The Weierstrass Approximation Theorem states that a function in $C[a,b]$ can be uniformly approximated by a polynomial. One way of expressing this fact is that given $f \in C[a,b]$ and $\epsilon > 0$, there exists $p \in P[a,b]$ such that $|f(x) - p(x)| \leq \epsilon$ for every $x \in [a,b]$. Using the vocabulary of norms, this is equivalent to

$$||f - p||_\infty \leq \epsilon$$

The same idea can also be expressed in terms of the closure of $P[a,b]$ in $C[a,b]$. Recall that given a set A in a metric space M, $x \in \bar{A}$ if and only if for every $\epsilon > 0$, $B_\epsilon(x) \cap A \neq \emptyset$. Hence the Weierstrass Approximation Theorem asserts that $C[a,b] \subseteq \overline{P[a,b]}$. But of course $\overline{P[a,b]} \subseteq C[a,b]$. Hence we have arrived at a concise statement of the theorem.

Theorem 1: (Weierstrass Approximation Theorem) $\overline{P[a,b]} = C[a,b]$, where closure is taken with respect to the uniform norm.

You are already familiar with the idea of writing certain functions as power series. For example,

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^{n+1} x^{2n+1}}{n!}.$$

This series converges pointwise on all of $\mathbb{R}$ (verify this with the ratio test) and therefore uniformly on any fixed interval $[-R,R]$. (Recall Theorem 10.10) Hence, given any $\epsilon > 0$, we can find an N such that

$$\left| \sin(x) - \sum_{n=0}^{N} \frac{(-1)^{n+1} x^{2n+1}}{n!} \right| \leq \epsilon$$

for every $x \in [-\pi,\pi]$. So sin can be approximated uniformly by polynomials on $[-\pi,\pi]$. But functions that can be written as power series are special; in particular they are infinitely differentiable – this is a consequence of Theorem 10.10.

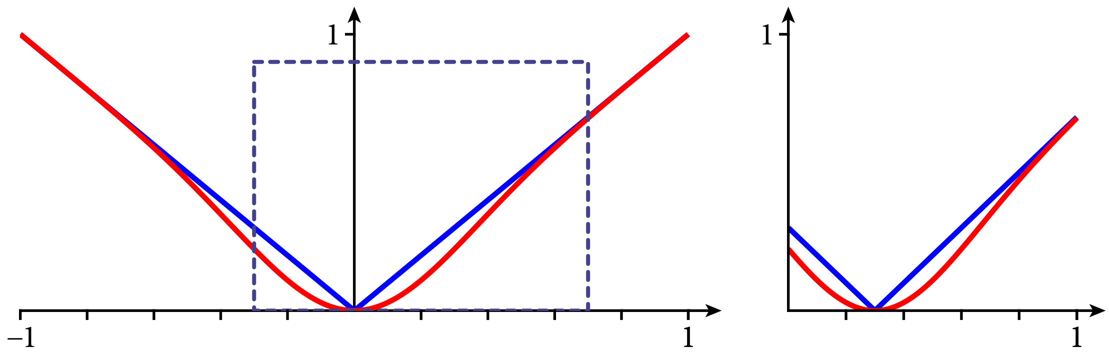
The remarkable part about the Weierstrass Approximation Theorem is that every continuous function, even the non-differentiable ones, can can be uniformly approximating by polynomials. Interestingly, the proof of this fact can be reduced to showing that just one non-smooth function, the absolute value function abs, can be uniformly approximated by polynomials.

Proposition 2: abs $\in \overline{P[-1,1]}$.

Supposing for the moment we have proved this result, let's see how this results in a fairly easy proof of Theorem 1. First, we show that any translate of the absolute value function is in $\overline{P[0,1]}$. We define

$$\text{abs}_a(x) = |x - a|.$$

Lemma 3: For any $a \in \mathbb{R}$, $\text{abs}_a \in \overline{P[0,1]}$.



Approximating abs_a on $[0,1]$.

Proof. If $a \le 0$ or $a \ge 1$, abs_a is linear on $[0,1]$ and hence in $P[0,1]$.

Suppose $0 < a < 1$ and let $\varepsilon > 0$. Let p be a polynomial such that

$$|p(x) - \mathrm{abs}(x)| < \varepsilon$$

for every $x \in [-1,1]$. Define $q(x) = p(x-a)$, so q is a polynomial. Then

$$\begin{aligned} \sup_{x\in[0,1]} |q(x) - \mathrm{abs}_a(x)| &= \sup_{x\in[0,1]} |p(x-a) - |x-a|| \\ &= \sup_{x\in[-a,1-a]} |p(x) - |x|| \\ &\le \sup_{x\in[-1,1]} |p(x) - \mathrm{abs}(x)| \le \varepsilon \end{aligned}$$

Hence $\|q - \mathrm{abs}_a\|_{C[0,1]} \le \varepsilon$ Since q is a polynomial and $\varepsilon > 0$ is arbitrary, $\mathrm{abs}_a \in \overline{P[0,1]}$. □

A function $f \in C[0,1]$ is called piecewise linear if there is a partition $0 = x_0 < x_1 < \cdots < x_n = 1$ such that the restriction of f to each interval $[x_{k-1}, x_k]$ is linear; we denote by $\mathrm{PL}[0,1]$ the collection of all such functions. Clearly any linear combination of functions of the form abs_a belongs to $\mathrm{PL}[0,1]$. We now show that these functions span all of $\mathrm{PL}[0,1]$.

Proposition 4: Let $f \in \mathrm{PL}[0,1]$, and let $0 = x_0 < x_1 < \cdots < x_n = 1$ be a partition such that f is linear on each interval $I_k = [x_{k-1}, x_k]$. Then f is a linear combination of the functions 1 and $\{\mathrm{abs}_{x_k} : 0 \le k \le n\}$.

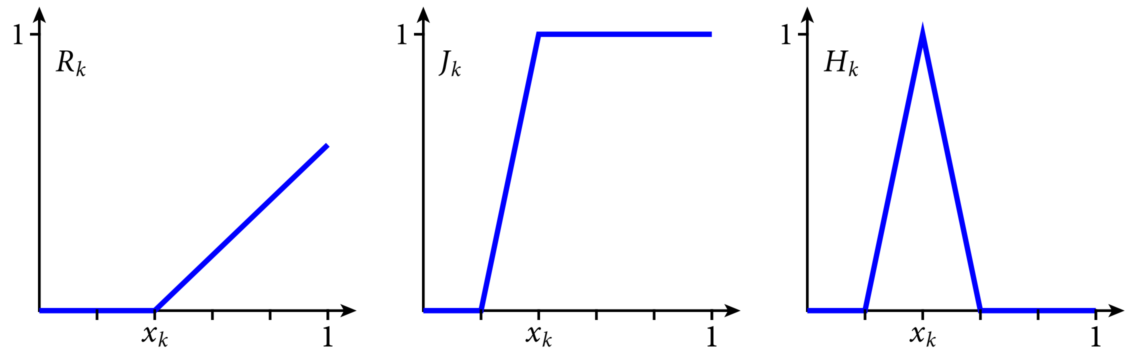
Proof. Let $S = \mathrm{span}\,\{\mathrm{abs}_{x_k} : 0 \le k \le n\}$. Notice that

$$\mathrm{abs}_{x_0}(x) + \mathrm{abs}_{x_n}(x) = x + (1-x) = 1.$$

Hence the constants belong to S.

For $0 \le k \le 1$, let

$$R_k(x) = \frac{1}{2}\left(\mathrm{abs}_{x_k}(x) + (x - x_k)\right).$$



The functions R_k, J_k, and H_k.

Then R_k is a linear combination of 1, abs_{x_0}, and abs_{x_k} and hence $R_k \in S$.

Notice that $R_k(x) = 0$ if $x \le x_k$ and $R_k(x) = x - x_k$ otherwise. For $1 \le k \le n$ let

$$J_k = \frac{R_k - R_{k-1}}{x_k - x_{k-1}},$$

and let $J_0 = 1$ and $J_{n+1} = 0$. Then each $J_k \in S$ and

$$J_k(x_j) = \begin{cases} 0 & j < k \\ 1 & j \ge k. \end{cases}$$

Finally, let $H_k = J_k - J_{k+1}$ for $0 \le k \le n$. Then $H_k \in S$ for each k, and

$$H_k(x_j) = \begin{cases} 1 & k = j \\ 0 & k \ne j. \end{cases}$$

Hence

$$\sum_{k=0}^{n} f(x_k) H_k$$

is a piecewise linear function that agrees with f at each point x_k. We conclude that

$$f = \sum_{k=0}^{n} f(x_k) H_k.$$

Since each $H_k \in S$, we conclude that $f \in S$. □

We have seen that each $\text{abs}_a \in \overline{P[0,1]}$ and that each $f \in PL[0,1]$ is a linear combination of functions abs_a. To show that $PL[0,1] \subseteq \overline{P[0,1]}$ we now take advantage of the idea that the metric space and the vector space structures of a normed vector space are compatible.

Proposition 5: Let X be a normed linear space and let W be a subspace of X. Then $\overline{W}$ is a subspace of X.

Proof. Let $x, y \in \overline{W}$. Let (x_n) and (y_n) be sequences in W converging to x and y. Then $\|(x+y) - (x_n + y_n)\| \leq \|x - x_n\| + \|y - y_n\|$ and therefore $(x_n + y_n) \to (x+y)$. Hence $x + y \in \overline{W}$. Similarly, $\alpha x_n \to \alpha x$ and hence $\alpha x \in \overline{W}$. So $\overline{W}$ is a subspace. □

We can now prove the Weierstrass Approximation Theorem, at least for the domain $[0,1]$.

Proposition 6: $C[0,1] = \overline{P[0,1]}$.

Proof. Proposition 5 implies that $\overline{P[0,1]}$ is a subspace of $C[0,1]$ since $P[0,1]$ is. Suppose $f \in PL[0,1]$. Proposition 4 shows that f can be written as a finite linear combination of functions abs_a, and Proposition 3 implies that each $\text{abs}_a \in \overline{P[0,1]}$. Since $\overline{P[0,1]}$ is a subspace, we conclude that $f \in \overline{P[0,1]}$ and hence $PL[0,1] \subseteq \overline{P[0,1]}$. Consequently $\overline{\text{PL}[0,1]} \subseteq \overline{P[0,1]}$. From the proof of Carothers 11.2 it follows that $C[0,1] = \overline{\text{PL}[0,1]}$. Hence $\overline{P[0,1]} = C[0,1]$. □

Exercise 1: Use Proposition 6 to prove the Weierstrass Approximation Theorem for an aribitrary interval $[a,b]$. Hint: Given $f \in C[a,b]$, define $g(x) = f(a + x(b-a)$. Approximate g in $C[0,1]$ by $p \in P[0,1]$, and define $q(x) = p((x-a)/(b-a)$.

It remains to prove Proposition 2, which we do now.

Proof. For $0 \leq x \leq 1$, define $P_0(x) = 0$ and for $k \geq 0$ define

$$P_{k+1}(x) = P_k(x) + \frac{x - P_k^2}{2}.$$

We claim that $0 \leq P_k(x) \leq \sqrt{x}$ for every $k \geq 0$ and that $P_{k+1} \geq P_k$ for every k. This is certainly true for $k = 0$. Suppose $0 \leq P_k(x) \leq \sqrt{x}$. Then

$$P_{k+1}(x) = P_k(x) + \frac{x - P_k^2}{2} \geq P_k(x)$$

so $P_{k+1}(x) \geq 0$. But also, since $0 \leq P_k(x) \leq \sqrt{x} \leq 1$, we have

$$P_{k+1}(x) = P_k(x) + \frac{x - P_k^2}{2} = P_k(x) + \frac{1}{2}(\sqrt{x} + P_k(x))(\sqrt{x} - P_k(x)) \leq P_k(x) + (\sqrt{x} - P_k(x)) = \sqrt{x}.$$

Hence $P_{k+1}(x) \leq \sqrt{x}$. We have therefore shown inductively that $0 \leq P_k(x) \leq \sqrt{x}$ for every $k \geq 0$. As seen above, this also implies that $P_{k+1} \geq P_k(x)$.

It follows that for any fixed $x \in [0,1]$, $\{P_k(x)\}$ is monotone increasing and bounded above by 1, and hence converges to a limit $P(x) \leq 1$. But then $P(x)$ satisfies

$$P(x) = P(x) + \frac{x - P(x)^2}{2}$$

and hence

$$P(x)^2 = x.$$

Since $P(x) \geq 0$, we conclude that $P(x) = \sqrt{x}$ and P_k converges pointwise to the square root function. Since the convergence is monotone and the limit function is continuous, Dini's theorem implies that the convergence is actually uniform.

Now let $\varepsilon > 0$. Pick k so that $\left|P_k(x) - \sqrt{x}\right| < \varepsilon$ for all $x \in [0,1]$. Define $q(y) = P_k(y^2)$ for $y \in [-1,1]$, so q is a polynomial. Then for any $y \in [-1,1]$,

$$|q(y) - \operatorname{abs} y| = \left|P_k(y^2) - \sqrt{y^2}\right| < \varepsilon$$

since $y^2 \in [0,1]$. Since $\varepsilon > 0$ is arbitrary, we conclude that $\operatorname{abs} \in P[0,1]$.