We now study the self-adjoint subalgebras of $C(X)$ for X a compact Hausdorff space. We begin with the generalization due to Stone of the classical theorem of Weierstrass on the density of polynomials. A subset $\mathfrak{U}$ of $C(X)$ is said to be self-adjoint if f in $\mathfrak{U}$ implies $\overline{f}$ is in $\mathfrak{U}$.

2.40 Theorem. (Stone-Weierstrass) Let X be a compact Hausdorff space. If $\mathfrak{U}$ is a closed self-adjoint subalgebra of $C(X)$ which separates the points of X and contains the constant function 1, then $\mathfrak{U} = C(X)$.

Proof If $\mathfrak{U}_r$ denotes the set of real functions in $\mathfrak{U}$, then $\mathfrak{U}_r$ is a closed subalgebra of the real algebra $C_r(X)$ of continuous functions on X which separates points and contains the function 1. Moreover, proof of the theorem reduces to showing that $\mathfrak{U}_r = C_r(X)$.

We begin by showing that f in $\mathfrak{U}_r$, implies that $|f|$ is in $\mathfrak{U}_r$. Recall that the binomial series for the function $\varphi(t) = (1-t)^{1/2}$ is $\sum_{n=0}^{\infty} \alpha_n t^n$, where $\alpha_n = (-1)^n \binom{1/2}{n}$. It is an easy consequence of the comparison theorem that the sequence $\{\sum_{n=0}^{N} \alpha_n t^n\}_{N=1}^{\infty}$ converges uniformly to φ on the closed interval $[0, 1-\delta]$ for $\delta > 0$. (The sequence actually converges uniformly to φ on $[-1, 1]$.) Let f be in $\mathfrak{U}_r$ such that $\|f\|_\infty \le 1$ and set $g_\delta = \delta + (1-\delta)f^2$ for δ in $(0, 1]$; then $0 \le 1 - g_\delta \le 1 - \delta$. For fixed $\delta > 0$, set $h_N = \sum_{n=0}^{N} \alpha_n (1 - g_\delta)^n$. Then h_N is in $\mathfrak{U}_r$ and

$$\left\| h_N - (g_\delta)^{1/2} \right\|_\infty = \sup_{x \in X} \left| \sum_{n=0}^{N} \alpha_n \big(1 - g_\delta(x)\big)^n - \varphi\big(1 - g_\delta(x)\big) \right|$$

$$\le \sup_{t \in [0, 1-\delta]} \left| \sum_{n=0}^{N} \alpha_n t^n - \varphi(t) \right| .$$

Therefore, $\lim_{N \to \infty} \left\| h_N - (g_\delta)^{1/2} \right\|_\infty = 0$, implying that $(g_\delta)^{1/2}$ is in $\mathfrak{U}_r$. Now since the square root function is uniformly continuous on [0,1], we have $\lim_{\delta \to 0} \left\| |f| - (g_\delta)^{1/2} \right\|_\infty = 0$, and thus $|f|$ is in $\mathfrak{U}_r$.

We next show that $\mathfrak{U}_r$ is a lattice, that is, for f and g in $\mathfrak{U}_r$ the functions $f \vee g$ and $f \wedge g$ are in $\mathfrak{U}_r$, where $(f \vee g)(x) = \max\{f(x), g(x)\}$, and $(f \wedge g)(x) = \min\{f(x), g(x)\}$. This follows from the identities

$$f \vee g = \tfrac{1}{2}\{f + g + |f - g|\}, \qquad \text{and} \qquad f \wedge g = \tfrac{1}{2}\{f + g - |f - g|\}$$

which can be verified pointwise.

Further, if x and y are distinct points in X and a and b arbitrary real numbers, and f is a function in $\mathfrak{U}_r$ such that $f(x) \ne f(y)$, then the function g defined by

$$g(z) = a + (b - a)\frac{f(z) - f(x)}{f(y) - f(x)}$$

is in $\mathfrak{U}_r$ and has the property that $g(x) = a$ and $g(y) = b$. Thus there exist functions in $\mathfrak{U}_r$ taking prescribed values at two points.

We now complete the proof. Take f in $C_r(X)$ and $\varepsilon > 0$. Fix x_0 in X. For each x in X, we can find a g_x in $\mathfrak{U}_r$ such that $g_x(x_0) = f(x_0)$ and $g_x(x) = f(x)$. Since f and g are continuous, there exists an open set U_x of x such that $g_x(y) \leq f(y) + \varepsilon$ for all y in U_x. The open sets $\{U_x\}_{x\in X}$ cover X and hence by compactness, there is a finite subcover $U_{x_1}, U_{x_2}, \ldots, U_{x_n}$. Let $h_{x_0} = g_{x_1} \wedge g_{x_2} \wedge \cdots \wedge g_{x_n}$. Then h_{x_0} is in $\mathfrak{U}_r$, $h_{x_0}(x_0) = f(x_0)$, and $h_{x_0}(y) \leq f(y) + \varepsilon$ for y in X.

Thus for each x_0 in X there exists h_{x_0} in $\mathfrak{U}_r$ such that $h_{x_0}(x_0) = f(x_0)$ and $h_{x_0}(y) \leq f(y) + \varepsilon$ for y in X. Since h_{x_0} and f are continuous, there exists an open set V_{x_0} of x_0 such that $h_{x_0}(y) \geq f(y) - \varepsilon$ for y in V_{x_0}. Again, the family $\{V_{x_0}\}_{x_0\in X}$ covers X, and hence there exists a finite subcover $V_{x_1}, V_{x_2}, \ldots, V_{x_m}$. If we set $k = h_{x_1}, \vee h_{x_2} \vee \cdots \vee h_{x_m}$, then k is in $\mathfrak{U}_r$ and $f(y) - \varepsilon \leq k(y) \leq f(y) + \varepsilon$ for y in X. Therefore, $\|f - k\|_\infty \leq \varepsilon$ and the proof is complete. ∎

13.3 THE STONE-WEIERSTRASS THEOREM

Theorem 4. *Let S be a compact Hausdorff space, $C(S)$ the set of all real-valued continuous functions on S. Let E be a subalgebra of $C(S)$, that is,*

(i) E is a linear subspace of $C(S)$.
(ii) The product of two functions in E belongs to E.

In addition we impose the following conditions on E:

(iii) E separates points of S, that is, given any pair of points p and q, $p \neq q$, there is a function f in E such that $f(p) \neq f(q)$.
(iv) All constant functions belong to E.

Conclusion: E is dense in $C(S)$ in the maximum norm.

The classical Weierstrass theorem is a special case of this proposition, with S an interval of the x axis, and E the set of all polynomials in x. We present Louis de Branges's elegant proof, based on the Krein-Milman theorem, of Stone's generalization of the Weierstrass theorem.

Proof. According to the spanning criterion, theorem 8 of chapter 8, E is dense in $C(S)$ if the only bounded linear functional ℓ on $C(S)$ that is zero on E is the zero functional. According to the Riesz-Kakutani representation theorem, theorem 14 of chapter 8, the bounded linear functionals on $C(S)$ are of the form

$$\ell(f) = \int_S f\,d\nu,$$

ν a signed measure of finite total variation $\|\nu\| = \int |d\nu|$. So what we have to show is that if $\int_S f\,d\nu = 0$ for all f in E, $\nu = 0$.

Suppose not; denote by U the set of signed measures of finite total mass is ≤ 1 that annihilate all functions in E. This is a convex set, and according to Alaoglu's theorem, theorem 3 in chapter 12, compact in the weak* topology. So according to the Krein-Milman theorem, if U contained a nonzero measure, it would contain a nonzero extreme point; call it μ. Since μ is extreme, $\|\mu\| = 1$. Since E is an algebra, if f and g belong to E, so does gf. Since μ annihilates every function in E,

$$\int (fg)d\mu = 0.$$

It follows that the measure $g\,d\mu$ also annihilates every function in E.

Let g be a function in E whose values lie between 0 and 1:

$$0 < g(p) < 1 \qquad \text{for all } p \text{ in } S.$$

Denote

$$a = \|g\mu\| = \int g|d\mu|, \quad b = \|(1-g)\mu\| = \int (1-g)|\,d\mu.$$

Clearly a and b are positive. Add them:

$$a + b = \int |d\mu| = 1.$$

The identity

$$\mu = a\frac{g\mu}{a} + b\frac{(1-g)\mu}{b}$$

represents μ as a nontrivial convex combination of $g\mu/a$ and $(1-g)\mu/b$, both points in U. Since μ is an extreme point, μ must be equal to $g\mu/a$.

Define the *support* of the measure μ to be the set of points p that have the property that $\int_N |d\mu| > 0$ for any open set N containing p. If $\mu = g\mu/a$, it follows that g has the same value at all points of the support of μ.

We claim that the support of μ consists of a single point. To see this, suppose that both p and q, $p \neq q$, belong to the support μ. Since the functions in E separate points of S, there is a function h in E, $h(p) \neq h(q)$. Adding a large enough constant to h and dividing it by another large constant, we obtain a function g whose values lie between 0 and 1, and $g(p) \neq g(q)$. This contradicts our previous conclusion.

A measure μ whose support consists of a single point p, and $\|\mu\| = 1$, is a unit point mass at p. Therefore

$$\int f \, d\mu = f(p) \text{ or } -f(p).$$

Since, by hypothesis, the constant 1 belong to E, $\int f \, d\mu \neq 0$ for $f \equiv 1$ in E, a contradiction. □