# Lecture 16: The Stone-Weierstrass theorem

updated 25/9

The subject of today's lecture is Weierstrass's approximation theorem and its generalisation, the Stone–Weierstrass theorem, which tell us in particular that any continuous function on $[a,b]$ (resp. $S^1$) may be approximated arbitrarily well by a polynomial (resp. a trigonometric polynomial), which is to say that polynomials give a dense subspace of $\mathrm{Cts}([a,b],\mathbb{R})$ (resp. $\mathrm{Cts}(S^1,\mathbb{R})$).

Recall: We have associated a space of functions $\mathrm{Cts}(X,Y)$ to any pair of topological spaces $X, Y$ (see Lecture 12) with a list of good properties:

- if $F: Z\times X \to Y$ is continuous, so is $Z \to \mathrm{Cts}(X,Y)$ defined by $z \mapsto F(z,-)$.
- if $X$ is locally compact Hausdorff $\mathrm{Cts}(Z\times X, Y) \cong \mathrm{Cts}(Z, \mathrm{Cts}(X,Y))$ (see Theorem L12-4 and Ex. L12-13).
- if $X$ is compact and $(Y, d_Y)$ is a metric space then $\mathrm{Cts}(X,Y)$ is a metric space with the sup metric, and moreover if $Y$ is complete so too is $\mathrm{Cts}(X,Y)$ (see Lecture 13, specifically Theorem L13-2 and Corollary L13-6).

We have applied this theory to prove the existence of solutions to ODEs (Lecture 15), and we observed that for polynomial ODEs the solutions could be approximated by polynomials (see Remark L15-2). Just as our ability to compute effectively with real numbers is predicated on $\overline{\mathbb{Q}} = \mathbb{R}$, our ability to work with function spaces $\mathrm{Cts}(X,\mathbb{R})$ is often predicated on identifying a class of "simple" functions

$$A \subseteq \mathrm{Cts}(X,\mathbb{R}) \qquad \text{with} \qquad \overline{A} = \mathrm{Cts}(X,\mathbb{R}).$$

If $X = [a,b]$ and $A$ is all polynomial functions, this works:

<u>Theorem L16-0</u> (Weierstrass, 1885) Let $f \in \mathrm{Cts}([a,b], \mathbb{R})$. Then there is a sequence of polynomials $p_n(x)$ which converges uniformly to $f(x)$ on $[a,b]$.

We need a few ingredients before we are ready for the proof (the proof we will give is not Weierstrass's original one: it is due to Bernstein, see K. Davidson and A. Donsig's "Real analysis with real applications" 2002).

<u>Exercise L16-0</u> Prove that if $f : (X, d_X) \longrightarrow (Y, d_Y)$ is continuous and $X$ is compact then $f$ is <u>uniformly continuous</u>, that is

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x_1, x_2 \in X \left( d_X(x_1, x_2) < \delta \Rightarrow d_Y(fx_1, fx_2) < \varepsilon \right).$$

<u>Def$^n$</u> Given a function $f : [0,1] \longrightarrow \mathbb{R}$ the nth <u>Bernstein polynomial</u> $B_n(f)$ is

$$B_n(f) := \sum_{k=0}^{n} f\left(\tfrac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}.$$

To avoid confusion we adopt the convention of writing $f$ as $f(z)$ to distinguish the input variable of $f$ from the $x$ in $B_n(f)$. Clearly $B_n(-)$ is linear, so $B_n(f+g) = B_n(f) + B_n(g)$ and $B_n(\lambda f) = \lambda B_n(f)$ for any scalar $\lambda \in \mathbb{R}$.

<u>Lemma L16-½</u> We have for $n \geq 1$

$$B_n(1) = 1, \qquad B_n(z) = x, \qquad B_n(z^2) = \frac{n-1}{n} x^2 + \frac{1}{n} x.$$

<u>Proof</u> The binomial theorem gives

$$B_n(1) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = (x + (1-x))^n = 1.$$

Note the following identity of polynomials in $x, y$ for $n \geq 1$

$$\frac{\partial}{\partial x}\left( \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \right) = \frac{\partial}{\partial x}\left( (x+y)^n \right) = n(x+y)^{n-1}$$

but computing differently, as $\sum_k \binom{n}{k} \frac{\partial}{\partial x}(x^k) y^{n-k}$ we obtain

$$\sum_{k=0}^{n} \binom{n}{k} k \cdot x^{k-1} y^{n-k} = n(x+y)^{n-1}$$

multiplying both sides by $\frac{x}{n}$ gives

$$\sum_{k=0}^{n} \binom{n}{k} \frac{k}{n} x^k y^{n-k} = x(x+y)^{n-1} \tag{3.1}$$

substituting $y = 1-x$ gives $B_n(z) = x$. For the remaining identity, we differentiate (3.1) again with respect to $x$, obtaining

zero if $n=1$

$$\sum_{k=0}^{n} \binom{n}{k} \frac{k}{n} \cdot k \cdot x^{k-1} y^{n-k} = (x+y)^{n-1} + (n-1)x(x+y)^{n-2}$$

again multiplying both sides by $\frac{x}{n}$ gives

$$\sum_{k=0}^{n} \binom{n}{k} \frac{k^2}{n^2} x^k y^{n-k} = \frac{x}{n}(x+y)^{n-1} + \frac{n-1}{n} x^2 (x+y)^{n-2} \tag{3.2}$$

substituting $y = 1-x$ gives the formula for $B_n(z^2)$. $\square$

Proof of Theorem L16-0 First we prove the $[a,b] = [0,1]$ case. Let continuous $f : [0,1] \longrightarrow \mathbb{R}$ be given. We claim $B_n(f) \longrightarrow f$ with respect to $d_\infty$. Since $f$ is continuous it is, by Ex. L16-0, uniformly continuous. Given $\varepsilon > 0$ let $\delta > 0$ be such that

$$|x-y| < \delta \implies |f(x) - f(y)| < \varepsilon/2 \qquad \forall x, y \in [0,1].$$

Since $[0,1]$ is compact $f$ is bounded, say $|f(x)| \leq M$ for all $x \in [0,1]$.

<u>Claim</u> For any $x, y \in [0,1]$, $|f(x) - f(y)| \leq 2M\left(\frac{x-y}{\delta}\right)^2 + \varepsilon/2$

<u>Proof of claim</u> if $|x-y| < \delta$ then $|f(x) - f(y)| < \varepsilon/2$ so this is clear.
Otherwise if $|x-y| \geqslant \delta$ then $\left(\frac{x-y}{\delta}\right)^2 \geqslant 1$ so

$$|f(x) - f(y)| \leq 2M \leq 2M\left(x-y/\delta\right)^2 + \varepsilon/2. \ \square$$

Now observe that for a constant $x_0 \in [0,1]$, we have an equality of polynomials in $x$,

$$B_n(f - f(x_0)) = B_n(f) - f(x_0)B_n(1) = B_n(f) - f(x_0).$$

Hence for $x \in [0,1]$ we have

$$|B_n(f) - f(x_0)| = |B_n(f - f(x_0))|$$

clearly if $f(z) \leq g(z)$ for all $z \in [0,1]$
then $B_n(f)(x) \leq B_n(g)(x)$, $x \in [0,1]$

$$\leq B_n(|f - f(x_0)|)$$

reading this as a function of $z$.

by the claim above

$$\leq B_n\left(2M\left(\frac{z - x_0}{\delta}\right)^2 + \varepsilon/2\right)$$

$$= \frac{2M}{\delta^2} B_n\left((z - x_0)^2\right) + \varepsilon/2$$

$$= \frac{2M}{\delta^2}\left[B_n\left(z^2 - 2x_0 z + x_0^2\right)\right] + \varepsilon/2$$

$$= \frac{2M}{\delta^2}\left[B_n(z^2) - 2x_0 B_n(z) + x_0^2 B_n(1)\right] + \varepsilon/2$$

$$= \frac{2M}{\delta^2}\left[\frac{n-1}{n}x^2 + \frac{1}{n}x - 2x_0 x + x_0^2\right] + \varepsilon/2$$

$$= \frac{2M}{\delta^2}\left[\frac{1}{n}(x - x^2) + (x - x_0)^2\right] + \varepsilon/2$$

Now substituting $x = x_0$, we have

$$\begin{aligned} |B_n(f)(x_0) - f(x_0)| &\leqslant \varepsilon/2 + \frac{2M}{\delta^2}\frac{1}{n}(x_0 - x_0^2) \\ &\leqslant \varepsilon/2 + \frac{2M}{\delta^2}\cdot\frac{1}{n}\cdot\frac{1}{4} \\ &= \varepsilon/2 + \frac{M}{2\delta^2 n} \end{aligned}$$

But this is true for all $x_0 \in [0,1]$, so

$$d_\infty(B_n(f), f) \leqslant \varepsilon/2 + \frac{M}{2\delta^2 n}$$

If we take $N \geqslant M/\delta^2\varepsilon$ then for all $n \geqslant N$, we have $M/2\delta^2 n \leq \varepsilon/2$ and so

$$d_\infty(B_n(f), f) \leqslant \varepsilon/2 + \varepsilon/2 \leq \varepsilon$$

which proves that $B_n(f) \longrightarrow f$ in $(Cts([0,1], \mathbb{R}), d_\infty)$. This completes the proof of the $[0,1]$ case. For the general case, observe that $\phi : [0,1] \longrightarrow [a,b]$ $\phi(x) = (b-a)x + a$ is a homeomorphism, and if $f : [a,b] \longrightarrow \mathbb{R}$ is continuous then $g = f \circ \phi$ is continuous and with $B_n(f) := B_n(g) \circ \phi^{-1}$

$$\begin{aligned} d_\infty(B_n(f), f) &= \sup\{ |B_n(f)(x) - f(x)| \mid x \in [a,b] \} \\ &= \sup\{ |B_n(g)(\phi^{-1}x) - g(\phi^{-1}x)| \mid x \in [a,b] \} \\ &= \sup\{ |B_n(g)(y) - g(y)| \mid y \in [0,1] \} \\ &= d_\infty(B_n(g), g). \end{aligned}$$

Hence $B_n(f) \longrightarrow f$ in $Cts([a,b], \mathbb{R})$ and moreover $B_n(f)$ is clearly a polynomial. $\square$

<u>Exercise L16-1</u> Let $X$ be compact, $(Y, d_Y)$ a metric space. Given a subset $A \subseteq Cts(X,Y)$ the following conditions on $f \in Cts(X,Y)$ are equivalent

(i) $f \in \overline{A}$
(ii) there is a sequence $(a_n)_{n=0}^{\infty}$ in $A$ converging uniformly to $f$
(iii) $f$ may be uniformly approximated by elements of $A$, that is, given $\varepsilon > 0$ there exists $a \in A$ such that $|f(x) - a(x)| < \varepsilon$ for all $x \in X$.

<u>Def$^n$</u> A subset $A$ of a topological space $X$ is <u>dense</u> if $\overline{A} = X$.

Next we turn to a generalisation of the Weierstrass approximation theorem which will apply to any compact $X \subseteq \mathbb{R}^n$, the <u>Stone-Weierstrass theorem</u>. But first we need to talk briefly about $Cts(X, \mathbb{R})$ as an <u>algebra</u>. Recall that the addition and multiplication give continuous maps

$$+ : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}, \qquad \bullet : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$$

and hence given $f, g \in Cts(X, \mathbb{R})$ (here $X$ is any space) we have continuous maps

$$fg : X \xrightarrow{\Delta} X \times X \xrightarrow{f \times g} \mathbb{R} \times \mathbb{R} \xrightarrow{\bullet} \mathbb{R} \qquad x \longmapsto f(x)g(x)$$

$$f+g : X \xrightarrow{\Delta} X \times X \xrightarrow{f \times g} \mathbb{R} \times \mathbb{R} \xrightarrow{+} \mathbb{R} \qquad x \longmapsto f(x) + g(x)$$

Here we are using the diagonal $\Delta(x) = (x, x)$, and the product $f \times g$ (see Ex. L12-2). Moreover for fixed $\lambda \in \mathbb{R}$ the map

$$\lambda f : X \xrightarrow{f} \mathbb{R} \xrightarrow{\lambda \cdot (-)} \mathbb{R} \qquad x \longmapsto \lambda \cdot f(x)$$

is continuous. Let $a : \mathbb{R} \times \mathrm{Cts}(X,\mathbb{R}) \longrightarrow \mathrm{Cts}(X,\mathbb{R})$ be $(\lambda, f) \longmapsto \lambda f$. For any $c \in \mathbb{R}$ the constant function is continuous:

$$X \longrightarrow \{*\} \xrightarrow{\ c\ } \mathbb{R} \qquad x \longmapsto c$$

Usually we denote this function again by $c$. Note it is $a(c, 1)$.

<u>Exercise L16-2</u> Check that $\mathrm{Cts}(X,\mathbb{R})$ with the above structures is a <u>commutative algebra</u> (over $\mathbb{R}$) for any space $X$, which is to say that

i.e. addition, multiplication are same operations

- $(\mathrm{Cts}(X,\mathbb{R}), +, a)$ is an $\mathbb{R}$-vector space.
- $f(gh) = (fg)h$ for all $f, g, h \in \mathrm{Cts}(X,\mathbb{R})$
- $1f = f1 = f$ for all $f \in \mathrm{Cts}(X,\mathbb{R})$ where $1(x) = 1 \in \mathbb{R}$.
- $fg = gf$ for all $f, g \in \mathrm{Cts}(X,\mathbb{R})$.
- $f(g+h) = fg + fh$ for all $f, g, h \in \mathrm{Cts}(X,\mathbb{R})$.
- $(\lambda f)g = f(\lambda g) = \lambda \cdot fg$

(Note: occurrences of brackets above do <u>not</u> mean evaluation). A subset $A \subseteq \mathrm{Cts}(X,\mathbb{R})$ is a <u>subalgebra</u> if $1 \in A$, and whenever $f, g \in A$ we have $f+g \in A$, $fg \in A$ and $\lambda f \in A$ for any $\lambda \in \mathbb{R}$. For example, the constant functions give a subalgebra of $\mathrm{Cts}(X,\mathbb{R})$ isomorphic to $\mathbb{R}$, and moreover <u>every</u> subalgebra contains the constant functions.

<u>Def$^n$</u> A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is <u>polynomial</u> if there exists a function $F : \mathbb{N}^n \longrightarrow \mathbb{R}$ (where $\mathbb{N} = \{0,1,\ldots\}$) with the property that $\{ \underline{N} \in \mathbb{N}^n \mid F(\underline{N}) \neq 0 \}$ is finite and for all $x \in \mathbb{R}^n$ (write $\underline{N}$ for $(N_1, \ldots, N_n)$)

$$f(x) = \sum_{\underline{N} \in \mathbb{N}^n} F(\underline{N}) \, \pi_1(x)^{N_1} \cdots \pi_n(x)^{N_n} \qquad (4.1)$$

where $\pi_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ are the projection maps $\pi_i(x_1, \ldots, x_n) = x_i$. We denote by $\mathrm{Poly}(\mathbb{R}^n, \mathbb{R})$ the set of polynomial functions $\mathbb{R}^n \longrightarrow \mathbb{R}$.

<u>Lemma L16-1</u> Every polynomial function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is continuous, and $\mathrm{Poly}(\mathbb{R}^n, \mathbb{R})$ is the smallest subalgebra of $\mathrm{Cts}(\mathbb{R}^n, \mathbb{R})$ containing $\pi_1, \ldots, \pi_n$. We say that $\mathrm{Poly}(\mathbb{R}^n, \mathbb{R})$ is <u>generated</u> as an algebra by the set $\{\pi_1, \ldots, \pi_n\}$.

<u>Proof</u> The polynomial function $f$ of (4.1) may be written as

$$f = \sum_{\underline{N} \in \mathbb{N}^n} F(\underline{N}) \, \pi_1^{N_1} \cdots \pi_n^{N_n}$$

where the products (e.g. $\pi_1^{N_1} = \pi_1 \cdots \pi_1$), scalar multiplications and sums are all the algebra operations in $\mathrm{Cts}(\mathbb{R}^n, \mathbb{R})$ as defined above. Since the set of continuous functions is <u>closed</u> under these operations (and the $\pi_i$ are continuous), $f$ must be continuous. Moreover if a subalgebra $A \subseteq \mathrm{Cts}(\mathbb{R}^n, \mathbb{R})$ contains $\{\pi_1, \ldots, \pi_n\}$ it must contain $f$, and the subset $\mathrm{Poly}(\mathbb{R}^n, \mathbb{R})$ is closed under addition, multiplication and scalar multiplication (and contains 1) so it is a subalgebra, implying the second claim. $\square$

<u>Def$^n$</u> An <u>embedding</u> is an injective continuous map $j : X \longrightarrow Y$ such that the induced continuous map $X \longrightarrow j(X)$ is a homeomorphism (where $j(X)$ has the subspace topology). We say $j$ is a homeomorphism onto its image. Roughly speaking we identify $X$ as a subspace of $Y$ via $j$.

**Example L16-1** Given a subspace $X \subseteq Y$ the inclusion $X \longrightarrow Y$ is an embedding.

**Def$^n$** Given an embedding $j : X \longrightarrow \mathbb{R}^n$ we define the subspace $\mathrm{Poly}(X, j, \mathbb{R})$ of $\mathrm{Cts}(X, \mathbb{R})$ to be the image of

$$\mathrm{Poly}(\mathbb{R}^n, \mathbb{R}) \overset{inc}{\hookrightarrow} \mathrm{Cts}(\mathbb{R}^n, \mathbb{R}) \xrightarrow{(-)\circ j} \mathrm{Cts}(X, \mathbb{R})$$

that is, the set of continuous maps which are "restrictions" to $X$ of polynomial functions on $\mathbb{R}^n$, where "restriction" means precomposition with $j$. If the embedding is clear from the context we write $\mathrm{Poly}(X, \mathbb{R})$ for $\mathrm{Poly}(X, j, \mathbb{R})$.

**Exercise L16-3** Prove $\mathrm{Poly}(X, j, \mathbb{R})$ is the smallest subalgebra of $\mathrm{Cts}(X, \mathbb{R})$ containing the functions $\{\pi_1 \circ j, \ldots, \pi_n \circ j\}$.

**Example L16-2** Let $S^1 = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$, and let $j_1 : S^1 \longrightarrow \mathbb{R}^2$ be the inclusion. Let $j_2$ be the composite

$$S^1 \xrightarrow{j_1} \mathbb{R}^2 \xrightarrow{R_\theta} \mathbb{R}^2$$

where $R_\theta$ is multiplication by $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$. Since $R_\theta$ is a homeomorphism this is again an embedding. Then

$$(\pi_1 \circ j_2)(x,y) = x\cos\theta - y\sin\theta$$
$$(\pi_2 \circ j_2)(x,y) = x\sin\theta + y\cos\theta$$

Since $\theta$ is fixed these are polynomial functions of $x, y$ and so $\mathrm{Poly}(S^1, j_2, \mathbb{R}) \subseteq \mathrm{Poly}(S^1, j_1, \mathbb{R})$. Since $j_1 = R_{-\theta} \circ j_2$ the same argument shows $\mathrm{Poly}(S^1, j_1, \mathbb{R}) = \mathrm{Poly}(S^1, j_2, \mathbb{R})$.

However in general $Poly(X, j, \mathbb{R})$ does depend on $j$:

Example L16-3 Let $j_1, j_2 : (0,1) \longrightarrow \mathbb{R}$ be $j_1(x) = x$, $j_2(x) = x^2$. These are both embeddings, but the function $x^3 : (0,1) \to \mathbb{R}$ lies in $Poly((0,1), j_1, \mathbb{R})$ but not in $Poly((0,1), j_2, \mathbb{R})$.

Def$^n$ We say a subalgebra $A \subseteq Cts(X, \mathbb{R})$ separates points if whenever $x, y \in X$ are distinct points there exists $f \in A$ with $f(x) \neq f(y)$.

Lemma L16-2 If $j : X \longrightarrow \mathbb{R}^n$ is an embedding then the subalgebra $Poly(X, j, \mathbb{R}) \subseteq Cts(X, \mathbb{R})$ separates points.

Proof If $x, y \in X$ are distinct, then for some $1 \leq i \leq n$ we have $\pi_i(jx) \neq \pi_i(jy)$, and so $\pi_i \circ j \in Poly(X, j, \mathbb{R})$ will do. $\square$

Example L16-4 Consider the embedding

$$j : \mathbb{R}/2\pi\mathbb{Z} \longrightarrow \mathbb{R}^2, \quad j(\theta) = (\cos\theta, \sin\theta)$$

where $\mathbb{R}/2\pi\mathbb{Z}$ is the quotient of $\mathbb{R}$ by the relation $\lambda \sim \mu$ if $\lambda - \mu \in 2\pi\mathbb{Z}$ (see Tutorial 4). We claim that $A = Poly(\mathbb{R}/2\pi\mathbb{Z}, j, \mathbb{R})$ is the smallest subalgebra of $Cts(\mathbb{R}/2\pi\mathbb{Z}, \mathbb{R})$ containing the set $\{\cos(n\theta), \sin(n\theta)\}_{n \in \mathbb{Z}}$. By definition $A$ is the smallest subalgebra containing $\cos\theta$, $\sin\theta$, so the claim follows from

$$\cos(n\theta) = Re(e^{in\theta}) = Re([\cos\theta + i\sin\theta]^n) \in A$$
$$\sin(n\theta) = Im(e^{in\theta}) = Im([\cos\theta + i\sin\theta]^n) \in A$$

using the binomial formula (this does $n > 0$, but this suffices).

<u>Lemma L16-3</u> With the above notation, the elements of $\mathrm{Poly}(\mathbb{R}/2\pi\mathbb{Z}, j, \mathbb{R})$ are precisely the functions

$$f(\theta) = a_0 + \sum_{n=1}^{N} \big( a_n \cos(n\theta) + b_n \sin(n\theta) \big) \qquad (7.1)$$

for some $a_0, a_1, \ldots, a_N, b_1, \ldots, b_N \in \mathbb{R}$, and $N \geq 1$. This collection of functions therefore separates points of $\mathbb{R}/2\pi\mathbb{Z}$. We call such functions <u>trigonometric polynomials</u>.

<u>Proof</u> Clearly these expressions give functions in $\mathrm{Poly}(\mathbb{R}/2\pi\mathbb{Z}, j, \mathbb{R})$, so it suffices to prove functions of this form compose a <u>subalgebra</u> of $\mathrm{Cts}(\mathbb{R}/2\pi\mathbb{Z}, \mathbb{R})$. For this it is enough to observe that these functions are closed under multiplication:

$$\begin{aligned}
\sin(mt)\cos(nt) &= \tfrac{1}{2}\big[\sin((m+n)t) + \sin((m-n)t)\big] \\
\sin(mt)\sin(nt) &= \tfrac{1}{2}\big[\cos((m-n)t) - \cos((m+n)t)\big] \\
\cos(mt)\cos(nt) &= \tfrac{1}{2}\big[\cos((m-n)t) + \cos((m+n)t)\big].
\end{aligned}$$

The claim about separating points is now immediate from Lemma L16-2. □

<u>Theorem L16-3</u> (Stone-Weierstrass) Let $X$ be a compact Hausdorff space and $A \subseteq \mathrm{Cts}(X, \mathbb{R})$ a subalgebra which separates points. Then we have $\overline{A} = \mathrm{Cts}(X, \mathbb{R})$.

<u>Corollary L16-4</u> Given $X \subseteq \mathbb{R}^n$ compact we have

$$\overline{\mathrm{Poly}(X, \mathbb{R})} = \mathrm{Cts}(X, \mathbb{R}).$$

<u>Proof</u> Immediate from the theorem and Lemma L16-2. □

<u>Corollary L16-5</u> The trigonometric polynomials are dense in $\mathrm{Cts}(\mathbb{R}/2\pi\mathbb{Z}, \mathbb{R})$, i.e.

$$\overline{\mathrm{Poly}(\mathbb{R}/2\pi\mathbb{Z}, j, \mathbb{R})} = \mathrm{Cts}(\mathbb{R}/2\pi\mathbb{Z}, \mathbb{R}).$$

<u>Proof</u> Again, immediate from the theorem and Lemma L16-3. $\square$

Of course $S^1 \cong \mathbb{R}/2\pi\mathbb{Z}$ so this actually computes a dense subset of $\mathrm{Cts}(S^1, \mathbb{R})$, namely the trigonometric polynomials, presented in terms of the angle rather than cartesian coordinates.

<u>Before proving the Stone-Weierstrass theorem</u> we need some preliminary results.

<u>Lemma L16-6</u> If $X$ is locally compact Hausdorff the functions

$$\begin{aligned} \mathrm{Cts}(X,\mathbb{R}) \times \mathrm{Cts}(X,\mathbb{R}) &\longrightarrow \mathrm{Cts}(X,\mathbb{R}), & (f,g) &\longmapsto f+g \\ \mathrm{Cts}(X,\mathbb{R}) \times \mathrm{Cts}(X,\mathbb{R}) &\longrightarrow \mathrm{Cts}(X,\mathbb{R}), & (f,g) &\longmapsto fg \\ \mathbb{R} \times \mathrm{Cts}(X,\mathbb{R}) &\xrightarrow{a} \mathrm{Cts}(X,\mathbb{R}), & (\lambda,f) &\longmapsto \lambda f \end{aligned}$$

are <u>continuous</u>. We say $\mathrm{Cts}(X,\mathbb{R})$ is a <u>topological $\mathbb{R}$-algebra</u>, to emphasise this. In particular $\mathrm{Cts}(X,\mathbb{R})$ is a topological abelian group under addition.

<u>Proof</u> Consider the map ($\delta$ denotes an interchange $X_1 \times X_2 \cong X_2 \times X_1$)

$$\begin{tikzcd} X \times \mathrm{Cts}(X,\mathbb{R}) \times \mathrm{Cts}(X,\mathbb{R}) \arrow[r, "\Delta \times 1 \times 1"] & X \times X \times \mathrm{Cts}(X,\mathbb{R}) \times \mathrm{Cts}(X,\mathbb{R}) \arrow[d, "1 \times \delta \times 1"] \\ \mathbb{R} & \mathbb{R} \times \mathbb{R} \arrow[l, "+"'] & X \times \mathrm{Cts}(X,\mathbb{R}) \times X \times \mathrm{Cts}(X,\mathbb{R}) \arrow[l, "ev_{X,\mathbb{R}} \times ev_{X,\mathbb{R}}"'] \end{tikzcd}$$

which is continuous since $X$ is locally compact Hausdorff and hence $ev_{X,\mathbb{R}}$ is continuous. Corresponding to this is the continuous map

$$Cts(X,\mathbb{R}) \times Cts(X,\mathbb{R}) \longrightarrow Cts(X,\mathbb{R}) \qquad (f,g) \longmapsto f+g.$$

The other claims are handled similarly. $\square$

**Lemma L16-7** Let $X$ be locally compact Hausdorff and $A \subseteq Cts(X,\mathbb{R})$ a subalgebra. Then $\overline{A} \subseteq Cts(X,\mathbb{R})$ is also a subalgebra.

*Proof* Clearly $1 \in \overline{A}$, so we have to show $\overline{A}$ is closed under the operations $+, \cdot$ and scalar multiplication. Suppose $f,g \in \overline{A}$ but $f+g \notin \overline{A}$. Then there is $U \subseteq Cts(X,\mathbb{R})$ open with $f+g \in U$ and $U \cap A = \phi$. But then by Lemma L16-6

$$Q := \{ (a,b) \in Cts(X,\mathbb{R})^2 \mid a+b \in U \}$$

is open, and we may therefore find $C,D \subseteq Cts(X,\mathbb{R})$ open with $(f,g) \in C \times D \subseteq Q$. Since $f,g \in \overline{A}$ we have $C \cap A \neq \phi$ and $D \cap A \neq \phi$, say $f' \in C \cap A$ and $g' \in D \cap A$. Then $f'+g' \in A$ and

$$(f',g') \in C \times D \subseteq Q \implies f'+g' \in U$$

which contradicts $U \cap A = \phi$. Hence $f+g \in \overline{A}$. Similarly we show $fg \in \overline{A}$ and $\lambda f \in \overline{A}$ for any $\lambda \in \mathbb{R}$. $\square$

**Exercise L16-2½** Give an alternative proof of the Lemma in the case where $X$ is compact using the $d_\infty$ metric.

<u>Def$^n$</u> Let $X$ be a topological space and $f \in Cts(X, \mathbb{R})$. Then $|f| \in Cts(X, \mathbb{R})$ is the composite

$$X \xrightarrow{f} \mathbb{R} \xrightarrow{|-|} \mathbb{R}, \qquad x \longmapsto |f(x)|.$$

Given $f, g \in Cts(X, \mathbb{R})$ we define

$$\min\{f, g\} : X \longrightarrow \mathbb{R}, \qquad x \longmapsto \min\{f(x), g(x)\}$$
$$\max\{f, g\} : X \longrightarrow \mathbb{R}, \qquad x \longmapsto \max\{f(x), g(x)\}.$$

These functions are continuous since

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

<u>Exercise L16-4</u> Prove that if $X$ is locally compact Hausdorff then

$$|-| : Cts(X, \mathbb{R}) \longrightarrow Cts(X, \mathbb{R})$$
$$\min, \max : Cts(X, \mathbb{R}) \times Cts(X, \mathbb{R}) \longrightarrow Cts(X, \mathbb{R})$$

are all continuous functions.

The most difficult part of proving Stone-Weierstrass is proving that a <u>closed</u> subalgebra $A \subseteq Cts(X, \mathbb{R})$ has the property that $|A| \subseteq A$, i.e. if $f \in A$ then also $|f| \in A$. To prove this we will use that $|-|$ can be approximated by polynomials (so we use Weierstrass to prove Stone-Weierstrass).

**Lemma L16-8** Let $X$ be a compact space and $A \subseteq \mathrm{Cts}(X, \mathbb{R})$ a closed subalgebra. If $f, g \in A$ then $|f|, \min\{f,g\}, \max\{f,g\} \in A$.

*Proof* It clearly suffices to prove that $|f| \in A$. Given $f \in \mathrm{Cts}(X,\mathbb{R})$ we know $f$ is bounded, since $X$ is compact. Say $|f(x)| \le M$ for all $x \in X$. Then the function $|f|$ may be written as

$$X \xrightarrow{\ f\ } [-M, M] \xrightarrow{\ |-|\ } \mathbb{R}$$

Let $p_n \in \mathrm{Cts}([-M,M], \mathbb{R})$ be a sequence of polynomials converging to $|-|$ (this exists by Theorem L16-0). The function

$$\mathrm{Cts}([-M,M], \mathbb{R}) \xrightarrow{\ (-)\circ f\ } \mathrm{Cts}(X, \mathbb{R})$$

is continuous by Lemma L12-1, and since $p_n \longrightarrow |-|$ we have $p_n \circ f \longrightarrow |f|$ as $n \to \infty$. But if for some fixed $n$ we have $p_n = a_0 + a_1 t + \cdots + a_k t^k$ for constants $a_i \in \mathbb{R}$ then

$$p_n \circ f = a_0 + a_1 f + \cdots + a_k f^k$$

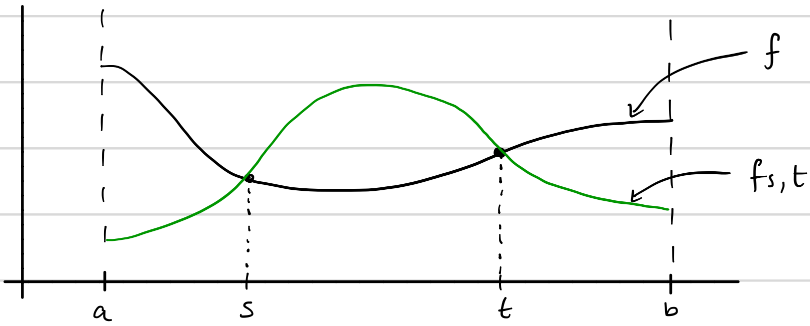
is an element of $A$. Hence $(p_n \circ f)_{n=0}^{\infty}$ is a sequence in $A$, and since $A$ is closed the limit $|f|$ also lies in $A$. $\square$

We are now prepared for the proof of the Stone-Weierstrass theorem. Our proof will use the Weierstrass theorem to prove the more general result. All the proofs of Stone-Weierstrass I am aware of hinge ultimately on a polynomial approximation of $|-|$, sometimes done "by hand" using a Taylor series of $\sqrt{1-t}$. This has its own complexities, and seems to me no easier than just proving the Weierstrass theorem.

<u>Proof of Theorem L16-3</u> Let $A \subseteq \mathrm{Cts}(X, \mathbb{R})$ be a subalgebra which separates points. Then by Lemma L16-7, $\bar{A}$ is also a subalgebra, and it clearly separates points since $A \subseteq \bar{A}$, so we may assume from the beginning that $A$ is <u>closed</u> and our goal is to show $A = \mathrm{Cts}(X, \mathbb{R})$.

Let $f \in \mathrm{Cts}(X, \mathbb{R})$ be given: we have to show $f \in A$. Given $\varepsilon > 0$ we will produce $g \in A$ such that $d_\infty(f, g) \leqslant \varepsilon$. This shows $f \in \bar{A} = A$. To produce $g$ we take distinct points $s, t \in X$ (if $X$ is empty or $X = \{*\}$ there is nothing to prove, as $\mathrm{Cts}(\{*\}, \mathbb{R}) \cong \mathbb{R}$ and any subalgebra contains the constants). We claim there exists $f_{s,t} \in A$ agreeing with $f$ on $\{s,t\}$, that is

$$f_{s,t}(s) = f(s), \qquad f_{s,t}(t) = f(t)$$




$X = [a,b]$

Since $A$ separates points there exists $h \in A$ such that $h(s) \neq h(t)$. Then we can just appropriately "massage" $h$ to produce $f_{s,t}$ with the desired property:

$$f_{s,t} := f(t) + \frac{f(s) - f(t)}{h(s) - h(t)} \left[ h - h(t) \right]$$

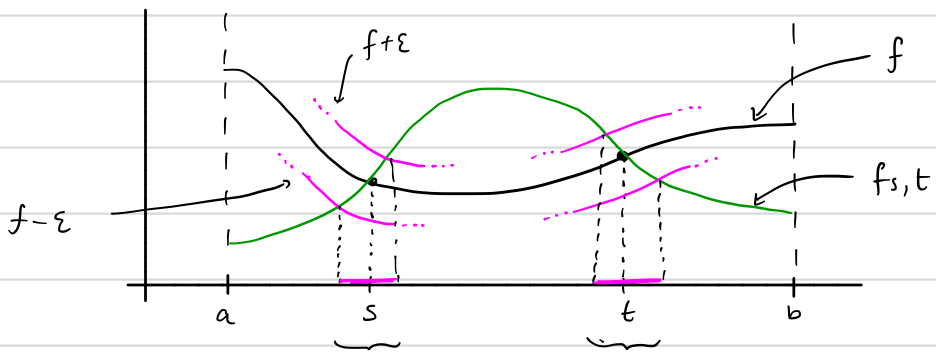
Moreover since $A$ is a subalgebra it is clear that $f_{s,t} \in A$. Now we construct $g$ from the collection $\{f_{s,t}\}_{s \neq t} \subseteq A$ (the construction involves for each $s, t$ choosing a $h$, but we don't care, any $f_{s,t} \in A$ agreeing with $f$ on $\{s,t\}$ will do).

The idea is to use the $f_{s,t}$ to construct the required approximation $g$ to $f$. Now, $f_{s,t}$ approximates $f$ only near $\{s,t\}$ (as far as we know) but

$$D_{s,t} = |f_{s,t} - f| : X \longrightarrow \mathbb{R}$$

is continuous, so the following set (where $f_{s,t}$ approximates $f$ sufficiently) is open:

$$U_{s,t} = D_{s,t}^{-1}((-\infty, \varepsilon)) = \{x \in X \mid f(x) - \varepsilon < f_{s,t}(x) < f(x) + \varepsilon\}$$




the set $U_{s,t}$ is the union of these two open intervals

We want to stitch $g$ together from the $f_{s,t}$ by <u>switching to a different pair</u> $(s', t')$ once we leave $U_{s,t}$, and we can use max, min to do the switching. But we have to be careful: in the context of the above picture, say $f_{s',t'} < f - \varepsilon$ on $U_{s,t}$, then $\min\{f_{s,t}, f_{s',t'}\}$ is <u>not</u> an approximation to $f$ on $U_{s,t} \cap U_{s',t'}$. The trick is to fix one of the points, say $s$, and compute instead $\min\{f_{s,t}, f_{s,t'}\}$ which <u>is</u> an approximation to $f$ near $s$, and is at least bounded above by $f + \varepsilon$ on $U_{s,t} \cup U_{s,t'}$. By compactness finitely such min's can arrange this to be the case on all of $X$ (still with $s$ fixed), so we'll have an approximation $h_s$ to $f$ near $s$ which is at least $< f + \varepsilon$ everywhere. But then we can take <u>max</u>'s of these $h_s$'s to impose a lower bound as well.

OK, so enough preamble, let's perform the construction.

For each $s \in X$, use compactness of $X$ to find $t_1, \ldots, t_r$ (depending on $s$) such that $U_{s,t_1}, \ldots, U_{s,t_r}$ cover $X$, and set

$$h_s := \min\{ f_{s,t_1}, \ldots, f_{s,t_r} \}.$$

By Lemma L16-8 (and hence ultimately by our polynomial approximation to $|\cdot|$) we have $h_s \in A$. Moreover $h_s(s) = f(s)$ and if $x \in X$ then $x \in U_{s,t_j}$ for some $j$ and hence

$$h_s(x) \leq f_{s,t_j}(x) < f(x) + \varepsilon$$

Also for $x$ in the open set $V_s = U_{s,t_1} \cap \cdots \cap U_{s,t_r}$ we have

$$h_s(x) = \min\{ f_{s,t_i}(x) \mid 1 \leq i \leq r \} > f(x) - \varepsilon .$$

The open sets $\{ V_s \}_{s \in X}$ cover $X$, and we may take a finite subcover $V_{s_1}, \ldots, V_{s_n}$. Then $g := \max\{ h_{s_1}, \ldots, h_{s_n} \}$ is by the same argument an element of $A$, and if $x \in X$ then

$$g(x) = \max\{ h_{s_1}(x), \ldots, h_{s_n}(x) \} < f(x) + \varepsilon$$

while there exists $1 \leq j \leq n$ with $x \in V_{s_j}$ and so

$$g(x) \geq h_{s_j}(x) > f(x) - \varepsilon$$

This shows that $d_\infty(g, f) \leq \varepsilon$ and completes the proof. $\square$

The construction of the approximating polynomials $B_n(f)$ in Weierstrass's theorem was explicit (although the $N$ we have to take s.t. $n \geq N$ ensures $d_\infty(B_n(f), f) < \varepsilon$ depends on $\delta$ which we may not be able to easily calculate). The Stone–Weierstrass theorem is less constructive, since it is not necessarily clear how to pick the finite subcovers involved, or how to choose the $h \in A$. However the other ingredients can be made constructive, in the way outlined by the following exercise:

Exercise L16-5 Let $X \subseteq \mathbb{R}^2$ be compact, $f: X \longrightarrow \mathbb{R}$ continuous, let $A = \mathrm{Poly}(X, \mathbb{R})$ and suppose $|f(x)| \leq M$ for all $x \in X$.

(i) Compute $B_n(|-|)$ on $[-M, M]$, as explained at the end of the proof of Theorem L16-0.

(ii) Set $s = (0,0)$ and $t_1 = (0,1)$, $t_2 = (0,2)$. Then $h(x,y) = y$ is a polynomial which separates both the pairs $(s, t_1)$ and $(s, t_2)$, and we may define $(\alpha = f(s), \beta = f(t_1), \gamma = f(t_2))$

$$f_{s,t_1}(x,y) := \beta - [\alpha - \beta](y-1)$$
$$f_{s,t_2}(x,y) := \gamma - \tfrac{1}{2}[\alpha - \gamma](y-2).$$

Compute using (i) a sequence of polynomial functions converging to $\min\{f_{s,t_1}, f_{s,t_2}\}$.

Def$^n$ A topological space is separable if it contains a countable dense subset.

Exercise L16-6 Prove that if $X \subseteq \mathbb{R}^n$ is compact then $C_b(X, \mathbb{R})$ is separable, and hence second-countable (this means that there is a basis $\mathcal{B}$ for the topology with $\mathcal{B}$ a countable set).

<u>Exercise L16-7</u> Recall from Ex. L12-11 that if $X$ is locally compact Hausdorff and $Y_1 \subseteq Y_2$ is a subspace then there is an embedding

$$\mathrm{Cts}(X, Y_1) \longrightarrow \mathrm{Cts}(X, Y_2)$$

given by post-composition with the inclusion $Y_1 \to Y_2$. We identify $\mathrm{Cts}(X, Y_1)$ with a subspace of $\mathrm{Cts}(X, Y_2)$ via this map. Prove

(i) If $X$ is compact and $Y_1 \subseteq Y_2$ is open, $\mathrm{Cts}(X,Y_1) \subseteq \mathrm{Cts}(X, Y_2)$ is open.
(ii) If $Y_1 \subseteq Y_2$ is closed, $\mathrm{Cts}(X, Y_1) \subseteq \mathrm{Cts}(X, Y_2)$ is closed.

We say a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is polynomial if each of the composites

$$\mathbb{R}^n \xrightarrow{f} \mathbb{R}^m \xrightarrow{\pi_i} \mathbb{R} \qquad 1 \le i \le m$$

is polynomial, and we write $\mathrm{Poly}(\mathbb{R}^n, \mathbb{R}^m) \subseteq \mathrm{Cts}(\mathbb{R}^n, \mathbb{R}^m)$ for the set of polynomial functions. If $j : X \to \mathbb{R}^n$ is an embedding then as above we define

$$\mathrm{Poly}(X, j, \mathbb{R}^m) := \{ f \circ j \in \mathrm{Cts}(X, \mathbb{R}^m) \mid f \text{ is polynomial} \}.$$

<u>Exercise L16-8</u> Prove that $\mathrm{Poly}(X, j, \mathbb{R}^m)$ is dense in $\mathrm{Cts}(X, \mathbb{R}^m)$.

<u>Exercise L16-9</u> Prove that for any space $X$ and $U \subseteq X$ open, $A \subseteq X$ dense that $U \cap A$ is a dense subset of $U$, with its subspace topology.

<u>Exercise L16-10</u> Prove that if $X \subseteq \mathbb{R}^n$ is compact and $Y \subseteq \mathbb{R}^m$ is open then the set of polynomial functions is dense in $\mathrm{Cts}(X, Y)$, where we call $f : X \to Y$ <u>polynomial</u> if $X \to Y \to \mathbb{R}^m$ is the restriction of a polynomial function.

quoted only for Exercise L16-11, not officially part of the course

<u>Theorem</u> (Urysohn lemma) Let $X$ be a normal space, $A, B$ disjoint closed subsets of $X$. Then there exists a continuous map $f : X \longrightarrow [0,1]$ such that $f(a) = 0$ for all $a \in A$ and $f(b) = 1$ for all $b \in B$.

<u>Exercise L16-11</u> Assuming the Urysohn lemma, prove that if $X, Y$ are compact Hausdorff spaces and $h : X \times Y \longrightarrow \mathbb{R}$ is continuous then for every $\varepsilon > 0$ there are continuous functions (for some $n$) $f_1, \ldots, f_n \in \mathrm{Cts}(X, \mathbb{R})$, $g_1, \ldots, g_n \in \mathrm{Cts}(Y, \mathbb{R})$ such that $d_\infty(h, \sum_i f_i g_i) < \varepsilon$, where given $f : X \to \mathbb{R}$ and $g : Y \to \mathbb{R}$ we write $fg$ for the function $(fg)(x,y) = f(x)g(y)$.

<u>Note</u> There is for $X$ locally compact Hausdorff a homeomorphism

$$\mathrm{Cts}(X, Y \times Z) \cong \mathrm{Cts}(X, Y) \times \mathrm{Cts}(X, Z).$$

It is <u>not</u> true that $\mathrm{Cts}(Y \times Z, X) \cong \mathrm{Cts}(Y, X) \times \mathrm{Cts}(Z, X)$ (what would a natural map relating LHS and RHS even be? It doesn't make sense). But if $X, Y$ are locally compact Hausdorff we have the continuous map

$$\begin{tikzcd}
X \times Y \times \mathrm{Cts}(X,\mathbb{R}) \times \mathrm{Cts}(Y,\mathbb{R}) \arrow[r, phantom, "\cong"] & (X \times \mathrm{Cts}(X,\mathbb{R})) \times (Y \times \mathrm{Cts}(Y,\mathbb{R})) \arrow[d, "ev_X \times ev_Y"] & \\
& \mathbb{R} \times \mathbb{R} \arrow[r, "\bullet"] & \mathbb{R}
\end{tikzcd}$$

associated to which is a continuous map (<u>not</u> injective if either $X \neq \phi$ or $Y \neq \phi$)

$$\Phi : \mathrm{Cts}(X, \mathbb{R}) \times \mathrm{Cts}(Y, \mathbb{R}) \longrightarrow \mathrm{Cts}(X \times Y, \mathbb{R})$$

The Exercise says: the subalgebra generated by the image of $\Phi$ is dense, if both $X, Y$ are compact (this is not the same as saying $\mathrm{Im}(\Phi)$ is dense).

**Exercise L16-12** Set $S^1 = \mathbb{R}/2\pi\mathbb{Z}$ and $\mathbb{T} = S^1 \times S^1$, with angular coordinates $(\theta, \psi)$. Give an appropriate class of trigonometric polynomials in $\mathrm{Cts}(\mathbb{T}, \mathbb{R})$ and prove that your set of polynomials is dense.

**Exercise L16-13** Let $X$ be locally compact Hausdorff, set $Y := X \amalg \{\infty\}$ (here $\infty$ denotes anything, $\infty = 0$ will do (although it looks nuts)) and define a topology on $Y$ as follows: the open subsets of $Y$ not containing $\infty$ are precisely the open subsets of $X$, and the open subsets of $Y$ containing $\infty$ are of the form $K^c \amalg \{\infty\}$ where $K \subseteq X$ is compact. The space $Y$ is called the <u>one-point compactification</u> of $X$:

(i) Prove $Y$ is compact Hausdorff and $X \longrightarrow Y$ is continuous

(ii) Prove that the one-point compactification of $\mathbb{R}$ is homeomorphic to $S^1$ (see Ex. L12-12).

The next exercise gives the generalisation of Stone–Weierstrass to locally compact spaces. We say $A \subseteq \mathrm{Cts}(X, \mathbb{R})$ is a <u>nonunital subalgebra</u> if whenever $f, g \in A$ we have $f+g$, $fg$, $\lambda f \in A$ for all $\lambda \in \mathbb{R}$ (but not necessarily $1 \in A$). If $X$ is locally compact Hausdorff we say $f: X \longrightarrow \mathbb{R}$ <u>vanishes at infinity</u> if

$$\forall \varepsilon > 0 \ \exists K \subseteq X \text{ compact } \forall x \notin K \left( |f(x)| < \varepsilon \right).$$

We write $\mathrm{Cts}_0(X, \mathbb{R}) \subseteq \mathrm{Cts}(X, \mathbb{R})$ for the subspace of functions vanishing at infinity.

**Exercise L16-14*** Suppose $X$ is locally compact Hausdorff and that $A$ is a nonunital subalgebra of $\mathrm{Cts}_0(X, \mathbb{R})$ which separates points and has the property that for every $x \in X$ there exists $f \in A$ with $f(x) \neq 0$. Then $\overline{A} = \mathrm{Cts}_0(X, \mathbb{R})$.

# Solutions to selected exercises

**L16-0** Suppose $f$ is continuous but not uniformly, so that for some $\varepsilon > 0$ no matter how small we make $\delta$, say $\delta = 1/n$, there exists a pair $x_n, y_n$ with $d_X(x_n, y_n) < 1/n$ but $d_Y(fx_n, fy_n) \geqslant \varepsilon$. Since $X$ is sequentially compact $(y_n)_{n=1}^{\infty}$ has a convergent subsequence $y_{n_k}$, with say $y_{n_k} \longrightarrow y$ as $k \to \infty$. We claim $x_{n_k} \longrightarrow y$ also, since

$$d_X(x_{n_k}, y) \leq d_X(x_{n_k}, y_{n_k}) + d_X(y_{n_k}, y)$$
$$< 1/n_k + d_X(y_{n_k}, y)$$

so given $\varepsilon' > 0$ let $K$ be s.t. $n_k \geqslant \frac{2}{\varepsilon'}$ if $k \geqslant K$ and $d_X(y_{n_k}, y) \leq \varepsilon'/2$ for $k \geqslant K$, then $d_X(x_{n_k}, y) < \varepsilon'/2 + \varepsilon'/2 = \varepsilon'$. But then since $f$ is continuous $fx_{n_k} \longrightarrow fy$ and $fy_{n_k} \longrightarrow fy$ as $k \to \infty$ and hence (again using a triangle inequality, or that $d_Y$ is continuous) we have $d_Y(fx_{n_k}, fy_{n_k}) \longrightarrow 0$ as $k \to \infty$. But this contradicts the lower bound $d_Y(fx_n, fy_n) \geqslant \varepsilon$. $\square$