The Stone-Weierstrass theorem

Throughout this section, X denotes a compact Hausdorff space, for example a compact metric space. In what follows, we take $C(X)$ to denote the algebra of real-valued continuous functions on X. We return to the complex valued case at the end.

Definition 12.1. We say a set of functions $\mathcal{A} \subset C(X)$ *separates points* if for every $x, y \in X$, there is a function $f \in \mathcal{A}$ so $f(x) \neq f(y)$.

Theorem 12.2 (Stone-Weierstrass (proved by Stone, published in 1948))**.**
Let $\mathcal{A}$ be a subalgebra of $C(X)$ which

- contains the constants, and
- separates points.

Then $\mathcal{A}$ is uniformly dense in $C(X)$.

Corollary 12.3 (Weierstrass approximation (1895))**.** Polynomials are uniformly dense in $C([a,b])$.

I'll give a proof here adapted from §4.3 of Pedersen's book *Analysis Now.*

Definition 12.4. Let $\mathcal{A}$ be a vector subspace of $C(X)$. If $\mathcal{A}$ contains $\max\{f,g\}$ and $\min\{f,g\}$ whenever $f, g \in \mathcal{A}$, then we call $\mathcal{A}$ a *function lattice.*

Definition 12.5. A set of functions $\mathcal{A} \subset C(X)$ *separates points strongly* if for $x, y \in X$ and $a, b \in \mathbb{R}$, there is a function $f \in \mathcal{A}$ so $f(x) = a$ and $f(y) = b$.

Lemma 12.6. If a subspace $\mathcal{A} \subset C(X)$ separates points and contains the constants, it separates points strongly.

Lemma 12.7. If $\mathcal{A}$ is a subalgebra of $C(X)$, then for $f, g \in \mathcal{A}$, $\max\{f,g\}$ and $\min\{f,g\}$ are in $\overline{\mathcal{A}}$, the uniform closure of $\mathcal{A}$. (That is, $\overline{\mathcal{A}}$ is a function lattice.)

Lemma 12.8. Suppose $\mathcal{A}$ is a function lattice which separates points strongly. Then $\mathcal{A}$ is uniformly dense in $C(X)$.

Proof of the Stone-Weierstrass theorem:
The algebra $\mathcal{A}$ separates points strongly, by Lemma 12.6. Clearly $\overline{\mathcal{A}}$ also separates points strongly, and by Lemma 12.7 it is also a function lattice. Finally, by Lemma 12.8 we have that $\overline{\mathcal{A}}$ is uniformly dense in $C(X)$, so $\overline{\mathcal{A}} = C(X)$, as desired. □

Proof of Lemma 12.6: Given $x, y \in X$, find $f' \in \mathcal{A}$ so $f'(x) = a'$ and $f'(y) = b'$, for some $a' \neq b'$. Then the function $f'' = \frac{f'-a'}{b'-a'}$ satisfies $f''(x) = 0$, and $f''(y) = 1$, so the function $f = (b-a)f'' + a$ has the desired property. □

Proof of Lemma 12.7: Let $\epsilon > 0$. The function $t \mapsto (\epsilon^2 + t)^{1/2}$ has a power series expansion that converges uniformly on $[0, 1]$ (e.g., the Taylor series at $t = 1/2$).

We can thus find a polynomial p so $|(\epsilon^2 + t)^{1/2} - p(t)| < \epsilon$ for all $t \in [0, 1]$.

Observe that at $t = 0$ this gives $|p(0)| < 2\epsilon$, and define $q(t) = p(t) - p(0)$ (still a polynomial). Certainly $q(f) \in \mathcal{A}$ for any $f \in \mathcal{A}$, as $\mathcal{A}$ is an algebra. If $f \in \mathcal{A}$ with $||f||_\infty \leq 1$, we have

$$\begin{aligned} ||q(f^2) - |f|||_\infty &= \sup_{x \in X} |q(f^2(x)) - f^2(x)^{1/2}| \\ &\leq \sup_{t \in [0,1]} |p(t) - p(0) - t^{1/2}| \\ &\leq 2\epsilon + \sup_{t \in [0,1]} |p(t) - t^{1/2}| \\ &\leq 3\epsilon + \sup_{t \in [0,1]} |(\epsilon^2 + t)^{1/2} - t^{1/2}| \\ &\leq 4\epsilon. \end{aligned}$$

Since $q(f^2) \in \mathcal{A}$, we have shown that $|f| \in \overline{\mathcal{A}}$.

Now

$$\max\{f, g\} = \frac{1}{2}(f + g + |f - g|)$$

and

$$\min\{f, g\} = \frac{1}{2}(f + g - |f - g|)$$

so we are finished. □

Proof of Lemma 12.8: Fix $\epsilon > 0$ and $f \in C(X)$. We will find $f_\epsilon \in \mathcal{A}$ with $||f - f_\epsilon||_\infty < \epsilon$.

For each $x, y \in X$, choose $f_{xy} \in \mathcal{A}$ with

$$f_{xy}(x) = f(x) \qquad \text{and} \qquad f_{xy}(y) = f(y)$$

(this is possible because $\mathcal{A}$ separates points strongly). Define the open sets

$$\begin{aligned} U_{xy} &= \{z \in X | f(z) < f_{xy}(z) + \epsilon\} \\ V_{xy} &= \{z \in X | f_{xy}(z) < f(z) + \epsilon\}. \end{aligned}$$

Observe $x, y \in U_{xy} \cap V_{xy}$.

Fix x for a moment. As y varies, the sets U_{xy} cover X. Since X is compact, we can find $y_1, \ldots, y_n$ so $X = \bigcup U_{xy_i}$. Define $f_x = \max\{f_{xy_i}\}$. Since $\mathcal{A}$ is a function lattice, $f_x \in \mathcal{A}$. Moreover, $f(z) < f_x(z) + \epsilon$ for every $z \in X$. Also, if we define $W_x = \bigcap V_{xy_i}$, we see W_x is an open neighbourhood of x, and $f_x(z) < f(z) + \epsilon$ for every $z \in W_x$.

The sets $\{W_x\}_{x\in X}$ cover X, so applying compactness again we find $x_1, \ldots, x_m$ so $X = \bigcup W_{x_i}$. Finally we define $f_\epsilon = \min\{f_{x_i}\}$, which is again in $\mathcal{A}$ as it is a function lattice. Observe that we still have

$$f(z) < f_\epsilon(z) + \epsilon,$$

and now

$$f_\epsilon(z) < f(z) + \epsilon$$

for every $z \in X$, giving the desired result. □

Finally, what about $C(X, \mathbb{C})$, the complex valued continuous functions? We give a slightly revised version of the main theorem:

Theorem 12.9. Let $\mathcal{A}$ be a (complex) subalgebra of $C(X, \mathbb{C})$ which

- is *self-adjoint*, i.e. for every $f \in \mathcal{A}$, the complex conjugate $\overline{f} \in \mathcal{A}$ also,
- contains the complex constants, and
- separates points.

Then $\overline{A} = C(X, \mathbb{C})$.

Proof: We can bootstrap from the real-valued theorem.

Since $\mathcal{A}$ is self-adjoint, if $f \in \mathcal{A}$ then $\Re f \in \mathcal{A}$ and $\Im f \in \mathcal{A}$, since $\Re f = \frac{1}{2}(f + \overline{f})$.

Let

$$A_{\Re} = \{f \in \mathcal{A} | f \text{ is real-valued}\}.$$

Easily, $A_{\Re}$ contains $\mathbb{R}$. We see that it still separates points, as follows. Suppose we have $x, y \in X$, and a complex valued function $f \in \mathcal{A}$ so $f(x) \neq f(y)$. Then for some constant c, $|f(x) + c| \neq |f(y) + c|$. Thus the real-valued function

$$z \mapsto (f(z) + c)\overline{(f(z) + c)}$$

which is still in $\mathcal{A}$ also separates x and y.

Thus by the real-valued version of the theorem we have that $\overline{A_{\Re}} = C(X, \mathbb{R})$. Finally, given $f \in C(X, \mathbb{C})$, we can write $f = \Re f + i\Im f$, and approximate separately the real and imaginary parts using $\mathcal{A}_{\Re}$. □

- Trigonometric polynomials are uniformly dense in $C([0, 1])$ even though the Fourier series need not converge uniformly.
- The hypothesis that $\mathcal{A} \subset C(X, \mathbb{C})$ be self-adjoint is essential. Consider, for example, the holomorphic functions on X the unit disc.