \quad \int |1_A - f|^p d\mu \leq \int 1_{V \setminus F} d\mu = \mu(V \setminus F) \leq \epsilon$$

or equivalently

$$\|1_A - f\| \leq \epsilon^{1/p}.$$

Since $\epsilon > 0$ is arbitrary, we have shown that 1_A can be approximated in $L^p(\mu)$ arbitrarily well by functions from $BC_f(X)$. ■

Proposition 11.6. *Let (X, τ) be a second countable locally compact Hausdorff space, $\mathcal{B}_X = \sigma(\tau)$ be the Borel σ -algebra and $\mu : \mathcal{B}_X \rightarrow [0, \infty]$ be a measure such that $\mu(K) < \infty$ when K is a compact subset of X . Then $C_c(X)$ (the space of continuous functions with compact support) is dense in $L^p(\mu)$ for all $p \in [1, \infty)$.*

Proof. First Proof. Let $\{K_k\}_{k=1}^\infty$ be a sequence of compact sets as in Lemma 10.10 and set $X_k = K_k^o$. Using Item 3. of Lemma 10.17, there exists $\{\psi_{n,k}\}_{n=1}^\infty \subset C_c(X)$ such that $\text{supp}(\psi_{n,k}) \subset X_k$ and $\lim_{n \rightarrow \infty} \psi_{n,k} = 1_{X_k}$. As in the first proof of Theorem 11.5, let \mathcal{H} denote those bounded \mathcal{B}_X -measurable functions, $f : X \rightarrow \mathbb{R}$, such that $\psi_{n,k}f \in \overline{C_c(X)}^{L^p(\mu)}$. It is easily seen that \mathcal{H} is a vector space closed under bounded convergence and this subspace contains $BC(X, \mathbb{R})$. By Corollary 10.18, \mathcal{H} is the set of all bounded real valued \mathcal{B}_X -measurable functions on X , i.e.

$\psi_{n,k}f \in \overline{C_c(X)}^{L^p(\mu)}$ for all bounded measurable f and $n, k \in \mathbb{N}$. Let f be a bounded measurable function, by the dominated convergence theorem, $\psi_{n,k}f \rightarrow 1_{X_k}f$ in $L^p(\mu)$ as $k \rightarrow \infty$, therefore $1_{X_k}f \in \overline{C_c(X)}^{L^p(\mu)}$. It now follows as in the first proof of Theorem 11.3 that $\overline{C_c(X)}^{L^p(\mu)} = L^p(\mu)$.

Second Proof. Following the second proof of Theorem 11.5, let $A \in \mathcal{M}$ with $\mu(A) < \infty$. Since $\lim_{k \rightarrow \infty} \|1_{A \cap K_k^o} - 1_A\|_p = 0$, it suffices to assume $A \subset K_k^o$ for some k . Given $\epsilon > 0$, by Item 2. of Lemma 10.17 and Exercises 8.4 there exists a closed set F and an open set V such that $F \subset A \subset V$ and $\mu(V \setminus F) < \epsilon$. Replacing V by $V \cap K_k^o$ we may assume that $V \subset K_k^o \subset K_k$. The function f defined in Eq. (10.1) is now in $C_c(X)$. The remainder of the proof now follows as in the second proof of Theorem 11.5. ■

Lemma 11.7. *Let (X, τ) be a second countable locally compact Hausdorff space, $\mathcal{B}_X = \sigma(\tau)$ be the Borel σ -algebra and $\mu : \mathcal{B}_X \rightarrow [0, \infty]$ be a measure such that $\mu(K) < \infty$ when K is a compact subset of X . If $h \in L^1_{loc}(\mu)$ is a function such that*

$$(11.3) \quad \int_X fhd\mu = 0 \text{ for all } f \in C_c(X)$$

then $h(x) = 0$ for μ -a.e. x .

Proof. First Proof. Let $d\nu(x) = |h(x)|dx$, then ν is a measure on X such that $\nu(K) < \infty$ for all compact subsets $K \subset X$ and hence $C_c(X)$ is dense in $L^1(\nu)$ by Proposition 11.6. Notice that

$$(11.4) \quad \int_X f \cdot \text{sgn}(h)d\nu = \int_X fhd\mu = 0 \text{ for all } f \in C_c(X).$$

Let $\{K_k\}_{k=1}^\infty$ be a sequence of compact sets such that $K_k \uparrow X$ as in Lemma 10.10. Then $1_{K_k} \overline{\text{sgn}(h)} \in L^1(\nu)$ and therefore there exists $f_m \in C_c(X)$ such that $f_m \rightarrow 1_{K_k} \overline{\text{sgn}(h)}$ in $L^1(\nu)$. So by Eq. (11.4),

$$\nu(K_k) = \int_X 1_{K_k} d\nu = \lim_{m \rightarrow \infty} \int_X f_m \overline{\text{sgn}(h)} d\nu = 0.$$

Since $K_k \uparrow X$ as $k \rightarrow \infty$, $0 = \nu(X) = \int_X |h| d\mu$, i.e. $h(x) = 0$ for μ -a.e. x .

Second Proof. Let K_k be as above and use Lemma 10.15 to find $\chi \in C_c(X, [0, 1])$ such that $\chi = 1$ on K_k . Let \mathcal{H} denote the set of bounded measurable real valued functions on X such that $\int_X \chi fhd\mu = 0$. Then it is easily checked that \mathcal{H} is linear subspace closed under bounded convergence which contains $C_c(X)$. Therefore by Corollary 10.18, $0 = \int_X \chi fhd\mu$ for all bounded measurable functions $f : X \rightarrow \mathbb{R}$ and then by linearity for all bounded measurable functions $f : X \rightarrow \mathbb{C}$. Taking $f = \overline{\text{sgn}(h)}$ then implies

$$0 = \int_X \chi |h| d\mu \geq \int_{K_k} |h| d\mu$$

and hence by the monotone convergence theorem,

$$0 = \lim_{k \rightarrow \infty} \int_{K_k} |h| d\mu = \int_X |h| d\mu.$$

■

Corollary 11.8. *Suppose $X \subset \mathbb{R}^n$ is an open set, \mathcal{B}_X is the Borel σ -algebra on X and μ is a measure on (X, \mathcal{B}_X) which is finite on compact sets. Then $C_c(X)$ is dense in $L^p(\mu)$ for all $p \in [1, \infty)$.*

11.1. Convolution and Young's Inequalities.

Definition 11.9. Let $f, g : \mathbb{R}^n \rightarrow \mathbb{C}$ be measurable functions. We define

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy$$

whenever the integral is defined, i.e. either $f(x-\cdot)g(\cdot) \in L^1(\mathbb{R}^n, m)$ or $f(x-\cdot)g(\cdot) \geq 0$. Notice that the condition that $f(x-\cdot)g(\cdot) \in L^1(\mathbb{R}^n, m)$ is equivalent to writing $|f| * |g|(x) < \infty$.

Notation 11.10. Given a multi-index $\alpha \in \mathbb{Z}_+^n$, let $|\alpha| = \alpha_1 + \cdots + \alpha_n$,

$$x^\alpha := \prod_{j=1}^n x_j^{\alpha_j}, \text{ and } \partial_x^\alpha = \left(\frac{\partial}{\partial x} \right)^\alpha := \prod_{j=1}^n \left(\frac{\partial}{\partial x_j} \right)^{\alpha_j}.$$

Remark 11.11 (The Significance of Convolution). Suppose that $L = \sum_{|\alpha| \leq k} a_\alpha \partial^\alpha$ is a constant coefficient differential operator and suppose that we can solve (uniquely) the equation $Lu = g$ in the form

$$u(x) = Kg(x) := \int_{\mathbb{R}^n} k(x, y)g(y)dy$$

where $k(x, y)$ is an “integral kernel.” (This is a natural sort of assumption since, in view of the fundamental theorem of calculus, integration is the inverse operation to differentiation.) Since $\tau_z L = L\tau_z$ for all $z \in \mathbb{R}^n$, (this is another way to characterize constant coefficient differential operators) and $L^{-1} = K$ we should have $\tau_z K = K\tau_z$. Writing out this equation then says

$$\begin{aligned} \int_{\mathbb{R}^n} k(x-z, y)g(y)dy &= (Kg)(x-z) = \tau_z Kg(x) = (K\tau_z g)(x) \\ &= \int_{\mathbb{R}^n} k(x, y)g(y-z)dy = \int_{\mathbb{R}^n} k(x, y+z)g(y)dy. \end{aligned}$$

Since g is arbitrary we conclude that $k(x-z, y) = k(x, y+z)$. Taking $y = 0$ then gives

$$k(x, z) = k(x-z, 0) =: \rho(x-z).$$

We thus find that $Kg = \rho * g$. Hence we expect the convolution operation to appear naturally when solving constant coefficient partial differential equations. More about this point later.

The following proposition is an easy consequence of Minkowski's inequality for integrals, Theorem 9.27.

Proposition 11.12. *Suppose $q \in [1, \infty]$, $f \in L^1$ and $g \in L^q$, then $f * g(x)$ exists for almost every x , $f * g \in L^q$ and*

$$\|f * g\|_p \leq \|f\|_1 \|g\|_p.$$

For $z \in \mathbb{R}^n$ and $f : \mathbb{R}^n \rightarrow \mathbb{C}$, let $\tau_z f : \mathbb{R}^n \rightarrow \mathbb{C}$ be defined by $\tau_z f(x) = f(x-z)$.

Proposition 11.13. *Suppose that $p \in [1, \infty)$, then $\tau_z : L^p \rightarrow L^p$ is an isometric isomorphism and for $f \in L^p$, $z \in \mathbb{R}^n \rightarrow \tau_z f \in L^p$ is continuous.*

Proof. The assertion that $\tau_z : L^p \rightarrow L^p$ is an isometric isomorphism follows from translation invariance of Lebesgue measure and the fact that $\tau_{-z} \circ \tau_z = id$. For the continuity assertion, observe that

$$\|\tau_z f - \tau_y f\|_p = \|\tau_{-y}(\tau_z f - \tau_y f)\|_p = \|\tau_{z-y} f - f\|_p$$

from which it follows that it is enough to show $\tau_z f \rightarrow f$ in L^p as $z \rightarrow 0 \in \mathbb{R}^n$.

When $f \in C_c(\mathbb{R}^n)$, $\tau_z f \rightarrow f$ uniformly and since the $K := \cup_{|z| \leq 1} \text{supp}(\tau_z f)$ is compact, it follows by the dominated convergence theorem that $\tau_z f \rightarrow f$ in L^p as $z \rightarrow 0 \in \mathbb{R}^n$. For general $g \in L^p$ and $f \in C_c(\mathbb{R}^n)$,

$$\|\tau_z g - g\|_p \leq \|\tau_z g - \tau_z f\|_p + \|\tau_z f - f\|_p + \|f - g\|_p = \|\tau_z f - f\|_p + 2\|f - g\|_p$$

and thus

$$\limsup_{z \rightarrow 0} \|\tau_z g - g\|_p \leq \limsup_{z \rightarrow 0} \|\tau_z f - f\|_p + 2\|f - g\|_p = 2\|f - g\|_p.$$

Because $C_c(\mathbb{R}^n)$ is dense in L^p , the term $\|f - g\|_p$ may be made as small as we please. ■

Definition 11.14. Suppose that (X, τ) is a topological space and μ is a measure on $\mathcal{B}_X = \sigma(\tau)$. For a measurable function $f : X \rightarrow \mathbb{C}$ we define the essential support of f by

(11.5)

$$\text{supp}_\mu(f) = \{x \in X : \mu(\{y \in V : f(y) \neq 0\}) > 0 \text{ for all neighborhoods } V \text{ of } x\}.$$

It is not hard to show that if $\text{supp}(\mu) = X$ (see Definition 9.41) and $f \in C(X)$ then $\text{supp}_\mu(f) = \text{supp}(f) := \overline{\{f \neq 0\}}$, see Exercise 11.5.

Lemma 11.15. *Suppose (X, τ) is second countable and $f : X \rightarrow \mathbb{C}$ is a measurable function and μ is a measure on \mathcal{B}_X . Then $X := U \setminus \text{supp}_\mu(f)$ may be described as the largest open set W such that $f1_W(x) = 0$ for μ -a.e. x . Equivalently put, $C := \text{supp}_\mu(f)$ is the smallest closed subset of X such that $f = f1_C$ a.e.*

Proof. To verify that the two descriptions of $\text{supp}_\mu(f)$ are equivalent, suppose $\text{supp}_\mu(f)$ is defined as in Eq. (11.5) and $W := X \setminus \text{supp}_\mu(f)$. Then

$$\begin{aligned} W &= \{x \in X : \mu(\{y \in V : f(y) \neq 0\}) = 0 \text{ for some neighborhood } V \text{ of } x\} \\ &= \cup \{V \subset_o X : \mu(f1_V \neq 0) = 0\} \\ &= \cup \{V \subset_o X : f1_V = 0 \text{ for } \mu\text{-a.e.}\}. \end{aligned}$$

So to finish the argument it suffices to show $\mu(f1_W \neq 0) = 0$. To this let \mathcal{U} be a countable base for τ and set

$$\mathcal{U}_f := \{V \in \mathcal{U} : f1_V = 0 \text{ a.e.}\}.$$

Then it is easily seen that $W = \cup \mathcal{U}_f$ and since \mathcal{U}_f is countable $\mu(f1_W \neq 0) \leq \sum_{V \in \mathcal{U}_f} \mu(f1_V \neq 0) = 0$. ■

Lemma 11.16. *Suppose $f, g, h : \mathbb{R}^n \rightarrow \mathbb{C}$ are measurable functions and assume that x is a point in \mathbb{R}^n such that $|f| * |g|(x) < \infty$ and $|f| * (|g| * |h|)(x) < \infty$, then*

- (1) $f * g(x) = g * f(x)$
- (2) $f * (g * h)(x) = (f * g) * h(x)$
- (3) *If $z \in \mathbb{R}^n$ and $\tau_z(|f| * |g|)(x) = |f| * |g|(x - z) < \infty$, then*

$$\tau_z(f * g)(x) = \tau_z f * g(x) = f * \tau_z g(x)$$

- (4) If $x \notin \overline{\text{supp}_m(f) + \text{supp}_m(g)}$ then $f * g(x) = 0$ and in particular, $\text{supp}_m(f * g) \subset \overline{\text{supp}_m(f) + \text{supp}_m(g)}$ where in defining $\text{supp}_m(f * g)$ we will use the convention that " $f * g(x) \neq 0$ " when $|f| * |g|(x) = \infty$.

Proof. For item 1.,

$$|f| * |g|(x) = \int_{\mathbb{R}^n} |f|(x-y) |g|(y) dy = \int_{\mathbb{R}^n} |f|(y) |g|(y-x) dy = |g| * |f|(x)$$

where in the second equality we made use of the fact that Lebesgue measure is invariant under the transformation $y \rightarrow x - y$. Similar computations prove all of the remaining assertions of the first three items of the lemma.

Item 4. Since $f * g(x) = \tilde{f} * \tilde{g}(x)$ if $f = \tilde{f}$ and $g = \tilde{g}$ a.e. we may, by replacing f by $f 1_{\text{supp}_m(f)}$ and g by $g 1_{\text{supp}_m(g)}$ if necessary, assume that $\{f \neq 0\} \subset \text{supp}_m(f)$ and $\{g \neq 0\} \subset \text{supp}_m(g)$. So if $x \notin (\text{supp}_m(f) + \text{supp}_m(g))$ then $x \notin (\{f \neq 0\} + \{g \neq 0\})$ and for all $y \in \mathbb{R}^n$, either $x - y \notin \{f \neq 0\}$ or $y \notin \{g \neq 0\}$. That is to say either $x - y \in \{f = 0\}$ or $y \in \{g = 0\}$ and hence $f(x-y)g(y) = 0$ for all y and therefore $f * g(x) = 0$. This shows that $f * g = 0$ on $\mathbb{R}^n \setminus \overline{(\text{supp}_m(f) + \text{supp}_m(g))}$ and therefore

$$\mathbb{R}^n \setminus \overline{(\text{supp}_m(f) + \text{supp}_m(g))} \subset \mathbb{R}^n \setminus \text{supp}_m(f * g),$$

i.e. $\text{supp}_m(f * g) \subset \overline{\text{supp}_m(f) + \text{supp}_m(g)}$. ■

Remark 11.17. Let A, B be closed sets of \mathbb{R}^n , it is not necessarily true that $A + B$ is still closed. For example, take

$$A = \{(x, y) : x > 0 \text{ and } y \geq 1/x\} \text{ and } B = \{(x, y) : x < 0 \text{ and } y \geq 1/|x|\},$$

then every point of $A + B$ has a positive y -component and hence is not zero. On the other hand, for $x > 0$ we have $(x, 1/x) + (-x, 1/x) = (0, 2/x) \in A + B$ for all x and hence $0 \in \overline{A + B}$ showing $A + B$ is not closed. Nevertheless if one of the sets A or B is compact, then $A + B$ is closed again. Indeed, if A is compact and $x_n = a_n + b_n \in A + B$ and $x_n \rightarrow x \in \mathbb{R}^n$, then by passing to a subsequence if necessary we may assume $\lim_{n \rightarrow \infty} a_n = a \in A$ exists. In this case

$$\lim_{n \rightarrow \infty} b_n = \lim_{n \rightarrow \infty} (x_n - a_n) = x - a \in B$$

exists as well, showing $x = a + b \in A + B$.

Proposition 11.18. Suppose that $p, q \in [1, \infty]$ and p and q are conjugate exponents, $f \in L^p$ and $g \in L^q$, then $f * g \in BC(\mathbb{R}^n)$, $\|f * g\|_u \leq \|f\|_p \|g\|_q$ and if $p, q \in (1, \infty)$ then $f * g \in C_0(\mathbb{R}^n)$.

Proof. The existence of $f * g(x)$ and the estimate $|f * g|(x) \leq \|f\|_p \|g\|_q$ for all $x \in \mathbb{R}^n$ is a simple consequence of Hölder's inequality and the translation invariance of Lebesgue measure. In particular this shows $\|f * g\|_u \leq \|f\|_p \|g\|_q$. By relabeling p and q if necessary we may assume that $p \in [1, \infty)$. Since

$$\|\tau_z(f * g) - f * g\|_u = \|\tau_z f * g - f * g\|_u \leq \|\tau_z f - f\|_p \|g\|_q \rightarrow 0 \text{ as } z \rightarrow 0$$

it follows that $f * g$ is uniformly continuous. Finally if $p, q \in (1, \infty)$, we learn from Lemma 11.16 and what we have just proved that $f_m * g_m \in C_c(\mathbb{R}^n)$ where

$f_m = f1_{|f| \leq m}$ and $g_m = g1_{|g| \leq m}$. Moreover,

$$\begin{aligned} \|f * g - f_m * g_m\|_u &\leq \|f * g - f_m * g\|_u + \|f_m * g - f_m * g_m\|_u \\ &\leq \|f - f_m\|_p \|g\|_q + \|f_m\|_p \|g - g_m\|_q \\ &\leq \|f - f_m\|_p \|g\|_q + \|f\|_p \|g - g_m\|_q \rightarrow 0 \text{ as } m \rightarrow \infty \end{aligned}$$

showing, with the aid of Proposition 10.30, $f * g \in C_0(\mathbb{R}^n)$. ■

Theorem 11.19 (Young's Inequality). *Let $p, q, r \in [1, \infty]$ satisfy*

$$(11.6) \quad \frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}.$$

*If $f \in L^p$ and $g \in L^q$ then $|f| * |g|(x) < \infty$ for m - a.e. x and*

$$(11.7) \quad \|f * g\|_r \leq \|f\|_p \|g\|_q.$$

*In particular L^1 is closed under convolution. (The space $(L^1, *)$ is an example of a "Banach algebra" without unit.)*

Remark 11.20. Before going to the formal proof, let us first understand Eq. (11.6) by the following scaling argument. For $\lambda > 0$, let $f_\lambda(x) := f(\lambda x)$, then after a few simple change of variables we find

$$\|f_\lambda\|_p = \lambda^{-1/p} \|f\| \quad \text{and} \quad (f * g)_\lambda = \lambda f_\lambda * g_\lambda.$$

Therefore if Eq. (11.7) holds for some $p, q, r \in [1, \infty]$, we would also have

$$\|f * g\|_r = \lambda^{1/r} \|(f * g)_\lambda\|_r \leq \lambda^{1/r} \lambda \|f_\lambda\|_p \|g_\lambda\|_q = \lambda^{(1+1/r-1/p-1/q)} \|f\|_p \|g\|_q$$

for all $\lambda > 0$. This is only possible if Eq. (11.6) holds.

Proof. Let $\alpha, \beta \in [0, 1]$ and $p_1, p_2 \in [0, \infty]$ satisfy $p_1^{-1} + p_2^{-1} + r^{-1} = 1$. Then by Hölder's inequality, Corollary 9.3,

$$\begin{aligned} |f * g(x)| &= \left| \int f(x-y)g(y)dy \right| \leq \int |f(x-y)|^{(1-\alpha)} |g(y)|^{(1-\beta)} |f(x-y)|^\alpha |g(y)|^\beta dy \\ &\leq \left(\int |f(x-y)|^{(1-\alpha)r} |g(y)|^{(1-\beta)r} dy \right)^{1/r} \left(\int |f(x-y)|^{\alpha p_1} dy \right)^{1/p_1} \left(\int |g(y)|^{\beta p_2} dy \right)^{1/p_2} \\ &= \left(\int |f(x-y)|^{(1-\alpha)r} |g(y)|^{(1-\beta)r} dy \right)^{1/r} \|f\|_{\alpha p_1}^\alpha \|g\|_{\beta p_2}^\beta. \end{aligned}$$

Taking the r^{th} power of this equation and integrating on x gives

$$\begin{aligned} \|f * g\|_r^r &\leq \int \left(\int |f(x-y)|^{(1-\alpha)r} |g(y)|^{(1-\beta)r} dy \right) dx \cdot \|f\|_{\alpha p_1}^\alpha \|g\|_{\beta p_2}^\beta \\ (11.8) \quad &= \|f\|_{(1-\alpha)r}^{(1-\alpha)r} \|g\|_{(1-\beta)r}^{(1-\beta)r} \|f\|_{\alpha p_1}^{\alpha r} \|g\|_{\beta p_2}^{\beta r}. \end{aligned}$$

Let us now suppose, $(1-\alpha)r = \alpha p_1$ and $(1-\beta)r = \beta p_2$, in which case Eq. (11.8) becomes,

$$\|f * g\|_r^r \leq \|f\|_{\alpha p_1}^r \|g\|_{\beta p_2}^r$$

which is Eq. (11.7) with

$$(11.9) \quad p := (1-\alpha)r = \alpha p_1 \quad \text{and} \quad q := (1-\beta)r = \beta p_2.$$

So to finish the proof, it suffices to show p and q are arbitrary indices in $[1, \infty]$ satisfying $p^{-1} + q^{-1} = 1 + r^{-1}$.

If α, β, p_1, p_2 satisfy the relations above, then

$$\alpha = \frac{r}{r+p_1} \text{ and } \beta = \frac{r}{r+p_2}$$

and

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{p_1} \frac{r+p_1}{r} + \frac{1}{p_2} \frac{r+p_2}{r} = \frac{1}{p_1} + \frac{1}{p_2} + \frac{2}{r} = 1 + \frac{1}{r}.$$

Conversely, if p, q, r satisfy Eq. (11.6), then let α and β satisfy $p = (1-\alpha)r$ and $q = (1-\beta)r$, i.e.

$$\alpha := \frac{r-p}{r} = 1 - \frac{p}{r} \leq 1 \text{ and } \beta = \frac{r-q}{r} = 1 - \frac{q}{r} \leq 1.$$

From Eq. (11.6), $\alpha = p(1-\frac{1}{q}) \geq 0$ and $\beta = q(1-\frac{1}{p}) \geq 0$, so that $\alpha, \beta \in [0, 1]$. We then define $p_1 := p/\alpha$ and $p_2 := q/\beta$, then

$$\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{r} = \beta \frac{1}{q} + \alpha \frac{1}{p} + \frac{1}{r} = \frac{1}{q} - \frac{1}{r} + \frac{1}{p} - \frac{1}{r} + \frac{1}{r} = 1$$

as desired. ■

Theorem 11.21 (Approximate δ -functions). *Let $p \in [1, \infty]$, $\phi \in L^1(\mathbb{R}^n)$, $a := \int_{\mathbb{R}^n} f(x)dx$, and for $t > 0$ let $\phi_t(x) = t^{-n}\phi(x/t)$. Then*

- (1) *If $f \in L^p$ with $p < \infty$ then $\phi_t * f \rightarrow af$ in L^p as $t \downarrow 0$.*
- (2) *If $f \in BC(\mathbb{R}^n)$ and f is uniformly continuous then $\|\phi_t * f - f\|_\infty \rightarrow 0$ as $t \downarrow 0$.*
- (3) *If $f \in L^\infty$ and f is continuous on $U \subset_o \mathbb{R}^n$ then $\phi_t * f \rightarrow af$ uniformly on compact subsets of U as $t \downarrow 0$.*

Proof. Making the change of variables $y = tz$ implies

$$\phi_t * f(x) = \int_{\mathbb{R}^n} f(x-y)\phi_t(y)dy = \int_{\mathbb{R}^n} f(x-tz)\phi(z)dz$$

so that

$$\begin{aligned} \phi_t * f(x) - af(x) &= \int_{\mathbb{R}^n} [f(x-tz) - f(x)]\phi(z)dz \\ (11.10) \qquad \qquad &= \int_{\mathbb{R}^n} [\tau_{tz}f(x) - f(x)]\phi(z)dz. \end{aligned}$$

Hence by Minkowski's inequality for integrals (Theorem 9.27), Proposition 11.13 and the dominated convergence theorem,

$$\|\phi_t * f - af\|_p \leq \int_{\mathbb{R}^n} \|\tau_{tz}f - f\|_p |\phi(z)| dz \rightarrow 0 \text{ as } t \downarrow 0.$$

Item 2. is proved similarly. Indeed, from Eq. (11.10)

$$\|\phi_t * f - af\|_\infty \leq \int_{\mathbb{R}^n} \|\tau_{tz}f - f\|_\infty |\phi(z)| dz$$

which again tends to zero by the dominated convergence theorem because $\lim_{t \downarrow 0} \|\tau_{tz}f - f\|_\infty = 0$ uniformly in z by the uniform continuity of f .

Item 3. Let $B_R = B(0, R)$ be a large ball in \mathbb{R}^n and $K \sqsubset\sqsubset U$, then

$$\begin{aligned} \sup_{x \in K} |\phi_t * f(x) - af(x)| &\leq \left| \int_{B_R} [f(x-tz) - f(x)] \phi(z) dz \right| + \left| \int_{B_R^c} [f(x-tz) - f(x)] \phi(z) dz \right| \\ &\leq \int_{B_R} |\phi(z)| dz \cdot \sup_{x \in K, z \in B_R} |f(x-tz) - f(x)| + 2 \|f\|_\infty \int_{B_R^c} |\phi(z)| dz \\ &\leq \|\phi\|_1 \cdot \sup_{x \in K, z \in B_R} |f(x-tz) - f(x)| + 2 \|f\|_\infty \int_{|z| > R} |\phi(z)| dz \end{aligned}$$

so that using the uniform continuity of f on compact subsets of U ,

$$\limsup_{t \downarrow 0} \sup_{x \in K} |\phi_t * f(x) - af(x)| \leq 2 \|f\|_\infty \int_{|z| > R} |\phi(z)| dz \rightarrow 0 \text{ as } R \rightarrow \infty.$$

■

See Theorem 8.15 if Folland for a statement about almost everywhere convergence.

Exercise 11.1. Let

$$f(t) = \begin{cases} e^{-1/t} & \text{if } t > 0 \\ 0 & \text{if } t \leq 0. \end{cases}$$

Show $f \in C^\infty(\mathbb{R}, [0, 1])$.

Lemma 11.22. *There exists $\phi \in C_c^\infty(\mathbb{R}^n, [0, \infty))$ such that $\phi(0) > 0$, $\text{supp}(\phi) \subset \bar{B}(0, 1)$ and $\int_{\mathbb{R}^n} \phi(x) dx = 1$.*

Proof. Define $h(t) = f(1-t)f(t+1)$ where f is as in Exercise 11.1. Then $h \in C_c^\infty(\mathbb{R}, [0, 1])$, $\text{supp}(h) \subset [-1, 1]$ and $h(0) = e^{-2} > 0$. Define $c = \int_{\mathbb{R}^n} h(|x|^2) dx$. Then $\phi(x) = c^{-1}h(|x|^2)$ is the desired function. ■

Definition 11.23. Let $X \subset \mathbb{R}^n$ be an open set. A **Radon** measure on \mathcal{B}_X is a measure μ which is finite on compact subsets of X . For a Radon measure μ , we let $L_{loc}^1(\mu)$ consists of those measurable functions $f : X \rightarrow \mathbb{C}$ such that $\int_K |f| d\mu < \infty$ for all compact subsets $K \subset X$.

The reader asked to prove the following proposition in Exercise 11.6 below.

Proposition 11.24. *Suppose that $f \in L_{loc}^1(\mathbb{R}^n, m)$ and $\phi \in C_c^1(\mathbb{R}^n)$, then $f * \phi \in C^1(\mathbb{R}^n)$ and $\partial_i(f * \phi) = f * \partial_i \phi$. Moreover if $\phi \in C_c^\infty(\mathbb{R}^n)$ then $f * \phi \in C^\infty(\mathbb{R}^n)$.*

Corollary 11.25 (C^∞ - Uryhson's Lemma). *Given $K \sqsubset\sqsubset U \subset_o \mathbb{R}^n$, there exists $f \in C_c^\infty(\mathbb{R}^n, [0, 1])$ such that $\text{supp}(f) \subset U$ and $f = 1$ on K .*

Proof. Let ϕ be as in Lemma 11.22, $\phi_t(x) = t^{-n}\phi(x/t)$ be as in Theorem 11.21, d be the standard metric on \mathbb{R}^n and $\epsilon = d(K, U^c)$. Since K is compact and U^c is closed, $\epsilon > 0$. Let $V_\delta = \{x \in \mathbb{R}^n : d(x, K) < \delta\}$ and $f = \phi_{\epsilon/3} * 1_{V_{\epsilon/3}}$, then

$$\text{supp}(f) \subset \overline{\text{supp}(\phi_{\epsilon/3}) + V_{\epsilon/3}} \subset \bar{V}_{2\epsilon/3} \subset U.$$

Since $\bar{V}_{2\epsilon/3}$ is closed and bounded, $f \in C_c^\infty(U)$ and for $x \in K$,

$$f(x) = \int_{\mathbb{R}^n} 1_{d(y, K) < \epsilon/3} \cdot \phi_{\epsilon/3}(x-y) dy = \int_{\mathbb{R}^n} \phi_{\epsilon/3}(x-y) dy = 1.$$

The proof will be finished after the reader (easily) verifies $0 \leq f \leq 1$. ■

Here is an application of this corollary whose proof is left to the reader, Exercise 11.7.

Lemma 11.26 (Integration by Parts). *Suppose f and g are measurable functions on \mathbb{R}^n such that $t \rightarrow f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)$ and $t \rightarrow g(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)$ are continuously differentiable functions on \mathbb{R} for each fixed $x = (x_1, \dots, x_n) \in \mathbb{R}^n$. Moreover assume $f \cdot g$, $\frac{\partial f}{\partial x_i} \cdot g$ and $f \cdot \frac{\partial g}{\partial x_i}$ are in $L^1(\mathbb{R}^n, m)$. Then*

$$\int_{\mathbb{R}^n} \frac{\partial f}{\partial x_i} \cdot g dm = - \int_{\mathbb{R}^n} f \cdot \frac{\partial g}{\partial x_i} dm.$$

With this result we may give another proof of the Riemann Lebesgue Lemma.

Lemma 11.27. *For $f \in L^1(\mathbb{R}^n, m)$ let*

$$\hat{f}(\xi) := (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x) e^{-i\xi \cdot x} dm(x)$$

be the Fourier transform of f . Then $\hat{f} \in C_0(\mathbb{R}^n)$ and $\|\hat{f}\|_u \leq (2\pi)^{-n/2} \|f\|_1$. (The choice of the normalization factor, $(2\pi)^{-n/2}$, in \hat{f} is for later convenience.)

Proof. The fact that \hat{f} is continuous is a simple application of the dominated convergence theorem. Moreover,

$$|\hat{f}(\xi)| \leq \int |f(x)| dm(x) \leq (2\pi)^{-n/2} \|f\|_1$$

so it only remains to see that $\hat{f}(\xi) \rightarrow 0$ as $|\xi| \rightarrow \infty$.

First suppose that $f \in C_c^\infty(\mathbb{R}^n)$ and let $\Delta = \sum_{j=1}^n \frac{\partial^2}{\partial x_j^2}$ be the Laplacian on \mathbb{R}^n . Notice that $\frac{\partial}{\partial x_j} e^{-i\xi \cdot x} = -i\xi_j e^{-i\xi \cdot x}$ and $\Delta e^{-i\xi \cdot x} = -|\xi|^2 e^{-i\xi \cdot x}$. Using Lemma 11.26 repeatedly,

$$\begin{aligned} \int \Delta^k f(x) e^{-i\xi \cdot x} dm(x) &= \int f(x) \Delta_x^k e^{-i\xi \cdot x} dm(x) = -|\xi|^{2k} \int f(x) e^{-i\xi \cdot x} dm(x) \\ &= -(2\pi)^{n/2} |\xi|^{2k} \hat{f}(\xi) \end{aligned}$$

for any $k \in \mathbb{N}$. Hence $(2\pi)^{n/2} |\hat{f}(\xi)| \leq |\xi|^{-2k} \|\Delta^k f\|_1 \rightarrow 0$ as $|\xi| \rightarrow \infty$ and $\hat{f} \in C_0(\mathbb{R}^n)$. Suppose that $f \in L^1(m)$ and $f_k \in C_c^\infty(\mathbb{R}^n)$ is a sequence such that $\lim_{k \rightarrow \infty} \|f - f_k\|_1 = 0$, then $\lim_{k \rightarrow \infty} \|\hat{f} - \hat{f}_k\|_u = 0$. Hence $\hat{f} \in C_0(\mathbb{R}^n)$ by an application of Proposition 10.30. ■

Corollary 11.28. *Let $X \subset \mathbb{R}^n$ be an open set and μ be a Radon measure on \mathcal{B}_X .*

- (1) *Then $C_c^\infty(X)$ is dense in $L^p(\mu)$ for all $1 \leq p < \infty$.*
- (2) *If $h \in L_{loc}^1(\mu)$ satisfies*

$$(11.11) \quad \int_X f h d\mu = 0 \text{ for all } f \in C_c^\infty(X)$$

then $h(x) = 0$ for μ - a.e. x .

Proof. Let $f \in C_c(X)$, ϕ be as in Lemma 11.22, ϕ_t be as in Theorem 11.21 and set $\psi_t := \phi_t * (f1_X)$. Then by Proposition 11.24 $\psi_t \in C^\infty(X)$ and by Lemma 11.16 there exists a compact set $K \subset X$ such that $\text{supp}(\psi_t) \subset K$ for all t sufficiently small. By Theorem 11.21, $\psi_t \rightarrow f$ uniformly on X as $t \downarrow 0$

- (1) The dominated convergence theorem (with dominating function being $\|f\|_\infty 1_K$), shows $\psi_t \rightarrow f$ in $L^p(\mu)$ as $t \downarrow 0$. This proves Item 1., since Proposition 11.6 guarantees that $C_c(X)$ is dense in $L^p(\mu)$.

- (2) Keeping the same notation as above, the dominated convergence theorem (with dominating function being $\|f\|_\infty |h| 1_K$) implies

$$0 = \lim_{t \downarrow 0} \int_X \psi_t h d\mu = \int_X \lim_{t \downarrow 0} \psi_t h d\mu = \int_X f h d\mu.$$

The proof is now finished by an application of Lemma 11.7.

■

11.1.1. *Smooth Partitions of Unity.* We have the following smooth variants of Proposition 10.24, Theorem 10.26 and Corollary 10.27. The proofs of these results are the same as their continuous counterparts. One simply uses the smooth version of Urysohn's Lemma of Corollary 11.25 in place of Lemma 10.15.

Proposition 11.29 (Smooth Partitions of Unity for Compacts). *Suppose that X is an open subset of \mathbb{R}^n , $K \subset X$ is a compact set and $\mathcal{U} = \{U_j\}_{j=1}^n$ is an open cover of K . Then there exists a smooth (i.e. $h_j \in C^\infty(X, [0, 1])$) partition of unity $\{h_j\}_{j=1}^n$ of K such that $h_j \prec U_j$ for all $j = 1, 2, \dots, n$.*

Theorem 11.30 (Locally Compact Partitions of Unity). *Suppose that X is an open subset of \mathbb{R}^n and \mathcal{U} is an open cover of X . Then there exists a smooth partition of unity of $\{h_i\}_{i=1}^N$ ($N = \infty$ is allowed here) subordinate to the cover \mathcal{U} such that $\text{supp}(h_i)$ is compact for all i .*

Corollary 11.31. *Suppose that X is an open subset of \mathbb{R}^n and $\mathcal{U} = \{U_\alpha\}_{\alpha \in A} \subset \tau$ is an open cover of X . Then there exists a smooth partition of unity of $\{h_\alpha\}_{\alpha \in A}$ subordinate to the cover \mathcal{U} such that $\text{supp}(h_\alpha) \subset U_\alpha$ for all $\alpha \in A$. Moreover if \bar{U}_α is compact for each $\alpha \in A$ we may choose h_α so that $h_\alpha \prec U_\alpha$.*

11.2. **Classical Weierstrass Approximation Theorem.** Let $\mathbb{Z}_+ := \mathbb{N} \cup \{0\}$.

Notation 11.32. For $x \in \mathbb{R}^d$ and $\alpha \in \mathbb{Z}_+^d$ let $x^\alpha = \prod_{i=1}^d x_i^{\alpha_i}$ and $|\alpha| = \sum_{i=1}^d \alpha_i$. A polynomial on \mathbb{R}^d is a function $p : \mathbb{R}^d \rightarrow \mathbb{C}$ of the form

$$p(x) = \sum_{\alpha: |\alpha| \leq N} p_\alpha x^\alpha \text{ with } p_\alpha \in \mathbb{C} \text{ and } N \in \mathbb{Z}_+.$$

If $p_\alpha \neq 0$ for some α such that $|\alpha| = N$, then we define $\deg(p) := N$ to be the degree of p . The function p has a natural extension to $z \in \mathbb{C}^d$, namely $p(z) = \sum_{\alpha: |\alpha| \leq N} p_\alpha z^\alpha$ where $z^\alpha = \prod_{i=1}^d z_i^{\alpha_i}$.

Remark 11.33. The mapping $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d \rightarrow z = x + iy \in \mathbb{C}^d$ is an isomorphism of vector spaces. Letting $\bar{z} = x - iy$ as usual, we have $x = \frac{z + \bar{z}}{2}$ and $y = \frac{z - \bar{z}}{2i}$. Therefore under this identification any polynomial $p(x, y)$ on $\mathbb{R}^d \times \mathbb{R}^d$ may be written as a polynomial q in (z, \bar{z}) , namely

$$q(z, \bar{z}) = p\left(\frac{z + \bar{z}}{2}, \frac{z - \bar{z}}{2i}\right).$$

Conversely a polynomial q in (z, \bar{z}) may be thought of as a polynomial p in (x, y) , namely $p(x, y) = q(x + iy, x - iy)$.

Theorem 11.34 (Weierstrass Approximation Theorem). *Let $a, b \in \mathbb{R}^d$ with $a \leq b$ (i.e. $a_i \leq b_i$ for $i = 1, 2, \dots, d$) and set $[a, b] := [a_1, b_1] \times \dots \times [a_d, b_d]$. Then for $f \in C([a, b], \mathbb{C})$ there exists polynomials p_n on \mathbb{R}^d such that $p_n \rightarrow f$ uniformly on $[a, b]$.*

We will give two proofs of this theorem below. The first proof is based on the “weak law of large numbers,” while the second is based on using a certain sequence of approximate δ -functions.

Corollary 11.35. *Suppose that $K \subset \mathbb{R}^d$ is a compact set and $f \in C(K, \mathbb{C})$. Then there exists polynomials p_n on \mathbb{R}^d such that $p_n \rightarrow f$ uniformly on K .*

Proof. Choose $a, b \in \mathbb{R}^d$ such that $a \leq b$ and $K \subset (a, b) := (a_1, b_1) \times \cdots \times (a_d, b_d)$. Let $\tilde{f} : K \cup (a, b)^c \rightarrow \mathbb{C}$ be the continuous function defined by $\tilde{f}|_K = f$ and $\tilde{f}|_{(a, b)^c} \equiv 0$. Then by the Tietze extension Theorem (either of Theorems 10.2 or 10.16 will do) there exists $F \in C(\mathbb{R}^d, \mathbb{C})$ such that $\tilde{f} = F|_{K \cup (a, b)^c}$. Apply the Weierstrass Approximation Theorem 11.34 to $F|_{[a, b]}$ to find polynomials p_n on \mathbb{R}^d such that $p_n \rightarrow F$ uniformly on $[a, b]$. Clearly we also have $p_n \rightarrow f$ uniformly on K . ■

Corollary 11.36 (Complex Weierstrass Approximation Theorem). *Suppose that $K \subset \mathbb{C}^d$ is a compact set and $f \in C(K, \mathbb{C})$. Then there exists polynomials $p_n(z, \bar{z})$ for $z \in \mathbb{C}^d$ such that $\sup_{z \in K} |p_n(z, \bar{z}) - f(z)| \rightarrow 0$ as $n \rightarrow \infty$.*

Proof. This is an immediate consequence of Remark 11.33 and Corollary 11.35. ■

Example 11.37. Let $K = S^1 = \{z \in \mathbb{C} : |z| = 1\}$ and \mathcal{A} be the set of polynomials in (z, \bar{z}) restricted to S^1 . Then \mathcal{A} is dense in $C(S^1)$.²³ Since $\bar{z} = z^{-1}$ on S^1 , we have shown polynomials in z and z^{-1} are dense in $C(S^1)$. This example generalizes in an obvious way to $K = (S^1)^d \subset \mathbb{C}^d$.

11.2.1. *First proof of the Weierstrass Approximation Theorem 11.34.* **Proof.** Let $\mathbf{0} := (0, 0, \dots, 0)$ and $\mathbf{1} := (1, 1, \dots, 1)$. By considering the real and imaginary parts of f separately, it suffices to assume f is real valued. By replacing f by $g(x) = f(a_1 + x_1(b_1 - a_1), \dots, a_d + x_d(b_d - a_d))$ for $x \in [0, 1]$, it suffices to prove the theorem for $f \in C([0, 1])$.

For $x \in [0, 1]$, let ν_x be the measure on $\{0, 1\}$ such that $\nu_x(\{0\}) = 1 - x$ and $\nu_x(\{1\}) = x$. Then

$$(11.12) \quad \int_{\{0,1\}} y d\nu_x(y) = 0 \cdot (1 - x) + 1 \cdot x = x \text{ and}$$

$$(11.13) \quad \int_{\{0,1\}} (y - x)^2 d\nu_x(y) = x^2(1 - x) + (1 - x)^2 \cdot x = x(1 - x).$$

For $x \in [0, 1]$ let $\mu_x = \nu_{x_1} \otimes \cdots \otimes \nu_{x_d}$ be the product of $\nu_{x_1}, \dots, \nu_{x_d}$ on $\Omega := \{0, 1\}^d$. Alternatively the measure μ_x may be described by

$$(11.14) \quad \mu_x(\{\epsilon\}) = \prod_{i=1}^d (1 - x_i)^{1 - \epsilon_i} x_i^{\epsilon_i}$$

for $\epsilon \in \Omega$. Notice that $\mu_x(\{\epsilon\})$ is a degree d polynomial in x for each $\epsilon \in \Omega$. For $n \in \mathbb{N}$ and $x \in [0, 1]$, let μ_x^n denote the n -fold product of μ_x with itself on Ω^n , $X_i(\omega) = \omega_i \in \Omega \subset \mathbb{R}^d$ for $\omega \in \Omega^n$ and let

$$S_n = (S_n^1, \dots, S_n^d) := (X_1 + X_2 + \cdots + X_n)/n,$$

²³Note that it is easy to extend $f \in C(S^1)$ to a function $F \in C(\mathbb{C})$ by setting $F(z) = zf(\frac{z}{|z|})$ for $z \neq 0$ and $F(0) = 0$. So this special case does not require the Tietze extension theorem.

so $S_n : \Omega^n \rightarrow \mathbb{R}^d$. The reader is asked to verify (Exercise 11.2) that

$$(11.15) \quad \int_{\Omega^n} S_n d\mu_x^n = \left(\int_{\Omega^n} S_n^1 d\mu_x^n, \dots, \int_{\Omega^n} S_n^d d\mu_x^n \right) = (x_1, \dots, x_d) = x$$

and

$$(11.16) \quad \int_{\Omega^n} |S_n - x|^2 d\mu_x^n = \frac{1}{n} \sum_{i=1}^d x_i(1 - x_i) \leq \frac{d}{n}.$$

From these equations it follows that S_n is concentrating near x as $n \rightarrow \infty$, a manifestation of the law of large numbers. Therefore it is reasonable to expect

$$(11.17) \quad p_n(x) := \int_{\Omega^n} f(S_n) d\mu_x^n$$

should approach $f(x)$ as $n \rightarrow \infty$.

Let $\epsilon > 0$ be given, $M = \sup \{|f(x)| : x \in [0, 1]\}$ and

$$\delta_\epsilon = \sup \{|f(y) - f(x)| : x, y \in [0, 1] \text{ and } |y - x| \leq \epsilon\}.$$

By uniform continuity of f on $[0, 1]$, $\lim_{\epsilon \downarrow 0} \delta_\epsilon = 0$. Using these definitions and the fact that $\mu_x^n(\Omega^n) = 1$,

$$(11.18) \quad \begin{aligned} |f(x) - p_n(x)| &= \left| \int_{\Omega^n} (f(x) - f(S_n)) d\mu_x^n \right| \leq \int_{\Omega^n} |f(x) - f(S_n)| d\mu_x^n \\ &\leq \int_{\{|S_n - x| > \epsilon\}} |f(x) - f(S_n)| d\mu_x^n + \int_{\{|S_n - x| \leq \epsilon\}} |f(x) - f(S_n)| d\mu_x^n \\ &\leq 2M\mu_x^n(|S_n - x| > \epsilon) + \delta_\epsilon. \end{aligned}$$

By Chebyshev's inequality,

$$\mu_x^n(|S_n - x| > \epsilon) \leq \frac{1}{\epsilon^2} \int_{\Omega^n} (S_n - x)^2 d\mu_x^n = \frac{d}{n\epsilon^2},$$

and therefore, Eq. (11.18) yields the estimate

$$\|f - p_n\|_u \leq \frac{2dM}{n\epsilon^2} + \delta_\epsilon$$

and hence

$$\limsup_{n \rightarrow \infty} \|f - p_n\|_u \leq \delta_\epsilon \rightarrow 0 \text{ as } \epsilon \downarrow 0.$$

This completes the proof since, using Eq. (11.14),

$$p_n(x) = \sum_{\omega \in \Omega^n} f(S_n(\omega)) \mu_x^n(\{\omega\}) = \sum_{\omega \in \Omega^n} f(S_n(\omega)) \prod_{i=1}^n \mu_x(\{\omega_i\}),$$

is an nd -degree polynomial in $x \in \mathbb{R}^d$. ■

Exercise 11.2. Verify Eqs. (11.15) and (11.16). This is most easily done using Eqs. (11.12) and (11.13) and Fubini's theorem repeatedly. (Of course Fubini's theorem here is overkill since these are only finite sums after all. Nevertheless it is convenient to use this formulation.)

11.2.2. *Second proof of the Weierstrass Approximation Theorem 11.34.* For the second proof we will first need two lemmas.

Lemma 11.38 (Approximate δ - sequences). *Suppose that $\{Q_n\}_{n=1}^\infty$ is a sequence of positive functions on \mathbb{R}^d such that*

$$(11.19) \quad \int_{\mathbb{R}^d} Q_n(x) dx = 1 \text{ and}$$

$$(11.20) \quad \lim_{n \rightarrow \infty} \int_{|x| \geq \epsilon} Q_n(x) dx = 0 \text{ for all } \epsilon > 0.$$

For $f \in BC(\mathbb{R}^d)$, $Q_n * f$ converges to f uniformly on compact subsets of \mathbb{R}^d .

Proof. Let $x \in \mathbb{R}^d$, then because of Eq. (11.19),

$$|Q_n * f(x) - f(x)| = \left| \int_{\mathbb{R}^d} Q_n(y) (f(x-y) - f(x)) dy \right| \leq \int_{\mathbb{R}^d} Q_n(y) |f(x-y) - f(x)| dy.$$

Let $M = \sup \{|f(x)| : x \in \mathbb{R}^d\}$ and $\epsilon > 0$, then by and Eq. (11.19)

$$\begin{aligned} |Q_n * f(x) - f(x)| &\leq \int_{|y| \leq \epsilon} Q_n(y) |f(x-y) - f(x)| dy \\ &\quad + \int_{|y| > \epsilon} Q_n(y) |f(x-y) - f(x)| dy \\ &\leq \sup_{|z| \leq \epsilon} |f(x+z) - f(x)| + 2M \int_{|y| > \epsilon} Q_n(y) dy. \end{aligned}$$

Let K be a compact subset of \mathbb{R}^d , then

$$\sup_{x \in K} |Q_n * f(x) - f(x)| \leq \sup_{|z| \leq \epsilon, x \in K} |f(x+z) - f(x)| + 2M \int_{|y| > \epsilon} Q_n(y) dy$$

and hence by Eq. (11.20),

$$\limsup_{n \rightarrow \infty} \sup_{x \in K} |Q_n * f(x) - f(x)| \leq \sup_{|z| \leq \epsilon, x \in K} |f(x+z) - f(x)|.$$

This finishes the proof since the right member of this equation tends to 0 as $\epsilon \downarrow 0$ by uniform continuity of f on compact subsets of \mathbb{R}^n . ■

Let $q_n : \mathbb{R} \rightarrow [0, \infty)$ be defined by

$$(11.21) \quad q_n(x) \equiv \frac{1}{c_n} (1-x^2)^n 1_{|x| \leq 1} \text{ where } c_n := \int_{-1}^1 (1-x^2)^n dx.$$

Figure 26 displays the key features of the functions q_n .

Define

$$(11.22) \quad Q_n : \mathbb{R}^n \rightarrow [0, \infty) \text{ by } Q_n(x) = q_n(x_1) \dots q_n(x_d).$$

Lemma 11.39. *The sequence $\{Q_n\}_{n=1}^\infty$ is an approximate δ - sequence, i.e. they satisfy Eqs. (11.19) and (11.20).*

Proof. The fact that Q_n integrates to one is an easy consequence of Tonelli's theorem and the definition of c_n . Since all norms on \mathbb{R}^d are equivalent, we may assume that $|x| = \max \{|x_i| : i = 1, 2, \dots, d\}$ when proving Eq. (11.20). With this norm

$$\{x \in \mathbb{R}^d : |x| \geq \epsilon\} = \cup_{i=1}^d \{x \in \mathbb{R}^d : |x_i| \geq \epsilon\}$$

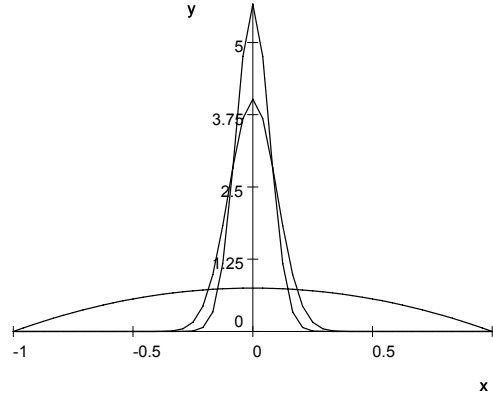


FIGURE 26. A plot of q_1 , q_{50} , and q_{100} . The most peaked curve is q_{100} and the least is q_1 . The total area under each of these curves is one.

and therefore by Tonelli's theorem and the definition of c_n ,

$$\int_{\{|x| \geq \epsilon\}} Q_n(x) dx \leq \sum_{i=1}^d \int_{\{|x_i| \geq \epsilon\}} Q_n(x) dx = d \int_{\{x \in \mathbb{R}^d : |x| \geq \epsilon\}} q_n(x) dx.$$

Since

$$\begin{aligned} \int_{|x| \geq \epsilon} q_n(x) dx &= \frac{2 \int_{\epsilon}^1 (1-x^2)^n dx}{2 \int_0^{\epsilon} (1-x^2)^n dx + 2 \int_{\epsilon}^1 (1-x^2)^n dx} \\ &\leq \frac{\int_{\epsilon}^1 \frac{x}{\epsilon} (1-x^2)^n dx}{\int_0^{\epsilon} \frac{x}{\epsilon} (1-x^2)^n dx} = \frac{(1-x^2)^{n+1} \Big|_{\epsilon}^1}{(1-x^2)^{n+1} \Big|_0^{\epsilon}} = \frac{(1-\epsilon^2)^{n+1}}{1 - (1-\epsilon^2)^{n+1}} \rightarrow 0 \text{ as } n \rightarrow \infty, \end{aligned}$$

the proof is complete. ■

We will now prove Corollary 11.35 which clearly implies Theorem 11.34.

Proof. Proof of Corollary 11.35. As in the beginning of the proof already given for Corollary 11.35, we may assume that $K = [a, b]$ for some $a \leq b$ and $f = F|_K$ where $F \in C(\mathbb{R}^d, \mathbb{C})$ is a function such that $F|_{K^c} \equiv 0$. Moreover, by replacing $F(x)$ by $G(x) = F(a_1 + x_1(b_1 - a_1), \dots, a_d + x_d(b_d - a_d))$ for $x \in \mathbb{R}^n$ we may further assume $K = [\mathbf{0}, \mathbf{1}]$.

Let $Q_n(x)$ be defined as in Eq. (11.22). Then by Lemma 11.39 and 11.38, $p_n(x) := (Q_n * F)(x) \rightarrow F(x)$ uniformly for $x \in [\mathbf{0}, \mathbf{1}]$ as $n \rightarrow \infty$. So to finish the

proof it only remains to show $p_n(x)$ is a polynomial when $x \in [0, 1]$. For $x \in [0, 1]$,

$$\begin{aligned} p_n(x) &= \int_{\mathbb{R}^d} Q_n(x-y)f(y)dy \\ &= \frac{1}{c_n} \int_{[0,1]} f(y) \prod_{i=1}^d [c_n^{-1}(1-(x_i-y_i)^2)^n 1_{|x_i-y_i| \leq 1}] dy \\ &= \frac{1}{c_n} \int_{[0,1]} f(y) \prod_{i=1}^d [c_n^{-1}(1-(x_i-y_i)^2)^n] dy. \end{aligned}$$

Since the product in the above integrand is a polynomial if $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, it follows easily that $p_n(x)$ is polynomial in x . ■

11.3. Stone-Weierstrass Theorem. We now wish to generalize Theorem 11.34 to more general topological spaces. We will first need some definitions.

Definition 11.40. Let X be a topological space and $\mathcal{A} \subset C(X) = C(X, \mathbb{R})$ or $C(X, \mathbb{C})$ be a collection of functions. Then

- (1) \mathcal{A} is said to **separate points** if for all distinct points $x, y \in X$ there exists $f \in \mathcal{A}$ such that $f(x) \neq f(y)$.
- (2) \mathcal{A} is an **algebra** if \mathcal{A} is a vector subspace of $C(X)$ which is closed under pointwise multiplication.
- (3) \mathcal{A} is called a **lattice** if $f \vee g := \max(f, g)$ and $f \wedge g = \min(f, g) \in \mathcal{A}$ for all $f, g \in \mathcal{A}$.
- (4) $\mathcal{A} \subset C(X)$ is closed under conjugation if $\bar{f} \in \mathcal{A}$ whenever $f \in \mathcal{A}$.²⁴

Remark 11.41. If X is a topological space such that $C(X, \mathbb{R})$ separates points then X is Hausdorff. Indeed if $x, y \in X$ and $f \in C(X, \mathbb{R})$ such that $f(x) \neq f(y)$, then $f^{-1}(J)$ and $f^{-1}(I)$ are disjoint open sets containing x and y respectively when I and J are disjoint intervals containing $f(x)$ and $f(y)$ respectively.

Lemma 11.42. *If $\mathcal{A} \subset C(X, \mathbb{R})$ is a closed algebra then $|f| \in \mathcal{A}$ for all $f \in \mathcal{A}$ and \mathcal{A} is a lattice.*

Proof. Let $f \in \mathcal{A}$ and let $M = \sup_{x \in X} |f(x)|$. Using Theorem 11.34 or Exercise 11.8, there are polynomials $p_n(t)$ such that

$$\lim_{n \rightarrow \infty} \sup_{|t| \leq M} ||t| - p_n(t)| = 0.$$

By replacing p_n by $p_n - p_n(0)$ if necessary we may assume that $p_n(0) = 0$. Since \mathcal{A} is an algebra, it follows that $f_n = p_n(f) \in \mathcal{A}$ and $|f| \in \mathcal{A}$, because $|f|$ is the uniform limit of the f_n 's. Since

$$\begin{aligned} f \vee g &= \frac{1}{2} (f + g + |f - g|) \text{ and} \\ f \wedge g &= \frac{1}{2} (f + g - |f - g|), \end{aligned}$$

we have shown \mathcal{A} is a lattice. ■

²⁴This is of course no restriction when $C(X) = C(X, \mathbb{R})$.

Lemma 11.43. *Let $\mathcal{A} \subset C(X, \mathbb{R})$ be an algebra which separates points and $x, y \in X$ be distinct points such that*

$$(11.23) \quad \exists f, g \in \mathcal{A} \quad \ni \quad f(x) \neq 0 \text{ and } g(y) \neq 0.$$

Then

$$(11.24) \quad V := \{(f(x), f(y)) : f \in \mathcal{A}\} = \mathbb{R}^2.$$

Proof. It is clear that V is a non-zero subspace of \mathbb{R}^2 . If $\dim(V) = 1$, then $V = \text{span}(a, b)$ with $a \neq 0$ and $b \neq 0$ by the assumption in Eq. (11.23). Since $(a, b) = (f(x), f(y))$ for some $f \in \mathcal{A}$ and $f^2 \in \mathcal{A}$, it follows that $(a^2, b^2) = (f^2(x), f^2(y)) \in V$ as well. Since $\dim V = 1$, (a, b) and (a^2, b^2) are linearly dependent and therefore

$$0 = \det \begin{pmatrix} a & a^2 \\ b & b^2 \end{pmatrix} = ab^2 - ba^2 = ab(b - a)$$

which implies that $a = b$. But this implies that $f(x) = f(y)$ for all $f \in \mathcal{A}$, violating the assumption that \mathcal{A} separates points. Therefore we conclude that $\dim(V) = 2$, i.e. $V = \mathbb{R}^2$. ■

Theorem 11.44 (Stone-Weierstrass Theorem). *ppose X is a compact Hausdorff space and $\mathcal{A} \subset C(X, \mathbb{R})$ is a **closed** subalgebra which separates points. For $x \in X$ let*

$$\begin{aligned} \mathcal{A}_x &\equiv \{f(x) : f \in \mathcal{A}\} \text{ and} \\ \mathcal{I}_x &= \{f \in C(X, \mathbb{R}) : f(x) = 0\}. \end{aligned}$$

Then either one of the following two cases hold.

- (1) $\mathcal{A}_x = \mathbb{R}$ for all $x \in X$, i.e. for all $x \in X$ there exists $f \in \mathcal{A}$ such that $f(x) \neq 0$.²⁵
- (2) There exists a unique point $x_0 \in X$ such that $\mathcal{A}_{x_0} = \{0\}$.

Moreover in case (1) $\mathcal{A} = C(X, \mathbb{R})$ and in case (2) $\mathcal{A} = \mathcal{I}_{x_0} = \{f \in C(X, \mathbb{R}) : f(x_0) = 0\}$.

Proof. If there exists x_0 such that $\mathcal{A}_{x_0} = \{0\}$ (x_0 is unique since \mathcal{A} separates points) then $\mathcal{A} \subset \mathcal{I}_{x_0}$. If such an x_0 exists let $\mathcal{C} = \mathcal{I}_{x_0}$ and if $\mathcal{A}_x = \mathbb{R}$ for all x , set $\mathcal{C} = C(X, \mathbb{R})$. Let $f \in \mathcal{C}$, then by Lemma 11.43, for all $x, y \in X$ such that $x \neq y$ there exists $g_{xy} \in \mathcal{A}$ such that $f = g_{xy}$ on $\{x, y\}$.²⁶ The basic idea of the proof is contained in the following identity,

$$(11.25) \quad f(z) = \inf_{x \in X} \sup_{y \in X} g_{xy}(z) \text{ for all } z \in X.$$

To prove this identity, let $g_x := \sup_{y \in X} g_{xy}$ and notice that $g_x \geq f$ since $g_{xy}(y) = f(y)$ for all $y \in X$. Moreover, $g_x(x) = f(x)$ for all $x \in X$ since $g_{xy}(x) = f(x)$ for all x . Therefore,

$$\inf_{x \in X} \sup_{y \in X} g_{xy} = \inf_{x \in X} g_x = f.$$

The rest of the proof is devoted to replacing the inf and the sup above by min and max over finite sets at the expense of Eq. (11.25) becoming only an approximate identity.

²⁵If \mathcal{A} contains the constant function 1, then this hypothesis holds.

²⁶If $\mathcal{A}_{x_0} = \{0\}$ and $x = x_0$ or $y = x_0$, then g_{xy} exists merely by the fact that \mathcal{A} separates points.

Claim 2. Given $\epsilon > 0$ and $x \in X$ there exists $g_x \in \mathcal{A}$ such that $g_x(x) = f(x)$ and $f < g_x + \epsilon$ on X .

To prove the claim, let V_y be an open neighborhood of y such that $|f - g_{xy}| < \epsilon$ on V_y so in particular $f < \epsilon + g_{xy}$ on V_y . By compactness, there exists $\Lambda \subset\subset X$ such that $X = \bigcup_{y \in \Lambda} V_y$. Set

$$g_x(z) = \max\{g_{xy}(z) : y \in \Lambda\},$$

then for any $y \in \Lambda$, $f < \epsilon + g_{xy} < \epsilon + g_x$ on V_y and therefore $f < \epsilon + g_x$ on X . Moreover, by construction $f(x) = g_x(x)$, see Figure 27 below.

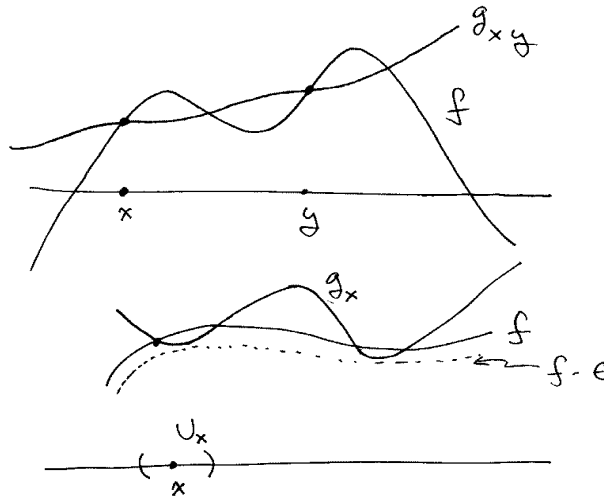


FIGURE 27. Constructing the functions g_x .

We now will finish the proof of the theorem. For each $x \in X$, let U_x be a neighborhood of x such that $|f - g_x| < \epsilon$ on U_x . Choose $\Gamma \subset\subset X$ such that $X = \bigcup_{x \in \Gamma} U_x$ and define

$$g = \min\{g_x : x \in \Gamma\} \in \mathcal{A}.$$

Then $f < g + \epsilon$ on X and for $x \in \Gamma$, $g_x < f + \epsilon$ on U_x and hence $g < f + \epsilon$ on U_x . Since $X = \bigcup_{x \in \Gamma} U_x$, we conclude

$$f < g + \epsilon \text{ and } g < f + \epsilon \text{ on } X,$$

i.e. $|f - g| < \epsilon$ on X . Since $\epsilon > 0$ is arbitrary it follows that $f \in \bar{\mathcal{A}} = \mathcal{A}$. ■

Theorem 11.45 (Complex Stone-Weierstrass Theorem). *Let X be a compact Hausdorff space. Suppose $\mathcal{A} \subset C(X, \mathbb{C})$ is closed in the uniform topology, separates points, and is closed under conjugation. Then either $\mathcal{A} = C(X, \mathbb{C})$ or $\mathcal{A} = \mathcal{I}_{x_0}^{\mathbb{C}} := \{f \in C(X, \mathbb{C}) : f(x_0) = 0\}$ for some $x_0 \in X$.*

Proof. Since

$$\operatorname{Re} f = \frac{f + \bar{f}}{2} \text{ and } \operatorname{Im} f = \frac{f - \bar{f}}{2i},$$

$\operatorname{Re} f$ and $\operatorname{Im} f$ are both in \mathcal{A} . Therefore

$$\mathcal{A}_{\mathbb{R}} = \{\operatorname{Re} f, \operatorname{Im} f : f \in \mathcal{A}\}$$

is a real sub-algebra of $C(X, \mathbb{R})$ which separates points. Therefore either $\mathcal{A}_{\mathbb{R}} = C(X, \mathbb{R})$ or $\mathcal{A}_{\mathbb{R}} = \mathcal{I}_{x_0} \cap C(X, \mathbb{R})$ for some x_0 and hence $\mathcal{A} = C(X, \mathbb{C})$ or $\mathcal{I}_{x_0}^{\mathbb{C}}$ respectively. ■

As an easy application, Theorems 11.44 and 11.45 imply Corollaries 11.35 and 11.36 respectively.

Corollary 11.46. *Suppose that X is a compact subset of \mathbb{R}^n and μ is a finite measure on (X, \mathcal{B}_X) , then polynomials are dense in $L^p(X, \mu)$ for all $1 \leq p < \infty$.*

Proof. Consider X to be a metric space with usual metric induced from \mathbb{R}^n . Then X is a locally compact separable metric space and therefore $C_c(X, \mathbb{C}) = C(X, \mathbb{C})$ is dense in $L^p(\mu)$ for all $p \in [1, \infty)$. Since, by the dominated convergence theorem, uniform convergence implies $L^p(\mu)$ -convergence, it follows from the Stone-Weierstrass theorem that polynomials are also dense in $L^p(\mu)$. ■

Here are a couple of more applications.

Example 11.47. Let $f \in C([a, b])$ be a positive function which is injective. Then functions of the form $\sum_{k=1}^N a_k f^k$ with $a_k \in \mathbb{C}$ and $N \in \mathbb{N}$ are dense in $C([a, b])$. For example if $a = 1$ and $b = 2$, then one may take $f(x) = x^\alpha$ for any $\alpha \neq 0$, or $f(x) = e^x$, etc.

Exercise 11.3. Let (X, d) be a separable compact metric space. Show that $C(X)$ is also separable. **Hint:** Let $E \subset X$ be a countable dense set and then consider the algebra, $\mathcal{A} \subset C(X)$, generated by $\{d(x, \cdot)\}_{x \in E}$.

11.4. Locally Compact Version of Stone-Weierstrass Theorem.

Theorem 11.48. *Let X be non-compact locally compact Hausdorff space. If \mathcal{A} is a closed subalgebra of $C_0(X, \mathbb{R})$ which separates points. Then either $\mathcal{A} = C_0(X, \mathbb{R})$ or there exists $x_0 \in X$ such that $\mathcal{A} = \{f \in C_0(X, \mathbb{R}) : f(x_0) = 0\}$.*

Proof. There are two cases to consider.

Case 1. There is no point $x_0 \in X$ such that $\mathcal{A} \subset \{f \in C_0(X, \mathbb{R}) : f(x_0) = 0\}$. In this case let $X^* = X \cup \{\infty\}$ be the one point compactification of X . Because of Proposition 10.31 to each $f \in \mathcal{A}$ there exists a unique extension $\tilde{f} \in C(X^*, \mathbb{R})$ such that $f = \tilde{f}|_X$ and moreover this extension is given by $\tilde{f}(\infty) = 0$. Let $\tilde{\mathcal{A}} := \{\tilde{f} \in C(X^*, \mathbb{R}) : f \in \mathcal{A}\}$. Then $\tilde{\mathcal{A}}$ is a closed (you check) sub-algebra of $C(X^*, \mathbb{R})$ which separates points. An application of Theorem 11.44 implies $\tilde{\mathcal{A}} = \{F \in C(X^*, \mathbb{R}) : F(\infty) = 0\}$ and therefore by Proposition 10.31 $\mathcal{A} = \{F|_X : F \in \tilde{\mathcal{A}}\} = C_0(X, \mathbb{R})$.

Case 2. There exists $x_0 \in X$ such that $\mathcal{A} \subset \{f \in C_0(X, \mathbb{R}) : f(x_0) = 0\}$. In this case let $Y := X \setminus \{x_0\}$ and $\mathcal{A}_Y := \{f|_Y : f \in \mathcal{A}\}$. Since X is locally compact, one easily checks $\mathcal{A}_Y \subset C_0(Y, \mathbb{R})$ is a closed subalgebra which separates points. By Case 1. it follows that $\mathcal{A}_Y = C_0(Y, \mathbb{R})$. So if $f \in C_0(X, \mathbb{R})$ and $f(x_0) = 0$, $f|_Y \in C_0(Y, \mathbb{R}) = \mathcal{A}_Y$, i.e. there exists $g \in \mathcal{A}$ such that $g|_Y = f|_Y$. Since $g(x_0) = f(x_0) = 0$, it follows that $f = g \in \mathcal{A}$ and therefore $\mathcal{A} = \{f \in C_0(X, \mathbb{R}) : f(x_0) = 0\}$. ■

Example 11.49. Let $X = [0, \infty)$, $\lambda > 0$ be fixed, \mathcal{A} be the algebra generated by $t \rightarrow e^{-\lambda t}$. So the general element $f \in \mathcal{A}$ is of the form $f(t) = p(e^{-\lambda t})$, where $p(x)$

is a polynomial. Since $\mathcal{A} \subset C_0(X, \mathbb{R})$ separates points and $e^{-\lambda t} \in \mathcal{A}$ is pointwise positive, $\bar{\mathcal{A}} = C_0(X, \mathbb{R})$.

As an application of this example, we will show that the Laplace transform is injective.

Theorem 11.50. *For $f \in L^1([0, \infty), dx)$, the Laplace transform of f is defined by*

$$\mathcal{L}f(\lambda) \equiv \int_0^\infty e^{-\lambda x} f(x) dx \text{ for all } \lambda > 0.$$

If $\mathcal{L}f(\lambda) \equiv 0$ then $f(x) = 0$ for m -a.e. x .

Proof. Suppose that $f \in L^1([0, \infty), dx)$ such that $\mathcal{L}f(\lambda) \equiv 0$. Let $g \in C_0([0, \infty), \mathbb{R})$ and $\epsilon > 0$ be given. Choose $\{a_\lambda\}_{\lambda>0}$ such that $\#\{\lambda > 0 : a_\lambda \neq 0\} < \infty$ and

$$|g(x) - \sum_{\lambda>0} a_\lambda e^{-\lambda x}| < \epsilon \text{ for all } x \geq 0.$$

Then

$$\begin{aligned} \left| \int_0^\infty g(x) f(x) dx \right| &= \left| \int_0^\infty \left(g(x) - \sum_{\lambda>0} a_\lambda e^{-\lambda x} \right) f(x) dx \right| \\ &\leq \int_0^\infty \left| g(x) - \sum_{\lambda>0} a_\lambda e^{-\lambda x} \right| |f(x)| dx \leq \epsilon \|f\|_1. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, it follows that $\int_0^\infty g(x) f(x) dx = 0$ for all $g \in C_0([0, \infty), \mathbb{R})$. The proof is finished by an application of Lemma 11.7. ■

11.5. Dynkin's Multiplicative System Theorem. This section is devoted to an extension of Theorem 8.12 based on the Weierstrass approximation theorem. In this section X is a set.

Definition 11.51 (Multiplicative System). A collection of real valued functions Q on a set X is a **multiplicative system** provided $f \cdot g \in Q$ whenever $f, g \in Q$.

Theorem 11.52 (Dynkin's Multiplicative System Theorem). *Let \mathcal{H} be a linear subspace of $B(X, \mathbb{R})$ which contains the constant functions and is closed under bounded convergence. If $Q \subset \mathcal{H}$ is multiplicative system, then \mathcal{H} contains all bounded real valued $\sigma(Q)$ -measurable functions.*

Theorem 11.53 (Complex Multiplicative System Theorem). *Let \mathcal{H} be a complex linear subspace of $B(X, \mathbb{C})$ such that: $1 \in \mathcal{H}$, \mathcal{H} is closed under complex conjugation, and \mathcal{H} is closed under bounded convergence. If $Q \subset \mathcal{H}$ is multiplicative system which is closed under conjugation, then \mathcal{H} contains all bounded complex valued $\sigma(Q)$ -measurable functions.*

Proof. Let \mathbb{F} be \mathbb{R} or \mathbb{C} . Let \mathcal{C} be the family of all sets of the form:

$$(11.26) \quad B := \{x \in X : f_1(x) \in R_1, \dots, f_m(x) \in R_m\}$$

where $m = 1, 2, \dots$, and for $k = 1, 2, \dots, m$, $f_k \in Q$ and R_k is an open interval if $\mathbb{F} = \mathbb{R}$ or R_k is an open rectangle in \mathbb{C} if $\mathbb{F} = \mathbb{C}$. The family \mathcal{C} is easily seen to be a π -system such that $\sigma(Q) = \sigma(\mathcal{C})$. So By Theorem 8.12, to finish the proof it suffices to show $1_B \in \mathcal{H}$ for all $B \in \mathcal{C}$.

It is easy to construct, for each k , a uniformly bounded sequence of continuous functions $\{\phi_n^k\}_{n=1}^\infty$ on \mathbb{F} converging to the characteristic function 1_{R_k} . By Weierstrass' theorem, there exists polynomials $p_n^k(x)$ such that $|p_n^k(x) - \phi_n^k(x)| \leq 1/n$ for $|x| \leq \|\phi_k\|_\infty$ in the real case and polynomials $p_n^k(z, \bar{z})$ in z and \bar{z} such that $|p_n^k(z, \bar{z}) - \phi_n^k(z)| \leq 1/n$ for $|z| \leq \|\phi_k\|_\infty$ in the complex case. The functions

$$F_n := p_n^1(f_1)p_n^2(f_2) \cdots p_n^m(f_m) \quad (\text{real case})$$

$$F_n := p_n^1(f_1, \bar{f}_1)p_n^2(f_2, \bar{f}_2) \cdots p_n^m(f_m, \bar{f}_m) \quad (\text{complex case})$$

on X are uniformly bounded, belong to \mathcal{H} and converge pointwise to 1_B as $n \rightarrow \infty$, where B is the set in Eq. (11.26). Thus $1_B \in \mathcal{H}$ and the proof is complete. ■

Remark 11.54. Given any collection of bounded real valued functions \mathcal{F} on X , let $\mathcal{H}(\mathcal{F})$ be the subspace of $B(X, \mathbb{R})$ generated by \mathcal{F} , i.e. $\mathcal{H}(\mathcal{F})$ is the smallest subspace of $B(X, \mathbb{R})$ which is closed under bounded convergence and contains \mathcal{F} . With this notation, Theorem 11.52 may be stated as follows. If \mathcal{F} is a multiplicative system then $\mathcal{H}(\mathcal{F}) = B_{\sigma(\mathcal{F})}(X, \mathbb{R})$ – the space of bounded $\sigma(\mathcal{F})$ – measurable real valued functions on X .

11.6. Exercises.

Exercise 11.4. Let (X, τ) be a topological space, μ a measure on $\mathcal{B}_X = \sigma(\tau)$ and $f : X \rightarrow \mathbb{C}$ be a measurable function. Letting ν be the measure, $d\nu = |f| d\mu$, show $\text{supp}(\nu) = \text{supp}_\mu(f)$, where $\text{supp}_\mu(f)$ is defined in Definition 9.41).

Exercise 11.5. Let (X, τ) be a topological space, μ a measure on $\mathcal{B}_X = \sigma(\tau)$ such that $\text{supp}(\mu) = X$ (see Definition 9.41). Show $\text{supp}_\mu(f) = \text{supp}(f) = \overline{\{f \neq 0\}}$ for all $f \in C(X)$.

Exercise 11.6. Prove Proposition 11.24 by appealing to Corollary 7.43.

Exercise 11.7 (Integration by Parts). Suppose that $(x, y) \in \mathbb{R} \times \mathbb{R}^{n-1} \rightarrow f(x, y) \in \mathbb{C}$ and $(x, y) \in \mathbb{R} \times \mathbb{R}^{n-1} \rightarrow g(x, y) \in \mathbb{C}$ are measurable functions such that for each fixed $y \in \mathbb{R}^{n-1}$, $x \rightarrow f(x, y)$ and $x \rightarrow g(x, y)$ are continuously differentiable. Also assume $f \cdot g$, $\partial_x f \cdot g$ and $f \cdot \partial_x g$ are integrable relative to Lebesgue measure on $\mathbb{R} \times \mathbb{R}^{n-1}$, where $\partial_x f(x, y) := \frac{d}{dt} f(x+t, y)|_{t=0}$. Show

$$(11.27) \quad \int_{\mathbb{R} \times \mathbb{R}^{n-1}} \partial_x f(x, y) \cdot g(x, y) dx dy = - \int_{\mathbb{R} \times \mathbb{R}^{n-1}} f(x, y) \cdot \partial_x g(x, y) dx dy.$$

(Note: this result and Fubini's theorem proves Lemma 11.26.)

Hints: Let $\psi \in C_c^\infty(\mathbb{R})$ be a function which is 1 in a neighborhood of $0 \in \mathbb{R}$ and set $\psi_\epsilon(x) = \psi(\epsilon x)$. First verify Eq. (11.27) with $f(x, y)$ replaced by $\psi_\epsilon(x)f(x, y)$ by doing the x – integral first. Then use the dominated convergence theorem to prove Eq. (11.27) by passing to the limit, $\epsilon \downarrow 0$.

Exercise 11.8. Let $M < \infty$, show there are polynomials $p_n(t)$ such that

$$\lim_{n \rightarrow \infty} \sup_{|t| \leq M} ||t| - p_n(t)| = 0$$

as follows. Let $f(t) = \sqrt{1-t}$ for $|t| \leq 1$. By Taylor's theorem with integral remainder (see Eq. A.15 of Appendix A) or by analytic function theory, there are

constants²⁷ $\alpha_n > 0$ for $n \in \mathbb{N}$ such that $\sqrt{1-x} = 1 - \sum_{n=1}^{\infty} \alpha_n x^n$ for all $|x| < 1$. Use this to prove $\sum_{n=1}^{\infty} \alpha_n = 1$ and therefore $q_m(x) := 1 - \sum_{n=1}^m \alpha_n x^n$

$$\lim_{m \rightarrow \infty} \sup_{|x| \leq 1} |\sqrt{1-x} - q_m(x)| = 0.$$

Let $1-x = t^2/M^2$, i.e. $x = 1 - t^2/M^2$, then

$$\lim_{m \rightarrow \infty} \sup_{|t| \leq M} \left| \frac{|t|}{M} - q_m(1 - t^2/M^2) \right| = 0$$

so that $p_m(t) := Mq_m(1 - t^2/M^2)$ are the desired polynomials.

Exercise 11.9. Given a continuous function $f : \mathbb{R} \rightarrow \mathbb{C}$ which is 2π -periodic and $\epsilon > 0$. Show there exists a trigonometric polynomial, $p(\theta) = \sum_{n=-N}^n \alpha_n e^{in\theta}$, such that $|f(\theta) - P(\theta)| < \epsilon$ for all $\theta \in \mathbb{R}$. **Hint:** show that there exists a unique function $F \in C(S^1)$ such that $f(\theta) = F(e^{i\theta})$ for all $\theta \in \mathbb{R}$.

Remark 11.55. Exercise 11.9 generalizes to 2π -periodic functions on \mathbb{R}^d , i.e. functions such that $f(\theta + 2\pi e_i) = f(\theta)$ for all $i = 1, 2, \dots, d$ where $\{e_i\}_{i=1}^d$ is the standard basis for \mathbb{R}^d . A trigonometric polynomial $p(\theta)$ is a function of $\theta \in \mathbb{R}^d$ of the form

$$p(\theta) = \sum_{n \in \Gamma} \alpha_n e^{in \cdot \theta}$$

where Γ is a finite subset of \mathbb{Z}^d . The assertion is again that these trigonometric polynomials are dense in the 2π -periodic functions relative to the supremum norm.

Exercise 11.10. Let μ be a finite measure on $\mathcal{B}_{\mathbb{R}^d}$, then $\mathbb{D} := \text{span}\{e^{i\lambda \cdot x} : \lambda \in \mathbb{R}^d\}$ is a dense subspace of $L^p(\mu)$ for all $1 \leq p < \infty$. **Hints:** By Proposition 11.6, $C_c(\mathbb{R}^d)$ is a dense subspace of $L^p(\mu)$. For $f \in C_c(\mathbb{R}^d)$ and $N \in \mathbb{N}$, let

$$f_N(x) := \sum_{n \in \mathbb{Z}^d} f(x + 2\pi Nn).$$

Show $f_N \in BC(\mathbb{R}^d)$ and $x \rightarrow f_N(Nx)$ is 2π -periodic, so by Exercise 11.9, $x \rightarrow f_N(Nx)$ can be approximated uniformly by trigonometric polynomials. Use this fact to conclude that $f_N \in \mathbb{D}^{L^p(\mu)}$. After this show $f_N \rightarrow f$ in $L^p(\mu)$.

Exercise 11.11. Suppose that μ and ν are two finite measures on \mathbb{R}^d such that

$$(11.28) \quad \int_{\mathbb{R}^d} e^{i\lambda \cdot x} d\mu(x) = \int_{\mathbb{R}^d} e^{i\lambda \cdot x} d\nu(x)$$

for all $\lambda \in \mathbb{R}^d$. Show $\mu = \nu$.

Hint: Perhaps the easiest way to do this is to use Exercise 11.10 with the measure μ being replaced by $\mu + \nu$. Alternatively, use the method of proof of Exercise 11.9 to show Eq. (11.28) implies $\int_{\mathbb{R}^d} f d\mu(x) = \int_{\mathbb{R}^d} f d\nu(x)$ for all $f \in C_c(\mathbb{R}^d)$.

Exercise 11.12. Again let μ be a finite measure on $\mathcal{B}_{\mathbb{R}^d}$. Further assume that $C_M := \int_{\mathbb{R}^d} e^{M|x|} d\mu(x) < \infty$ for all $M \in (0, \infty)$. Let $\mathcal{P}(\mathbb{R}^d)$ be the space of polynomials, $\rho(x) = \sum_{|\alpha| \leq N} \rho_\alpha x^\alpha$ with $\rho_\alpha \in \mathbb{C}$, on \mathbb{R}^d . (Notice that $|\rho(x)|^p \leq C(\rho, p, M)e^{M|x|}$, so that $\mathcal{P}(\mathbb{R}^d) \subset L^p(\mu)$ for all $1 \leq p < \infty$.) Show $\mathcal{P}(\mathbb{R}^d)$ is dense in $L^p(\mu)$ for all $1 \leq p < \infty$. Here is a possible outline.

²⁷In fact $\alpha_n := \frac{(2n-3)!!}{2^n n!}$, but this is not needed.

Outline: For $\lambda \in \mathbb{R}^d$ and $n \in \mathbb{N}$ let $f_\lambda^n(x) = (\lambda \cdot x)^n / n!$

- (1) Use calculus to verify $\sup_{t \geq 0} t^\alpha e^{-Mt} = (\alpha/M)^\alpha e^{-\alpha}$ for all $\alpha \geq 0$ where $(0/M)^0 := 1$. Use this estimate along with the identity

$$|\lambda \cdot x|^{pn} \leq |\lambda|^{pn} |x|^{pn} = \left(|x|^{pn} e^{-M|x|}\right) |\lambda|^{pn} e^{M|x|}$$

to find an estimate on $\|f_\lambda^n\|_p$.

- (2) Use your estimate on $\|f_\lambda^n\|_p$ to show $\sum_{n=0}^\infty \|f_\lambda^n\|_p < \infty$ and conclude

$$\lim_{N \rightarrow \infty} \left\| e^{i\lambda \cdot (\cdot)} - \sum_{n=0}^N f_\lambda^n \right\|_p = 0.$$

- (3) Now finish by appealing to Exercise 11.10.

Exercise 11.13. Again let μ be a finite measure on $\mathcal{B}_{\mathbb{R}^d}$ but now assume there exists an $\epsilon > 0$ such that $C := \int_{\mathbb{R}^d} e^{\epsilon|x|} d\mu(x) < \infty$. Also let $q > 1$ and $h \in L^q(\mu)$ be a function such that $\int_{\mathbb{R}^d} h(x) x^\alpha d\mu(x) = 0$ for all $\alpha \in \mathbb{N}_0^d$. (As mentioned in Exercise 11.13, $\mathcal{P}(\mathbb{R}^d) \subset L^p(\mu)$ for all $1 \leq p < \infty$, so $x \rightarrow h(x)x^\alpha$ is in $L^1(\mu)$.) Show $h(x) = 0$ for μ -a.e. x using the following outline.

Outline: For $\lambda \in \mathbb{R}^d$ and $n \in \mathbb{N}$ let $f_n^\lambda(x) = (\lambda \cdot x)^n / n!$ and let $p = q/(q-1)$ be the conjugate exponent to q .

- (1) Use calculus to verify $\sup_{t \geq 0} t^\alpha e^{-\epsilon t} = (\alpha/\epsilon)^\alpha e^{-\alpha}$ for all $\alpha \geq 0$ where $(0/\epsilon)^0 := 1$. Use this estimate along with the identity

$$|\lambda \cdot x|^{pn} \leq |\lambda|^{pn} |x|^{pn} = \left(|x|^{pn} e^{-\epsilon|x|}\right) |\lambda|^{pn} e^{\epsilon|x|}$$

to find an estimate on $\|f_n^\lambda\|_p$.

- (2) Use your estimate on $\|f_n^\lambda\|_p$ to show there exists $\delta > 0$ such that $\sum_{n=0}^\infty \|f_n^\lambda\|_p < \infty$ when $|\lambda| \leq \delta$ and conclude for $|\lambda| \leq \delta$ that $e^{i\lambda \cdot x} = L^p(\mu)$ - $\sum_{n=0}^\infty f_n^\lambda(x)$. Conclude from this that

$$\int_{\mathbb{R}^d} h(x) e^{i\lambda \cdot x} d\mu(x) = 0 \text{ when } |\lambda| \leq \delta.$$

- (3) Let $\lambda \in \mathbb{R}^d$ ($|\lambda|$ not necessarily small) and set $g(t) := \int_{\mathbb{R}^d} e^{it\lambda \cdot x} h(x) d\mu(x)$ for $t \in \mathbb{R}$. Show $g \in C^\infty(\mathbb{R})$ and

$$g^{(n)}(t) = \int_{\mathbb{R}^d} (i\lambda \cdot x)^n e^{it\lambda \cdot x} h(x) d\mu(x) \text{ for all } n \in \mathbb{N}.$$

- (4) Let $T = \sup\{\tau \geq 0 : g|_{[0,\tau]} \equiv 0\}$. By Step 2., $T \geq \delta$. If $T < \infty$, then

$$0 = g^{(n)}(T) = \int_{\mathbb{R}^d} (i\lambda \cdot x)^n e^{iT\lambda \cdot x} h(x) d\mu(x) \text{ for all } n \in \mathbb{N}.$$

Use Step 3. with h replaced by $e^{iT\lambda \cdot x} h(x)$ to conclude

$$g(T+t) = \int_{\mathbb{R}^d} e^{i(T+t)\lambda \cdot x} h(x) d\mu(x) = 0 \text{ for all } t \leq \delta/|\lambda|.$$

This violates the definition of T and therefore $T = \infty$ and in particular we may take $T = 1$ to learn

$$\int_{\mathbb{R}^d} h(x) e^{i\lambda \cdot x} d\mu(x) = 0 \text{ for all } \lambda \in \mathbb{R}^d.$$

(5) Use Exercise 11.10 to conclude that

$$\int_{\mathbb{R}^d} h(x)g(x)d\mu(x) = 0$$

for all $g \in L^p(\mu)$. Now choose g judiciously to finish the proof.