THE WEIERSTRASS and STONE APPROXIMATION THEOREMS

Theorem 1. *Weierstrass approximation theorem (1885).* The space of polynomials is dense in $C[a, b]$, for the uniform topology.

About sixty years later, a generalization was found to algebras of continuous functions on any (Hausdorff) space. (An *algebra* is a linear subspace of $C(X)$ which is closed under pointwise multiplication.).

Theorem 2. (M.H. Stone) Let X be a Hausdorff space. Suppose $\mathcal{A} \subset C(X)$ is a subalgebra of the algebra of continuous real-valued functions on X, satisfying:

(i) $\mathcal{A}$ vanishes nowhere: for all $p \in X$ there exists $f \in \mathcal{A}$ so that $f(p) \neq 0$;

(ii) $\mathcal{A}$ separates points of X: if $p \neq q$, there exists $f \in \mathcal{A}$ so that $f(p) \neq f(q)$. (Note this implies X is Hausdorff.)

Then $\mathcal{A}$ is *dense* in $C(X)$, for the topology of uniform convergence on compact sets.

The proof follows from Weierstrass's theorem and the following even more general fact, of interest in itself:

Theorem 3. *Lattices in $C(X)$ with two-point interpolation are dense.*

Let X be a Hausdorff space, $\mathcal{F} \subset C(X)$ be a subset with the following two properties:

(i) two-point interpolation property: given $p \neq q$ in X and $c_p, c_q \in R$, there exists $f \in \mathcal{F}$ so that $f(p) = c_p, f(q) = c_q$.

(ii) $\mathcal{F}$ is a *lattice*: if $f, g \in \mathcal{F}$, their pointwise min $f \wedge g$ and pointwise max $f \vee g$ are also in $\mathcal{F}$ (and therefore the same is true for any finite number of functions in $\mathcal{F}$).

Then $\mathcal{F}$ is dense in $C(X)$, for the topology of uniform convergence on compact sets.

Proof of Theorem 3. Given $f \in C(X), K \subset X$ compact and $\epsilon > 0$, we must find $g \in \mathcal{F}$ so that $|f(x) - g(x)| < \epsilon$, for all $x \in K$.

For each $p, q \in K$, let $h_{pq} \in \mathcal{F}$ satisfy $h_{pq}(p) = f(p), h_{pq}(q) = f(q)$ ($p = q$ not excluded.) Now fix $p \in K$.

For each $q \in K$ we have a neighborhood V_{pq} so that $f(x) - \epsilon < h_{pq}(x)$, for $x \in V_{pq}$. Thus $K \subset V_{pq_1} \cup \ldots \cup V_{pq_N}$, for some finite set $\{q_1, \ldots, q_N\} \subset K$. In particular, for the pointwise max of the h_{pq_i}:

$$f(x) - \epsilon < (h_{pq_1} \vee \ldots \vee h_{pq_N})(x), x \in K; \quad \text{and } h_p = h_{pq_1} \vee \ldots \vee h_{pq_N} \in \mathcal{F}, h_p(p) = f(p).$$

Now we vary p. For each $p \in K$ we have a neighborhood U_p such that $h_p(x) < f(x)+\epsilon$, for all $x \in U_p$. Thus $K \subset U_{p_1} \cup \ldots \cup U_{p_M}$, for some finite set $\{p_1, \ldots, p_M\} \subset K$. Therefore their pointwise minimum, $g = h_{p_1} \wedge \ldots \wedge h_{p_M}$ is in $\mathcal{F}$ and satisfies, for all $x \in K$:

$$f(x) - \epsilon < g(x) < f(x) + \epsilon,$$

as desired.

The proof of Stone's theorem from Theorem 3 amounts to proving the following two lemmas:

Lemma 1. Any algebra $\mathcal{A} \subset C(X)$ satisfying the non-vanishing and separation properties automatically has the two-point interpolation property.

Proof of Lemma 1. Let $p \neq q$ be points in X, $h \in \mathcal{A}$ with $h(p) \neq h(q)$, $g_1, g_2 \in \mathcal{A}$ with $g_1(p) \neq 0$, $g_2(p) \neq 0$. Then $g = g_1^2 + g_2^2 \in \mathcal{A}$ and $g(p) \neq 0, g(q) \neq 0$. Look for a solution of $f(p) = c_1, f(q) = c_2$ of the form:

$$f = xg + ygh \in \mathcal{A}.$$

The 2×2 linear system for (x, y) defined by $f(p) = c_1, f(q) = c_2$ has determinant $g(p)g(q)(h(q) - h(p)) \neq 0$, and hence has a unique solution.

Lemma 2. Let $\mathcal{A} \subset C(X)$ be an algebra. Then $f \in \mathcal{A} \Rightarrow |f| \in \overline{\mathcal{A}}$ (closure in the topology of uniform convergence on compact sets.)

Indeed Lemma 2 implies the u.o.c closure of any algebra $\mathcal{A}$ in $C(X)$ is a lattice, since, for $f, g \in \mathcal{A}$:

$$f \vee g = \frac{1}{2}(f + g + |f - g|); \quad f \wedge g = \frac{1}{2}(f + g - |f - g|).$$

Proof of Lemma 2.

Let $f \in C(X)$, $K \subset X$ compact, $||f||_K = \sup\{|f(x)|; x \in K\}$. Using the Weierstrass approximation theorem, we find a polynomial p so that $|p(t) - |t|| < \frac{\epsilon}{2}$, for all t with $|t| \leq ||f||_K$. In particular $|p(0)| < \frac{\epsilon}{2}$, so $q(t) = p(t) - p(0)$ is a polynomial with zero constant term ($q(0) = 0$), satisfying $|q(t) - |t|| < \epsilon$ whenever $|t| \leq ||f||_K$.

Thus $g = q(f) \in \mathcal{A}$ (we need $q(0) = 0$ for this, since we don't assume $\mathcal{A}$ includes constant functions other than 0), and:

$$|g(x) - |f(x)|| < \epsilon, \quad \text{for all } x \in K,$$

as we wished to show.

Proof of Stone's Theorem 2. We apply Theorem 3 to the algebra $\overline{\mathcal{A}} \subset C(X)$, which is a lattice (Lemma 2) with two-point interpolation (since $\mathcal{A}$ has this property, by Lemma 1.) The conclusion is $\overline{\mathcal{A}}$ is dense in $C(X)$. But this clearly implies $\mathcal{A}$ itself is dense in $C(X)$ (for the u.o.c. topology, in both cases.)

Proof of the Weierstrass approximation theorem. We consider $C[0,1]$ and $P_n[0,1]$, the subspace of polynomials of degree at most n. There is a "formula" for the approximation, given by *Bernstein polynomials*, the bounded linear operator $B_n : C[0,1] \to P_n[0,1]$, $||B_n|| = 1$:

$$B_n[f] = \sum_{k=0}^{n} C_{n,k} f(\frac{k}{n}) x^k (1-x)^{n-k},$$

where $C_{n,k} = \frac{n!}{(n-k)!k!}$ are binomial coefficients. This is reminiscent of the binomial formula; we see that $B_n[1] = 1$, and differentiating (twice) in x the binomial expansion of $(x+y)^n$ and setting $y = 1-x$, we see that:

$$n = \sum_{k=1}^{n} C_{n,k} k x^k (1-x)^{n-k}, \quad n(n-1)x^2 = \sum_{k=2}^{n} C_{n,k} k(k-1) x^k (1-x)^{n-k}.$$

Thus, we have:

$$B_n[x] = x, \quad B_n[x^2] = x^2 + \frac{1}{n} x(1-x).$$

Exercise 1. Restricted to the space $span_R\{1, x, x^2\}$ of quadratic polynomials, the operator B_n is diagonalizable, with eigenbasis $\{1, x, x(1-x)\}$, eigenvalues $\{1, 1, 1-\frac{1}{n}\}$ (resp.)

Exercise 2. The operator B_n preserves convexity and pointwise ordering of functions.

The proof of the theorem gives a quantitative estimate. Recall the *modulus of continuity* of a function f on $[0,1]$ is defined by:

$$\omega_f(\delta) = \sup\{|f(x) - f(y)|; x, y \in [0,1], |x-y| < \delta\}.$$

By uniform continuity, $f \in C[0,1]$ iff $\omega_f(\delta) \downarrow 0$ as $\delta \downarrow 0$.

Lemma 3. Let $f \in C[0,1]$. Then for all $\delta > 0$,

$$||B_n[f] - f|| < (1 + \frac{1}{4n\delta^2}) \omega_f(\delta)$$

(sup norm). As a corollary, taking $\delta = \frac{1}{\sqrt{n}}$, we have the estimate:

$$||B_n[f] - f|| < \frac{5}{4}\omega_f(\frac{1}{\sqrt{n}}),$$

a quantitative estimate that clearly implies Weierstrass's theorem.

Proof of Lemma 3. We let $r_{nk}(x) = C_{n,k}x^k(1-x)^{n-k} \in P_n[0,1]$, so $\sum_{k=0}^{n} r_{nk}(x) \equiv 1$ on [0,1] and:

$$(f - B_n[f])(x) = \sum_{k=0}^{n}(f(x) - f(\frac{k}{n}))r_{nk}(x).$$

Fix $\delta > 0$ and n. For a given $x \in [0,1]$ split the set of k from 0 to n into two:

$$K_1(x) = \{k; |x - \frac{k}{n}| < \delta\}, \quad K_2(x) = \{k; |x - \frac{k}{n}| \geq \delta.\}.$$

The sum over $K_1(x)$ is estimated in absolute value by $\omega_f(\delta)$.

To estimate the sum over $k \in K_2(x)$, divide the segment from x to each $\frac{k}{n}$ into segments of length δ. We need $\frac{|x-\frac{k}{n}|}{\delta}$ segments, and following the change of f along the chain we find, for $k \in K_2(x)$:

$$|f(x) - f(\frac{k}{n})| \leq \frac{|x - \frac{k}{n}|}{\delta}\omega_f(\delta) \leq \frac{|x - \frac{k}{n}|^2}{\delta^2}\omega_f(\delta),$$

since $\frac{|x-\frac{k}{n}|}{\delta} \geq 1$ for $k \in K_2(x)$. Thus the sum over $K_2(x)$ is estimated by:

$$\frac{\omega_f(\delta)}{\delta^2}\sum_{k=0}^{n}[x^2 - 2x\frac{k}{n} + \frac{k^2}{n^2}]r_{nk}(x)$$

$$= \frac{\omega_f(\delta)}{\delta^2}(x^2 - 2xB_n[x] + B_n[x^2]) = \frac{\omega_f(\delta)}{\delta^2}\frac{1}{n}x(1-x),$$

using the facts about B_n alluded to earlier. Since $x(1-x)$ is bounded above by $\frac{1}{4}$ when $x \in [0,1]$, this concludes the proof.

Remark on complex-valued functions. Let $X = D$, the open unit disk in the complex plane. The algebra $\mathcal{A}$ of complex-valued polynomials on D is non-vanishing and separate points; but its closure in the u.o.c. topology in $C(X, \mathbb{C})$ is the space of all holomorphic functions on D! Thus the Weierstrass theorem is not true in the complex-valued case. But it's true if we add a condition on $\mathcal{A}$ (a nonvanishing, separating algebra in $C(X, \mathbb{C})$), X any Hausdorff space:

$$f \in \mathcal{A} \Rightarrow \overline{f} \in \mathcal{A}.$$

Exercise 3. Under the above conditions, the set $\mathcal{B} = \{f + \overline{f}, i(f - \overline{f}); f \in \mathcal{A}\}$ is an algebra in $C(X, R)$, nonvanishing and separating points (and contained in $\mathcal{A}$).

Thus we can reduce to the real-valued Stone theorem, taking real and imaginary parts of functions in $\mathcal{A}$ (exercise: check this.) Namely, given $g \in C(X, \mathbb{C})$, ϵ-approximate $Re(g), Im(g) \in C(X, R)$ over $K \subset X$ compact by $f_1, f_2 \in \mathcal{B}$. Then with $f = f_1 + if_2 \in \mathcal{A}$, we have:

$$||g - f||_K \leq \sqrt{2}\epsilon.$$

The following are easily proven corollaries of Stone's approximation theorem.

Corollary. (i) Let X be a smooth manifold. If $f \in C(X)$, f can be approximated by smooth functions $g \in C^\infty(X)$, uniformly on compact subsets of X.

(ii) If X is a smooth manifold, any continuous map $f : X \to R^n$ may be approximated by smooth maps $g : X \to R^n$, uniformly on compact subsets of X. If f is a proper map, the approximating smooth maps g may be taken to be proper maps as well.

(iii) if X is a smooth manifold, any continuous map to the n-sphere, $f : X \to S^n$, may be approximated by smooth maps $g : X \to S^n$, uniformly over compact subsets of X. If X is compact, the approximating smooth maps g may be taken homotopic to f (as maps to S^n).