9. Approximation Theorems and Convolutions

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

Notation 9.1. Let $\mathbb{S}_f(\mathcal{A}, \mu)$ denote those simple functions $\phi : X \to \mathbb{C}$ 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 9.2 (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 5.12. 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,

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

■

Theorem 9.3 (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*

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

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 $k \to \infty$. 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 6.12, $\mathcal{H}_k$ is the set of all bounded $\mathcal{M}$ – measurable functions on X.

For $f \in L^p(\mu)$, the dominated convergence theorem implies $1_{X_k \cap \{|f| \leq k\}} f \to f$ in $L^p(\mu)$ as $k \to \infty$. 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 6.42, for all $E \in \mathcal{M}$ such that $\mu(E) < \infty$, there exists $A \in \mathcal{A}$ such that $\mu(E \triangle A) < \epsilon$. Therefore

$$\int |1_E - 1_A|^p d\mu = \mu(E \triangle A) < \epsilon. \tag{9.1}$$

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

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

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

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

where $a_k, b_k \in \bar{\mathbb{R}}$. By Theorem 9.3, 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

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

Combining these two equations with

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

we learn that

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

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

Theorem 9.5 (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 $\mu : \mathcal{B}_X \to [0, \infty]$ 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 $\psi_{n,k} \to 1_{d_{X_k^c} > 0} = 1_{X_k}$ as $n \to \infty$, see Figure 22 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, $f : X \to \mathbb{R}$, such that $\psi_{n,k} f \in \overline{BC_f(X)}^{L^p(\mu)}$. It is

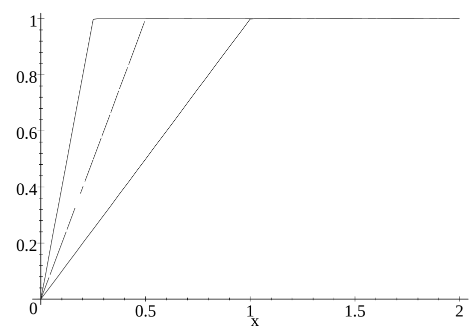

FIGURE 22. The plot of ϕ_n for $n = 1$, 2, and 4. Notice that $\phi_n \to 1_{(0,\infty)}$.

easily seen that $\mathcal{H}$ is a vector space closed under bounded convergence and this subspace contains $BC(X, \mathbb{R})$. By Corollary 6.13, $\mathcal{H}$ 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, $\psi_{n,k} f \to 1_{X_k} f$ in $L^p(\mu)$ as $n \to \infty$, therefore $1_{X_k} f \in \overline{BC_f(X)}^{L^p(\mu)}$. It now follows as in the first proof of Theorem 9.3 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 6.4 and 6.5, 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. (8.1), then $f \in BC_f(X)$ and since $|1_A - f| \le 1_{V \setminus F}$,

$$\int |1_A - f|^p \, d\mu \le \int 1_{V \setminus F} d\mu = \mu(V \setminus F) \le \epsilon \tag{9.2}$$

or equivalently

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

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 9.6. *Let (X, τ) be a second countable locally compact Hausdorff space, $\mathcal{B}_X = \sigma(\tau)$ be the Borel σ – algebra and $\mu : \mathcal{B}_X \to [0, \infty]$ 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 8.10 and set $X_k = K_k^o$. Using Item 3. of Lemma 8.17, there exists $\{\psi_{n,k}\}_{n=1}^{\infty} \subset C_c(X)$ such that $\mathrm{supp}(\psi_{n,k}) \subset X_k$ and $\lim_{n\to\infty} \psi_{n,k} = 1_{X_k}$. As in the first proof of Theorem 9.5, let $\mathcal{H}$ denote those bounded $\mathcal{B}_X$ – measurable functions, $f : X \to \mathbb{R}$, 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

8.18, $\mathcal{H}$ 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, $\psi_{n,k}f \to 1_{X_k}f$ in $L^p(\mu)$ as $k \to \infty$, therefore $1_{X_k}f \in \overline{C_c(X)}^{L^p(\mu)}$. It now follows as in the first proof of Theorem 9.3 that $\overline{C_c(X)}^{L^p(\mu)} = L^p(\mu)$.

Second Proof. Following the second proof of Theorem 9.5, let $A \in \mathcal{M}$ with $\mu(A) < \infty$. Since $\lim_{k\to\infty} ||1_{A\cap K_k^o} - 1_A||_p = 0$, it suffices to assume $A \subset K_k^o$ for some k. Given $\epsilon > 0$, by Item 2. of Lemma 8.17 and Exercises 6.4 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. (8.1) is now in $C_c(X)$. The remainder of the proof now follows as in the second proof of Theorem 9.5. ■

Lemma 9.7. *Let (X, τ) be a second countable locally compact Hausdorff space, $\mathcal{B}_X = \sigma(\tau)$ be the Borel σ – algebra and $\mu : \mathcal{B}_X \to [0, \infty]$ 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*

$$\int_X fhd\mu = 0 \text{ for all } f \in C_c(X) \tag{9.3}$$

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 9.6. Notice that

$$\int_X f \cdot \operatorname{sgn}(h) d\nu = \int_X fhd\mu = 0 \text{ for all } f \in C_c(X). \tag{9.4}$$

Let $\{K_k\}_{k=1}^{\infty}$ be a sequence of compact sets such that $K_k \uparrow X$ as in Lemma 8.10. Then $1_{K_k}\overline{\operatorname{sgn}(h)} \in L^1(\nu)$ and therefore there exists $f_m \in C_c(X)$ such that $f_m \to 1_{K_k}\overline{\operatorname{sgn}(h)}$ in $L^1(\nu)$. So by Eq. (9.4),

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

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

Second Proof. Let K_k be as above and use Lemma 8.15 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 8.18, $0 = \int_X \chi fhd\mu$ for all bounded measurable functions $f : X \to \mathbb{R}$ and then by linearity for all bounded measurable functions $f : X \to \mathbb{C}$. Taking $f = \overline{\operatorname{sgn}(h)}$ then implies

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

and hence by the monotone convergence theorem,

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

■

Corollary 9.8. *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 $p \in [1, \infty)$.*

9.1. Convolution and Young's Inequalities.

Definition 9.9. Let $f, g : \mathbb{R}^n \to \mathbb{C}$ 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 9.10. 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 9.11 (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

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

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 the Minkowski's inequality for integrals, Theorem 7.27.

Proposition 9.12. *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 $f : \mathbb{R}^n \to \mathbb{C}$, let $\tau_z f : \mathbb{R}^n \to \mathbb{C}$ be defined by $\tau_z f(x) = f(x-z)$.

Proposition 9.13. *Suppose that $p \in [1, \infty)$, then $\tau_z : L^p \to L^p$ is an isometric isomorphism and for $f \in L^p$, $z \in \mathbb{R}^n \to \tau_z f \in L^p$ is continuous.*

Proof. The assertion that $\tau_z : L^p \to L^p$ 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 $\tau_z f \to f$ in L^p as $z \to 0 \in \mathbb{R}^n$.

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

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

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

Lemma 9.14. *Suppose $f, g, h : \mathbb{R}^n \to \mathbb{C}$ 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. *Let $A = \overline{\mathrm{supp}(f) + \mathrm{supp}(g)}$, then for $x \notin A$ we have $f * g(x) = 0$.*

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 invariant under the transformation $y \to x - y$. Similar computations prove all of the remaining assertions of the first three items of the lemma.

For item 4., if $x \notin \mathrm{supp}(f) + \mathrm{supp}(g)$, then for all $y \in \mathbb{R}^n$, either $x - y \notin \mathrm{supp}(f)$ or $y \notin \mathrm{supp}(g)$ and hence $f(x - y)g(y) = 0$ for all y. Thus $f * g(x) = 0$. ■

Proposition 9.15. *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 Holders 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

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

it follows that $f * g$ is uniformly continuous. Finally if $p, q \in (1, \infty)$, we learn from Lemma 9.14 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,

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

showing, with the aid of Proposition 8.28, $f*g\in C_0(\mathbb{R}^n)$. ■

Theorem 9.16. *Let $p,q,r\in[1,\infty]$ satisfy*

$$\frac{1}{p}+\frac{1}{q}=1+\frac{1}{r}. \tag{9.5}$$

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

$$\|f*g\|_r\leq\|f\|_p\|g\|_q. \tag{9.6}$$

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

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 7.3,

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

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

$$\begin{aligned}\|f*g\|_r^r &\leq \int\left(\int |f(x-y)|^{(1-\alpha)r}|g(y)|^{(1-\beta)r}dy\right)dx\cdot\|f\|_{\alpha p_1}^{\alpha}\|g\|_{\beta p_2}^{\beta} \\ &= \|f\|_{(1-\alpha)r}^{(1-\alpha)r}\|g\|_{(1-\beta)r}^{(1-\beta)r}\|f\|_{\alpha p_1}^{\alpha r}\|g\|_{\beta p_2}^{\beta r}.\end{aligned} \tag{9.7}$$

Let us now suppose, $(1-\alpha)r=\alpha p_1$ and $(1-\beta)r=\beta p_2$, in which case Eq. (9.7) becomes,

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

which is Eq. (9.6) with

$$p:=(1-\alpha)r=\alpha p_1 \text{ and } q:=(1-\beta)r=\beta p_2. \tag{9.8}$$

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. (9.5), 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. (9.5), $\alpha=p(1-\frac{1}{q})\geq 0$ 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 9.17 (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 $f\in L^p$ then $\phi_t*f\to f$ in L^p as $t\downarrow 0$.*
2. *If $f\in C_b(\mathbb{R}^n)$ and f is uniformly continuous then $\|\phi_t*f-f\|_\infty\to 0$ as $t\downarrow 0$.*
3. *If $f\in L^\infty$ and f is continuous on $U\subset_o\mathbb{R}^n$ then $\phi_t*f\to af$ 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

$$\begin{aligned}\phi_t*f(x)-af(x)&=\int_{\mathbb{R}^n}\left[f(x-tz)-f(x)\right]\phi(z)dz\\ &=\int_{\mathbb{R}^n}\left[\tau_{tz}f(x)-f(x)\right]\phi(z)dz.\end{aligned}\tag{9.9}$$

Hence by Minkowski's inequality for integrals (Theorem 7.27), Proposition 9.12 and the dominated convergence theorem,

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

Item 2. is proved similarly. Indeed, form Eq. (9.9)

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

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

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

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

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

■

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

Exercise 9.1. Let

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

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

Lemma 9.18. *There exists $\phi\in C_c^\infty(\mathbb{R}^n,[0,\infty))$ such that $\phi(0)>0$, $\operatorname{supp}(\phi)\subset \bar{B}(0,1)$ 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 9.1. Then $h\in C_c^\infty(\mathbb{R},[0,1])$, $\operatorname{supp}(h)\subset[-1,1]$ 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 9.19. 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 $L^1_{loc}(\mu)$ consists of those measurable functions $f:X\to\mathbb{C}$ 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 9.4 below.

Proposition 9.20. *Suppose that $f\in L^1_{loc}(\mathbb{R}^n,m)$ 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 9.21 (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 $\operatorname{supp}(f)\subset U$ and $f=1$ on K.*

Proof. Let ϕ be as in Lemma 9.18, $\phi_t(x)=t^{-n}\phi(x/t)$ be as in Theorem 9.17, 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

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

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 $0\le f\le 1$. ■

Here is an application of this corollary whose proof is left to the reader, Exercise 9.5.

Lemma 9.22 (Integration by Parts)**.** *Suppose f and g are measurable functions on $\mathbb{R}^n$ such that $t\to f(x_1,\dots,x_{i-1},t,x_{i+1},\dots,x_n)$ and $t\to g(x_1,\dots,x_{i-1},t,x_{i+1},\dots,x_n)$ 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*

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

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

Lemma 9.23. *For $f \in L^1(\mathbb{R}^n, m)$ let*

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

be the Fourier transform of f. Then $\hat{f} \in C_0(\mathbb{R}^n)$ and $\left\|\hat{f}\right\|_u \le (2\pi)^{-n/2} \|f\|_1$. (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,

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

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

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 9.22 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 $(2\pi)^{n/2} \left|\hat{f}(\xi)\right| \le |\xi|^{-2k} \left\|\Delta^k f\right\|_1 \to 0$ as $|\xi| \to \infty$ 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 $\lim_{k\to\infty} \|f - f_k\|_1 = 0$, then $\lim_{k\to\infty} \left\|\hat{f} - \hat{f}_k\right\|_u = 0$. Hence $\hat{f} \in C_0(\mathbb{R}^n)$ by an application of Proposition 8.28. ■

Corollary 9.24. *Let $X \subset \mathbb{R}^n$ be an open set and μ be a Radon measure on $\mathcal{B}_X$. Then $C_c^\infty(X)$ is dense in $L^p(\mu)$ for all $1 \le p < \infty$.*

Proof. By Corollary 9.8 it suffices to show any $f \in C_c(X)$ may be approximated in $L^p(\mu)$ by $\psi \in C_c^\infty(X)$. To prove this let ϕ be as in Lemma 9.18, ϕ_t be as in Theorem 9.17 and set $\psi_t := \phi_t * (f 1_X)$. Then by Proposition 9.20 $\psi_t \in C^\infty(X)$ and by Lemma 9.14 there exists a compact set $K \subset X$ such that $\operatorname{supp}(\psi_t) \subset K$ for all t sufficiently small. By Theorem 9.17, $\psi_t \to f$ uniformly on X as $t \downarrow 0$ and hence by the dominated convergence theorem (the dominating function being $\|f\|_\infty 1_K$), $\psi_t \to f$ in $L^p(\mu)$ as $t \downarrow 0$. ■

Lemma 9.25. *Let $X \subset \mathbb{R}^n$ be an open set, μ be a Radon measure on $\mathcal{B}_X$ and $h \in L^1_{loc}(\mu)$. If*

$$\int_X f h d\mu = 0 \text{ for all } f \in C_c^\infty(X) \tag{9.10}$$

then $h(x) = 0$ for μ – a.e. x.

Proof. First Proof. Let $d\nu(x) = |h(x)|\, dx$, then ν is a Radon measure on X and hence $C_c^\infty(X)$ is dense in $L^1(\nu)$ by Corollary 9.24. Notice that

$$\int_X f \cdot \operatorname{sgn}(h) d\nu = \int_X f h d\mu = 0 \text{ for all } f \in C_c^\infty(X). \tag{9.11}$$

Let $\{K_k\}_{k=1}^{\infty}$ be a sequence of compact sets such that $K_k \uparrow X$ as in Lemma 8.10. Then $1_{K_k}\overline{\operatorname{sgn}(h)} \in L^1(\nu)$ and therefore there exists $f_m \in C_c^\infty(X)$ such that $f_m \to 1_{K_k}\overline{\operatorname{sgn}(h)}$ in $L^1(\nu)$. So by Eq. (9.11),

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

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

Second Proof. Approximating $f \in C_c(X)$ by $\psi_t \in C_c^\infty(X)$ as in the proof of Corollary 9.24. The dominated convergence theorem (with dominating function being $\|f\|_\infty 1_K |h|$) shows,

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

That is to say Eq. (9.10) holds for all $f \in C_c(X)$. Let K_k be as above and use Corollary 9.21 to find $\chi \in C_c^\infty(X, [0,1])$ such that $\chi = 1$ on K_k. Then

$$0 = \int_X \chi f h d\mu \text{ for all } f \in BC(X)$$

and a routine application of Corollary 6.13 gives $0 = \int_X \chi f h d\mu$ for all bounded measurable functions $f : X \to \mathbb{C}$. Taking $f = \overline{\operatorname{sgn}(h)}$ then implies

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

and hence by the monotone convergence theorem,

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

■

9.1.1. *Smooth Partitions of Unity.* We have the following smooth variants of Proposition 8.22, Theorem 8.24 and Corollary 8.25. The proofs of these results are the same as their continuous counterparts. One simply uses the smooth version of Urysohn's Lemma of Corollary 9.21 in place of Lemma 8.15.

Proposition 9.26 (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 $j = 1, 2, \ldots, n$.*

Theorem 9.27 (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* $\operatorname{supp}(h_i)$ *is compact for all i.*

Corollary 9.28. *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* $\operatorname{supp}(h_\alpha) \subset U_\alpha$ *for all $\alpha \in A$. Moreover if $\bar{U}_\alpha$ is compact for each $\alpha \in A$ we may choose h_α so that $h_\alpha \prec U_\alpha$.*

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

Notation 9.29. 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 $p : \mathbb{R}^d \to \mathbb{C}$ of the form

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

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 $p(z) = \sum_{\alpha:|\alpha|\le N} p_\alpha z^\alpha$ where $z^\alpha = \prod_{i=1}^d z_i^{\alpha_i}$.

Remark 9.30. The mapping $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d \to z = x+iy \in \mathbb{C}^d$ 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

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

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 9.31 (Weierstrass Approximation Theorem)**.** *Let $a, b \in \mathbb{R}^d$ with $a \le b$ (i.e. $a_i \le b_i$ for $i = 1, 2, \dots, d$) and set $[a, b] := [a_1, b_1] \times \cdots \times [a_d, b_d]$. Then for $f \in C([a, b], \mathbb{C})$ there exists polynomials p_n on $\mathbb{R}^d$ such that $p_n \to f$ 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 base on using a certain sequence of approximate δ – functions.

Corollary 9.32. *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 $p_n \to f$ uniformly on K.*

Proof. Choose $a, b \in \mathbb{R}^d$ such that $a \le b$ and $K \subset (a, b) := (a_1, b_1) \times \cdots \times (a_d, b_d)$. Let $\tilde{f} : K \cup (a, b)^c \to \mathbb{C}$ 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 8.2 or 8.16 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 9.31 to $F|_{[a,b]}$ to find polynomials p_n on $\mathbb{R}^d$ such that $p_n \to F$ uniformly on $[a, b]$. Clearly we also have $p_n \to f$ uniformly on K. ∎

Corollary 9.33 (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 $\sup_{z\in K} |p_n(z, \bar{z}) - f(z)| \to 0$ as $n \to \infty$.*

Proof. This is an immediate consequence of Remark 9.30 and Corollary 9.32. ∎

Example 9.34. 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 $C(S^1)$.[19] 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 $K = (S^1)^d \subset \mathbb{C}^d$.

[19]Note 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.

9.2.1. *First proof of the Weierstrass Approximation Theorem 9.31.* **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 $x\in[\mathbf{0},\mathbf{1}]$, it suffices to prove the theorem for $f\in C([\mathbf{0},\mathbf{1}])$.

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

$$\int_{\{0,1\}} y d\nu_x(y)=0\cdot(1-x)+1\cdot x=x \text{ and} \tag{9.12}$$

$$\int_{\{0,1\}}(y-x)^2 d\nu_x(y)=x^2(1-x)+(1-x)^2\cdot x=x(1-x). \tag{9.13}$$

For $x\in[\mathbf{0},\mathbf{1}]$ 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

$$\mu_x(\{\epsilon\})=\prod_{i=1}^d(1-x_i)^{1-\epsilon_i}x_i^{\epsilon_i} \tag{9.14}$$

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 $x\in[\mathbf{0},\mathbf{1}]$, 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

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

so $S_n:\Omega^n\to\mathbb{R}^d$. The reader is asked to verify (Exercise 9.2) that

$$\int_{\Omega^n} S_n d\mu_x^n=\left(\int_{\Omega^n}S_n^1 d\mu_x^n,\dots,\int_{\Omega^n}S_n^d d\mu_x^n\right)=(x_1,\dots,x_d)=x \tag{9.15}$$

and

$$\int_{\Omega^n}|S_n-x|^2 d\mu_x^n=\frac{1}{n}\sum_{i=1}^d x_i(1-x_i)\le\frac{d}{n}. \tag{9.16}$$

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

$$p_n(x):=\int_{\Omega^n}f(S_n)d\mu_x^n \tag{9.17}$$

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

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

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

By uniform continuity of f on $[\mathbf{0},\mathbf{1}]$, $\lim_{\epsilon\downarrow0}\delta_\epsilon=0$. Using these definitions and the fact that $\mu_x^n(\Omega^n)=1$,

$$\begin{aligned}|f(x)-p_n(x)|&=\left|\int_{\Omega^n}(f(x)-f(S_n))d\mu_x^n\right|\le\int_{\Omega^n}|f(x)-f(S_n)|d\mu_x^n\\&\le\int_{\{|S_n-x|>\epsilon\}}|f(x)-f(S_n)|d\mu_x^n+\int_{\{|S_n-x|\le\epsilon\}}|f(x)-f(S_n)|d\mu_x^n\\&\le 2M\mu_x^n(|S_n-x|>\epsilon)+\delta_\epsilon.\end{aligned} \tag{9.18}$$

By Chebyshev's inequality,

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

and therefore, Eq. (9.18) yields the estimate

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

and hence

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

This completes the proof since, using Eq. (9.14),

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

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

Exercise 9.2. Verify Eqs. (9.15) and (9.16). This is most easily done using Eqs. (9.12) and (9.13) and Fubini's theorem repeatedly. (Of course Fubini's theorem here is over kill since these are only finite sums after all. Nevertheless it is convenient to use this formulation.)

9.2.2. *Second proof of the Weierstrass Approximation Theorem 9.31.* For the second proof we will first need two lemmas.

Lemma 9.35 (Approximate δ – sequences)**.** *Suppose that $\{Q_n\}_{n=1}^\infty$ is a sequence of positive functions on $\mathbb{R}^d$ such that*

$$\int_{\mathbb{R}^d} Q_n(x)\, dx = 1 \text{ and} \tag{9.19}$$

$$\lim_{n\to\infty}\int_{|x|\geq\epsilon} Q_n(x)dx = 0 \text{ for all } \epsilon > 0. \tag{9.20}$$

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. (9.19),

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

Let $M = \sup\left\{|f(x)| : x \in \mathbb{R}^d\right\}$ and $\epsilon > 0$, then by and Eq. (9.19)

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

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. (9.20),

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

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

$$q_n(x) \equiv \frac{1}{c_n}(1-x^2)^n 1_{|x|\le 1} \text{ where } c_n := \int_{-1}^{1}(1-x^2)^n dx. \tag{9.21}$$

Figure 23 displays the key features of the functions q_n.

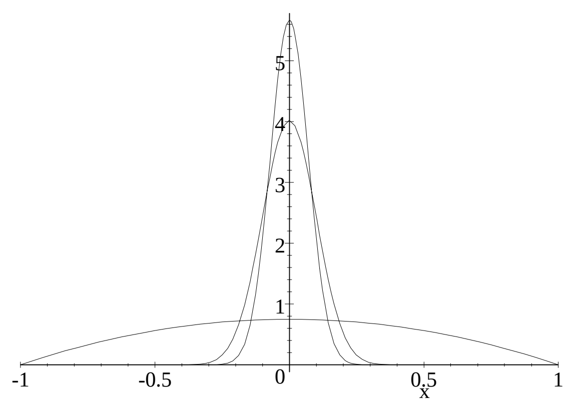

FIGURE 23. A plot of q_1, q_{50}, and q_{100}. The most peaked curve is q_{100} and the least is q_1. The total area under each of these curves is one.

Define

$$Q_n : \mathbb{R}^n \to [0,\infty) \text{ by } Q_n(x) = q_n(x_1)\dots q_n(x_d). \tag{9.22}$$

Lemma 9.36. *The sequence $\{Q_n\}_{n=1}^{\infty}$ is an approximate δ – sequence, i.e. they satisfy Eqs. (9.19) and (9.20).*

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. (9.20). With this norm

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

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

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

Since

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

the proof is complete. ■

We will now prove Corollary 9.32 which clearly implies Theorem 9.31.

Proof. Proof of Corollary 9.32. As in the beginning of the proof already given for Corollary 9.32, 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 $G(x) = F(a_1 + x_1(b_1 - a_1), \ldots, a_d + x_d(b_d - a_d))$ for $x \in \mathbb{R}^n$ we may further assume $K = [\mathbf{0}, \mathbf{1}]$.

Let $Q_n(x)$ be defined as in Eq. (9.22). Then by Lemma 9.36 and 9.35, $p_n(x) := (Q_n * F)(x) \to F(x)$ uniformly for $x \in [\mathbf{0}, \mathbf{1}]$ as $n \to \infty$. So to finish the proof it only remains to show $p_n(x)$ is a polynomial when $x \in [\mathbf{0}, \mathbf{1}]$. For $x \in [\mathbf{0}, \mathbf{1}]$,

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

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. ■

9.3. **Stone-Weierstrass Theorem.** We now wish to generalize Theorem 9.31 to more general topological spaces. We will first need some definitions.

Definition 9.37. 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 $f \in \mathcal{A}$.[20]

Remark 9.38. 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 9.39. *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.*

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

Proof. Let $f \in \mathcal{A}$ and let $M = \sup_{x \in X} |f(x)|$. Using Theorem 9.31 or Exercise 9.6, there are polynomials $p_n(t)$ such that

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

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

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

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

Lemma 9.40. *Let $\mathcal{A} \subset C(X, \mathbb{R})$ be an algebra which separates points and $x, y \in X$ be distinct points such that*

$$\exists\ f, g \in \mathcal{A} \ \ni\ f(x) \ne 0 \text{ and } g(y) \ne 0. \tag{9.23}$$

Then

$$V := \{(f(x), f(y)) : f \in \mathcal{A}\} = \mathbb{R}^2. \tag{9.24}$$

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 $a \ne 0$ and $b \ne 0$ by the assumption in Eq. (9.23). 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 the 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 9.41 (Stone-Weierstrass Theorem)**.** *Suppose 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*

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

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 $f(x) \ne 0$.*[21]
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 9.40, for all $x, y \in X$ such that $x \ne y$

[21]If $\mathcal{A}$ contains the constant function 1, then this hypothesis holds.

there exists $g_{xy} \in \mathcal{A}$ such that $f = g_{xy}$ on $\{x, y\}$.[22] The basic idea of the proof is contained in the following identity,

$$f(z) = \inf_{x \in X} \sup_{y \in X} g_{xy}(z) \text{ for all } z \in X. \tag{9.25}$$

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. (9.25) becoming only an approximate identity.

Claim 1. *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 24 below.

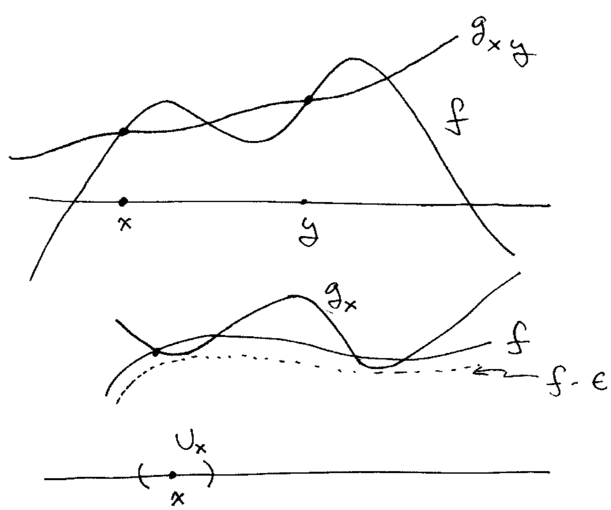

FIGURE 24. Constructing the funtions g_x.

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

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

$X = \bigcup_{x\in\Gamma} U_x$ 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 9.42 (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

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

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

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

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 9.41 and 9.42 imply Corollaries 9.32 and 9.33 respectively.

Corollary 9.43. *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 $1 \le p < \infty$.*

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 9.44. 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 9.3. 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 $\{d(x,\cdot)\}_{x\in E}$.

9.4. **Locally Compact Version of Stone-Weierstrass Theorem.**

Theorem 9.45. *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 8.29 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 $\widetilde{\mathcal{A}} := \{\tilde{f} \in C(X^*, \mathbb{R}) : f \in \mathcal{A}\}$. Then $\widetilde{\mathcal{A}}$ is a closed (you check) sub-algebra of $C(X^*, \mathbb{R})$ which separates points. An application of Theorem 9.41 implies $\widetilde{\mathcal{A}} = \{F \in C(X^*, \mathbb{R}) \ni F(\infty) = 0\}$ and therefore by Proposition 8.29 $\mathcal{A} = \{F|_X : F \in \widetilde{\mathcal{A}}\} = C_0(X, \mathbb{R})$.

Case 2. There exists $x_0 \in X$ such $\mathcal{A} \subset \{f \in C_0(X, \mathbb{R}) : f(x_0) = 0\}$. 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 9.46. Let $X = [0, \infty)$, $\lambda > 0$ be fixed, $\mathcal{A}$ be the algebra generated by $t \to e^{-\lambda t}$. 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 9.47. *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 9.7. ■

9.5. **Dynkin's Multiplicative System Theorem.** This section is devoted to an extension of Theorem 6.12 based on the Weierstrass approximation theorem. In this section X is a set.

Definition 9.48 (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 9.49 (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 9.50 (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:

$$B := \{x \in X : f_1(x) \in R_1, \ldots, f_m(x) \in R_m\} \tag{9.26}$$

where $m = 1, 2, \ldots$, and for $k = 1, 2, \ldots, m$, $f_k \in Q$ 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 6.12, 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 $p_m^k(x)$ such that $\left|p_n^k(x) - \phi_n^k(x)\right| \le 1/n$ for $|x| \le \|\phi_k\|_\infty$ in the real case and polynomials $p_m^k(z, \bar{z})$ in z and $\bar{z}$ such that $\left|p_n^k(z, \bar{z}) - \phi_n^k(z)\right| \le 1/n$ for $|z| \le \|\phi_k\|_\infty$ in the complex case. The functions

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

on X are uniformly bounded, belong to $\mathcal{H}$ and converge pointwise to 1_B as $n \to \infty$, where B is the set in Eq. (9.26). Thus $1_B \in \mathcal{H}$ and the proof is complete. ■

Remark 9.51*.* 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 9.49 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.

9.6. **Exercises.**

Exercise 9.4. Prove Proposition 9.20 by appealing to Corollary 5.43.

Exercise 9.5 (Integration by Parts)**.** Suppose that $(x, y) \in \mathbb{R} \times \mathbb{R}^{n-1} \to f(x, y) \in \mathbb{C}$ and $(x, y) \in \mathbb{R} \times \mathbb{R}^{n-1} \to g(x, y) \in \mathbb{C}$ are measurable functions such that for each fixed $y \in \mathbb{R}^{n-1}$, $x \to f(x, y)$ and $x \to g(x, y)$ 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

$$\int_{\mathbb{R} \times \mathbb{R}^{n-1}} \partial_x f(x, y) \cdot g(x, y) dx dy = -\int_{\mathbb{R} \times \mathbb{R}^{n-1}} f(x, y) \cdot \partial_x g(x, y) dx dy. \tag{9.27}$$

(Note: this result and Fubini's theorem proves Lemma 9.22.)

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. (9.27) 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. (9.27) by passing to the limit, $\epsilon \downarrow 0$.

Exercise 9.6. Let $M < \infty$, show there are polynomials $p_n(t)$ such that

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

as follows. Let $f(t) = \sqrt{1-t}$ for $|t| \le 1$. By Taylor's theorem with integral remainder (see Eq. A.15 of Appendix A) or by analytic function theory, there are constants[23] $\alpha_n > 0$ for $n \in \mathbb{N}$ such that

$$\sqrt{1-x} = 1 - \sum_{n=1}^\infty \alpha_n x^n \text{ for all } |x| < 1.$$

Use this to prove $\sum_{n=1}^\infty \alpha_n = 1$ and therefore $q_m(x) := 1 - \sum_{n=1}^m \alpha_n x^n$

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

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

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

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

Exercise 9.7. Given a continuous function $f : \mathbb{R} \to \mathbb{C}$ 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 9.52. Exercise 9.7 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 9.8. Let μ be a finite measure on $\mathcal{B}_{\mathbb{R}^d}$, then $\mathbb{D} := \operatorname{span}\{e^{i\lambda\cdot x} : \lambda \in \mathbb{R}^d\}$ is a dense subspace of $L^p(\mu)$ for all $1 \le p < \infty$. **Hints:** By Corollary 9.8, $C_c(\mathbb{R}^d)$ is a dense subspace of $L^p(\mu)$. For $f \in C_c(\mathbb{R}^d)$ and $N \in \mathbb{N}$, let

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

Show $f_N \in BC(\mathbb{R}^d)$ and, by Exercise 9.7, $f_N(x/N)$ can be approximated uniformly by trigonometric polynomials. Use this fact to conclude that $f_N \in \bar{\mathbb{D}}^{L^p(\mu)}$ for all N sufficiently large. After this show $f_N \to f$ in $L^p(\mu)$.

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

Exercise 9.9. Suppose that μ and ν are two finite measures on $\mathbb{R}^d$ such that

$$\int_{\mathbb{R}^d} e^{i\lambda\cdot x} d\mu(x) = \int_{\mathbb{R}^d} e^{i\lambda\cdot x} d\nu(x) \tag{9.28}$$

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

Hint: Perhaps the easiest way to do this is to use Exercise 9.8 with the measure μ being replaced by $\mu + \nu$. Alternatively, use the method of proof of Exercise 9.7 to show Eq. (9.28) 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 9.10. Again let μ be a finite measure on $\mathcal{B}_{\mathbb{R}^d}$. Further assume there exists an $\epsilon > 0$ such that $C := \int_{\mathbb{R}^d} e^{\epsilon|x|} d\mu(x) < \infty$. Show the space $\mathcal{P}(\mathbb{R}^d)$ of polynomials on $\mathbb{R}^d$ are dense in $L^p(\mu)$ for all $1 \le p < \infty$. Here is a possible outline.

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

1. Use calculus to verify $\sup_{t\ge 0} t^\alpha e^{-\epsilon t} = (\alpha/\epsilon)^\alpha e^{-\alpha}$ for all $\alpha \ge 0$ where $(0/\epsilon)^0 := 1$. Use this estimate along with the identity
$$|\lambda \cdot x|^{np} \le |\lambda|^{np} |x|^{pn} = \left(|x|^{pn} e^{-\epsilon|x|}\right) |\lambda|^{np} e^{\epsilon|x|}$$
to find an estimate on $\|f_n\|_p$.
2. Use this estimate to show there exists $\delta > 0$ such that $\sum_{n=0}^\infty \|f_n\|_p < \infty$ when $|\lambda| \le \delta$ and conclude for $|\lambda| \le \delta$ that $e^{i\lambda\cdot x} = L^p(\mu)\text{–}\sum_{n=0}^\infty f_n(x)$. From this it follows that $\int_{\mathbb{R}^d} e^{i\lambda\cdot x} d\mu(x) = 0$ when $|\lambda| \le \delta$.
3. Let $\lambda \in \mathbb{R}^d$ ($|\lambda|$ not necessarily small) and set $g(t) := \int_{\mathbb{R}^d} e^{it\lambda\cdot x} d\mu(x)$ for $t \in \mathbb{R}$. Show $g \in C^\infty(\mathbb{R})$ and
$$g^{(n)}(t) = \int_{\mathbb{R}^d} (i\lambda \cdot x)^n e^{it\lambda\cdot x} d\mu(x) \text{ for all } n \in \mathbb{N}.$$
4. Let $T = \sup\{\tau \ge 0 : g|_{[0,\tau]} \equiv 0\}$. By Step 2., $T \ge \delta$. If $T < \infty$, use Step 3. to conclude
$$\int_{\mathbb{R}^d} (i\lambda \cdot x)^n e^{iT\lambda\cdot x} d\mu(x) = 0 \text{ for all } n \in \mathbb{N}.$$
Then use Step 2. again to conclude
$$\int_{\mathbb{R}^d} e^{i(T+t)\lambda\cdot x} d\mu(x) = 0 \text{ for all } t \le \delta / |\lambda|$$
which violates the definition of T and therefore $T = \infty$.
5. Now finish by appealing to Exercise 9.8.