ECON 600 — Class Notes

This week has some mistakes, especially before class 5. TB solved soon.

What is a Set

Definition (Set)**.** A *set* is a well-defined collection of objects (called *elements*). Sets are written using curly brackets, for example $A = \{a, b, c\}$.

Definition (Membership)**.** If x is an element of A, we write $x \in A$; otherwise $x \notin A$.

Example. Standard number sets

- **Integers:** $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$.
- **Rationals:** $\mathbb{Q} = \{x : x = \frac{g}{r},\ g, r \in \mathbb{Z},\ r \neq 0\}$.
- **Reals:** $\mathbb{R}$ (not defined here; we use its usual properties).
- **Positive reals:** $\mathbb{R}_{>0} = \{x \in \mathbb{R} : x > 0\}$ (also $\mathbb{R}_{\geq 0} = \{x \in \mathbb{R} : x \geq 0\}$).

Intervals in $\mathbb{R}$

Definition (Intervals)**.** Given $a < b$ in $\mathbb{R}$:

$$(a, b) = \{x \in \mathbb{R} : a < x < b\}, \qquad [a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}.$$

Elements and the Empty Set

Example. If $A = \{a, b, c\}$, then a is an element of A, that is $a \in A$.

Definition (Empty set)**.** The *empty set* is the set with no elements. It is denoted $\varnothing$.

Fundamental Property: Order and Repetition Do Not Matter

Remark. A set is determined only by which elements it contains; order and repetition do not change the set. For example

$$\{a, b, c\} = \{a, c, b\} = \{a, a, b, c\}.$$

Equality of Sets

Definition (Equality)**.** We say $X = Y$ if and only if, for every x, $x \in X$ *if and only if* $x \in Y$.

Proposition (Extensionality)**.** *If $X \subseteq Y$ and $Y \subseteq X$, then $X = Y$.*

Subsets

Definition (Subset)**.** X is a *subset* of Y (denoted $X \subseteq Y$) if, for every $x \in X$, we have $x \in Y$.

Example (Quick examples)**.**

$$\{a, b\} \subseteq \{a, b, c\}, \qquad \mathbb{Z} \subseteq \mathbb{R}, \qquad \mathbb{Q} \subseteq \mathbb{R}, \qquad \mathbb{R}_{>0} \subseteq \mathbb{R}.$$

Definition (Proper subset)**.** We say X is a *proper subset* of Y if $X \subseteq Y$ but $X \neq Y$. Notation: $X \subsetneq Y$.

Remark. For every set X, we have $\varnothing \subseteq X$ and $X \subseteq X$.

Power Set

Definition (Power set)**.** The *power set* of X is the set of *all* subsets of X. It is denoted 2^X.

Example. If $A = \{a, b, c\}$, then

$$2^A = \{\varnothing,\ \{a\},\ \{b\},\ \{c\},\ \{a, b\},\ \{a, c\},\ \{b, c\},\ A\}.$$

Remark. If A is finite with $|A| = n$, then $|2^A| = 2^n$ (useful for counting subsets).

Summary of Notation

- $x \in X$ / $x \notin X$: x belongs / does not belong to X.
- $X \subseteq Y$ / $X \subsetneq Y$: X is a subset (not necessarily proper) / a *proper* subset of Y.
- $\varnothing$: empty set. 2^X: power set of X.
- Intervals: (a, b) open, $[a, b]$ closed.

Useful Mini–Proofs

1) *$\varnothing \subseteq X$ for all X.* If there were an $x \in \varnothing$ with $x \notin X$, we would have an element of the empty set, which is impossible; thus "for all $x \in \varnothing$, $x \in X$" is vacuously true.
2) *Criterion for equality.* If $X \subseteq Y$ and $Y \subseteq X$, then by the definition of equality of sets it follows that $X = Y$.

Set Operations

Union. Given X and Y,

$$X \cup Y = \{x : \ x \in X \text{ or } x \in Y\}.$$

For an indexed family $(X_i)_{i \in I}$,

$$\bigcup_{i \in I} X_i = \{x : \ \exists i \in I \text{ such that } x \in X_i\}.$$

Intersection. Given X and Y,

$$X \cap Y = \{x : \ x \in X \text{ and } x \in Y\}.$$

For a family $(X_i)_{i \in I}$,

$$\bigcap_{i \in I} X_i = \{x : \ \forall i \in I,\ x \in X_i\}.$$

Difference and complement. For Y and X, the *difference* is

$$Y \setminus X = \{x : \ x \in Y \text{ and } x \notin X\}.$$

If we fix a *universe* U with $E \subseteq U$, we define the *complement* of E (in U) as $E^c = U \setminus E$.

Disjoint sets. We say that X and Y are *disjoint* if $X \cap Y = \varnothing$. For a family $(X_i)_{i\in I}$:

- **Pairwise disjoint:** $\forall i \neq j, \ X_i \cap X_j = \varnothing$.
- **Empty total intersection:** $\bigcap_{i\in I} X_i = \varnothing$ (this condition does *not* necessarily imply pairwise disjointness).

Example (Difference between "pairwise" and "total intersection")**.** Let $X_1 = \{0, 1, 2\}$, $X_2 = \{1, 3\}$ and $X_3 = \{3, 0\}$. Then

$$X_1 \cap X_2 = \{1\} \neq \varnothing, \quad X_1 \cap X_3 = \{0\} \neq \varnothing, \quad X_2 \cap X_3 = \{3\} \neq \varnothing,$$

but

$$\bigcap_{i=1}^{3} X_i = \varnothing.$$

The family is *not* pairwise disjoint, even though the total intersection is empty.

Partitions

Definition (Partition)**.** Let X be a set. A collection $\mathcal{P} \subseteq 2^X$ is a *partition* of X if:

(i) $E \neq \varnothing$ for every $E \in \mathcal{P}$ **(nonempty blocks)**;

(ii) If $E, F \in \mathcal{P}$ and $E \cap F \neq \varnothing$, then $E = F$ **(pairwise disjoint)**;

(iii) $\bigcup_{E\in\mathcal{P}} E = X$ **(covering)**.

The elements of $\mathcal{P}$ are called *classes* or *blocks*.

Remark. A partition decomposes X into disjoint pieces that cover it entirely. Each $x \in X$ belongs to a unique block.

Cartesian Product

Definition (Cartesian product of two sets)**.** For X and Y,

$$X \times Y = \{(x, y) : \ x \in X, \ y \in Y\}.$$

Order matters: in general $(x, y) \neq (y, x)$.

Definition (Cartesian product of a family)**.** Given a family $(X_i)_{i\in I}$,

$$\prod_{i\in I} X_i = \big\{(x_i)_{i\in I} : \ \forall i \in I, \ x_i \in X_i\big\}.$$

Example. Let $C = \{1, 2, \ldots, c\}$ and let $\mathbb{R}_{>0}$ be the positive reals. Then

$$C \times \mathbb{R}_{>0} = \{(c, w) : \ c \in C, \ w \in \mathbb{R}_{>0}\}.$$

Example. If $X_1 = \{1, 2\}$ and $X_2 = \{\alpha, \beta\}$, then

$$X_1 \times X_2 = \{(1, \alpha), (1, \beta), (2, \alpha), (2, \beta)\}.$$

Definition (n-fold product)**.** For a set X and $n \in \mathbb{N}$,

$$X^n = \underbrace{X \times X \times \cdots \times X}_{n \text{ factors}}.$$

Examples: $\mathbb{R}^2$, $\mathbb{R}^n$.

Relations

Definition (Binary relation)**.** Given sets X and Y, a *relation* from X to Y is a subset $R \subseteq X \times Y$. We write $x\,R\,y$ to indicate $(x, y) \in R$. Formally:

$$R = \{(x, y) \in X \times Y : \varphi(x, y)\}$$

where $\varphi(x, y)$ is a property/predicate that decides when (x,y) is "related."

Definition (Domain, image, and inverse relation)**.** For $R \subseteq X \times Y$,

$$\mathrm{Dom}(R) = \{x \in X : \ \exists y \in Y \text{ with } (x, y) \in R\}, \quad \mathrm{Im}(R) = \{y \in Y : \ \exists x \in X \text{ with } (x, y) \in R\},$$

and the *inverse relation* is $R^{-1} = \{(y, x) \in Y \times X : \ (x, y) \in R\}$.

Remark. When $X = Y$, we speak of a relation *on* X; in that case, properties such as *reflexivity*, *symmetry*, *antisymmetry*, and *transitivity* are often of interest (to be introduced in the next class).

Minimal set algebra (useful).

For $A, B, C \subseteq U$:

$$(A \cup B)^c = A^c \cap B^c, \qquad (A \cap B)^c = A^c \cup B^c, \qquad A \setminus B = A \cap B^c.$$

"if and only if" statements are proved by double inclusion.

Properties of Relations in $X \times X$

Definition (Relation)**.** A relation R between X and Y is a subset $R \subseteq X \times Y$. We write $x\,R\,y$ for $(x, y) \in R$. In what follows, $R \subseteq X \times X$.

Definition. Basic properties

- **Reflexive:** $\forall x \in X,\ xRx$.
- **Irreflexive:** $\forall x \in X,\ \neg(xRx)$.
- **Symmetric:** $xRy \Rightarrow yRx$.
- **Antisymmetric:** $(xRy \ \&\ yRx) \Rightarrow x = y$.
- **Asymmetric:** $xRy \Rightarrow \neg(yRx)$ (implies irreflexive and antisymmetric).
- **Transitive:** $(xRy \ \&\ yRz) \Rightarrow xRz$.
- **Complete (or connected):** $\forall x, y \in X$, xRy *or* yRx.

Example A: $\geq$ on $\mathbb{R}$.

Define $x \geq y \iff x$ is greater than or equal to y.

- Reflexive: $x \geq x$.
- Transitive: $x \geq y$ and $y \geq z \Rightarrow x \geq z$.
- Antisymmetric: $x \geq y$ and $y \geq x \Rightarrow x = y$.
- *Not* symmetric: $3 \geq 2$ but $2 \not\geq 3$.
- Complete: for all x, y, either $x \geq y$ or $y \geq x$.

Your margin note: "not symmetric" is correct.

Example B: $>$ on $\mathbb{R}$. Define $x > y \iff x$ is strictly greater than y.

- *Not* reflexive (indeed *irreflexive*): $x \not> x$.
- Transitive: if $x > y$ and $y > z$, then $x > z$.
- **Asymmetric** $\Rightarrow$ **antisymmetric** vacuously: if $x > y$, then *never* $y > x$ (no pairs go both directions).
- *Not* complete: when $x = y$, neither $x > y$ nor $y > x$.

Margin marks:

- "think a bit more": the statement "it is antisymmetric because there are no x, y with $x > y$ and $y > x$" is correct: the implication of antisymmetry holds vacuously (no counterexamples).
- "not complete because it's not reflexive": this is a good intuition and indeed a **theorem**:

Proposition (Completeness implies reflexivity)**.** *If R is complete, then it is reflexive.*

Proof. Fix $x \in X$. By completeness applied to the pair (x, x) we must have xRx. □

Remark. The *converse* is false: a relation may be reflexive and yet *not* complete (e.g. equality $=$ on X).

Example (product order on $\mathbb{R}^2$)**.** Define for $x = (x_1, x_2)$, $y = (y_1, y_2)$:

$$x \succeq y \iff (x_1 \geq y_1 \text{ and } x_2 \geq y_2).$$

Then $\succeq$ is **reflexive**, **transitive**, and **antisymmetric** (a *partial order*), but *not* complete.

Example. $(2, 1)$ and $(1, 10)$ are **incomparable**: neither $(2, 1) \succeq (1, 10)$ (since $1 \not\geq 10$) nor $(1, 10) \succeq (2, 1)$ (since $1 \not\leq 2$).

Equivalence and Indifference Relations

Definition (Equivalence)**.** A relation E on X is an *equivalence* if it is **reflexive**, **symmetric**, and **transitive**. Its equivalence classes form a partition of X.

Example D (identity). Let $X = \{a, b, c\}$ and $E = \{(a, a), (b, b), (c, c)\}$. Then E is **reflexive**, **symmetric**, and **transitive**. *On your orange note ("I don't understand transitivity"):* transitivity requires: if xEy and yEz, then xEz. Here the only possible chains are xEx and xEx (with $x = a, b$ or c), so xEx holds; there are no "mixed" chains, hence the condition is true. Moreover, E is *not* complete (e.g. neither aEb nor bEa).

Indifference induced by a weak preference. Let $\succeq$ be a **reflexive** and **transitive** relation on X (a weak preference). Define

$$x \sim y \iff (x \succeq y \text{ and } y \succeq x).$$

Proposition. *The relation $\sim$ is an* ***equivalence*** *(reflexive, symmetric, and transitive).*

Proof. Reflexive: $x \succeq x$ gives $x \sim x$. Symmetric: the definition is commutative. Transitive: if $x \sim y$ and $y \sim z$, then $x \succeq y \succeq z$ and $z \succeq y \succeq x$; by transitivity of $\succeq$, $x \succeq z$ and $z \succeq x$, hence $x \sim z$. □

Guided summary (connection with your notes).

- $>$: irreflexive, asymmetric, transitive, *not* complete. Antisymmetry holds vacuously.
- $\geq$: total order (reflexive, antisymmetric, transitive, complete), but not symmetric.
- Completeness $\Rightarrow$ reflexivity; *not* conversely (e.g. $=$ or product order on $\mathbb{R}^2$).
- Identity and "indifference" are equivalence relations; equivalence does not require completeness.

Functions

Function as a Special Relation

Definition (Function)**.** Let X and Y be sets. A *function* f from X to Y is a relation $G \subseteq X \times Y$ that satisfies:

(i) **Existence:** for every $x \in X$ there exists $y \in Y$ such that $(x, y) \in G$;

(ii) **Uniqueness:** if $(x, y) \in G$ and $(x, y') \in G$, then $y = y'$.

In that case we write $f : X \to Y$ and $f(x) = y$ when $(x, y) \in G$. The set G is called the *graph* of f and is denoted

$$\mathrm{Gr}(f) = \{(x, y) \in X \times Y : \; y = f(x)\}.$$

Remark (On your orange note: "Is this the same as reflexive?")**.** No. **Reflexivity** of a relation on X says $(x, x) \in R$ for every $x \in X$, that is, "each x is related to itself." Here condition (i) is the *existence of some image* y for each x, and it need not be the case that $y = x$.

Domains, Codomains, Image (and the Word "Range")

Definition. For $f : X \to Y$:

- X is the **domain**;
- Y is the **codomain**;
- $\mathrm{Im}(f) = f(X) = \{f(x) : x \in X\} \subseteq Y$ is the **image** (the values actually attained).

Remark (Your green note)**.** Some authors use "*range*" to mean the *codomain* and others to mean the *image*. To avoid ambiguity, in this course we adopt: *codomain* $= Y$, *image* $= f(X)$.

Two Baseline Examples from Class

Example (Finite domain)**.** Let $X = \{-9, 200\}$, $Y = \mathbb{R}$, and $f(-9) = 10$, $f(200) = -1$. Then

$$\mathrm{Gr}(f) = \{(-9, 10), (200, -1)\}, \qquad \mathrm{Im}(f) = \{10, -1\}.$$

Example (Increasing parabola on $X = \mathbb{R}_{>0}$)**.** Let $f : \mathbb{R}_{>0} \to \mathbb{R}$, $f(x) = x^2$. For each x there is a unique y ("*vertical line test*"): it is a function. It is **injective** on $\mathbb{R}_{>0}$ (strictly increasing), but it is *not* **surjective** onto $\mathbb{R}$ (it does not take negative values).

Example (Concave parabola $f(x) = ax - x^2$ on $X = \mathbb{R}_{>0}$)**.** The graph passes the vertical line test (it is a function), but it is typically *not* injective: for certain positive y there are two distinct x with $f(x) = y$ (your orange note "*there might be multiple x's with* $f(x) = y$").

Direct Image and Inverse Image

Definition (Image of a subset)**.** For $E \subseteq X$, the **image of** E under f is

$$f(E) = \{\, y \in Y : \; \exists x \in E \text{ with } y = f(x) \,\} \subseteq Y.$$

Definition (Inverse image of a subset)**.** For $F \subseteq Y$, the **inverse image** (or *preimage*) is

$$f^{-1}(F) = \{\, x \in X : \; f(x) \in F \,\} \subseteq X.$$

In particular, $f^{-1}(\{y\})$ (the *fiber* over y) is the set of all x that map to y.

Example (Your example 10b)**.** If $f(x) = ax - x^2$ and y^* lies below the maximum of the parabola, then

$$f^{-1}(\{y^*\}) = \{x_1, x_2\} \quad \text{with } x_1 \neq x_2.$$

This shows that the *inverse image* is a *set*; it is **not**, in general, a *functional inverse*.

Injectivity, Surjectivity, Bijectivity

Definition (Types of functions)**.** Let $f : X \to Y$.

- **Injective** (one-to-one): for all $x_1 \neq x_2$, $f(x_1) \neq f(x_2)$. Equivalent formulation: $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$. Equivalent in terms of fibers: $|f^{-1}(\{y\})| \leq 1$ for every y.
- **Surjective** (onto): $\mathrm{Im}(f) = Y$. Equivalent in terms of fibers: $f^{-1}(\{y\}) \neq \varnothing$ for every $y \in Y$.
- **Bijective**: both injective and surjective.

Inverse Function

Proposition (Characterization of invertibility)**.** *A function $f : X \to Y$ is **invertible** if and only if it is **bijective**. In that case there exists a unique function $g : Y \to X$ (the* inverse *of f) such that*

$$g(f(x)) = x \quad \text{for every } x \in X, \qquad \text{and} \qquad f(g(y)) = y \quad \text{for every } y \in Y.$$

Proof idea. If f is bijective, then for each $y \in Y$ there is exactly one $x \in X$ with $f(x) = y$; defining $g(y) = x$ satisfies the identities. Conversely, if there exists g with $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$, then each y equals $f(x)$ for a unique x, hence f is bijective. □

Notation and Graphical "Tests"

- $f : X \to Y$ denotes the function; $f(x)$ denotes its *value* at x.
- $\mathcal{F}(X, Y)$: the set of all functions from X to Y.
- **Vertical line test:** a curve in the plane describes a function $x \mapsto y$ if *every* vertical line intersects the curve in at most one point (your orange note "for each x there is a unique y").
- **Horizontal line test:** a function is injective if *every* horizontal line intersects the graph in at most one point (your orange note "there may exist multiple x with the same y" indicates failure of injectivity).

Types of Functions

Let $f : X \to Y$.

- **Injective** (one-to-one): if $x_1 \neq x_2$ then $f(x_1) \neq f(x_2)$.
- **Surjective** (onto): for every $y \in Y$ there exists $x \in X$ such that $f(x) = y$ (equivalently: $f(X) = Y$).
- **Bijective**: both injective and surjective (equivalently: invertible).

Basic Rules for Images and Preimages

Proposition (Monotonicity and inclusions)**.** *Let $f : X \to Y$. Then:*

(1) If $E \subseteq E' \subseteq X$, then $f(E) \subseteq f(E')$.

(2) If $F \subseteq F' \subseteq Y$, then $f^{-1}(F) \subseteq f^{-1}(F')$.

(3) For every $E \subseteq X$, one has $E \subseteq f^{-1}(f(E))$.

(4) For every $F \subseteq Y$, one has $f\big(f^{-1}(F)\big) \subseteq F$.

Proof. All proofs proceed by *element chasing.*

(1) Let $y \in f(E)$. By definition, there exists $x \in E$ with $f(x) = y$. Since $E \subseteq E'$, then $x \in E'$, hence $y = f(x) \in f(E')$. Therefore $f(E) \subseteq f(E')$.

(2) Let $x \in f^{-1}(F)$. By definition, $f(x) \in F$. Since $F \subseteq F'$, we also have $f(x) \in F'$, so $x \in f^{-1}(F')$. Therefore $f^{-1}(F) \subseteq f^{-1}(F')$.

(3) Let $x \in E$. Then $f(x) \in f(E)$ by definition of image. Hence $x \in f^{-1}(f(E))$. Thus $E \subseteq f^{-1}(f(E))$.

(4) Let $y \in f(f^{-1}(F))$. Then there exists $x \in f^{-1}(F)$ with $f(x) = y$. But $x \in f^{-1}(F)$ means $f(x) \in F$, that is, $y \in F$. Hence $f(f^{-1}(F)) \subseteq F$. □

Important remarks (when equalities may fail).

- **(3) may be strict if f is not injective.** Classic example: $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$. Take $E = \{1\}$. Then $f(E) = \{1\}$ and
$$f^{-1}(f(E)) = f^{-1}(\{1\}) = \{-1, 1\} \supsetneq E.$$
This matches your sketch of the concave parabola: a single y-value may correspond to multiple x's in the preimage.
- **(4) may be strict if F contains values outside $\mathrm{Im}(f)$.** Again with $f(x) = x^2$, take $F = (-1, 1)$. Then $f^{-1}(F) = (-1, 1)$ and
$$f\big(f^{-1}(F)\big) = f((-1, 1)) = [0, 1) \subsetneq (-1, 1) = F.$$
The intuition: f does not "reach" negative values, so it cannot recover all of F.
- **Useful characterizations.**
  a) f is **injective** $\iff$ for *every* $E \subseteq X$ one has $f^{-1}(f(E)) = E$.
  b) f is **surjective** $\iff$ for *every* $F \subseteq Y$ one has $f(f^{-1}(F)) = F$.

  Proofs: the forward direction follows from (3) and (4) plus the definition of injectivity/surjectivity; the reverse direction follows by taking $E = \{x\}$ in a) and $F = \{y\}$ in b).

Distribution of preimage and image over unions and intersections. For $F, G \subseteq Y$ and $E, H \subseteq X$:
$$f^{-1}(F \cup G) = f^{-1}(F) \cup f^{-1}(G), \qquad f^{-1}(F \cap G) = f^{-1}(F) \cap f^{-1}(G).$$
For images:
$$f(E \cup H) = f(E) \cup f(H), \qquad f(E \cap H) \subseteq f(E) \cap f(H) \quad \text{(equality may fail if } f \text{ is not injective).}$$

When (3) and (4) are Equalities

Corollary. Let $f : X \to Y$.

a) f is **injective** $\iff$ for every $E \subseteq X$ one has
$$f^{-1}(f(E)) = E.$$

b) f is **surjective** $\iff$ for every $F \subseteq Y$ one has
$$f\big(f^{-1}(F)\big) = F.$$

Counterexamples.

- **(3) may be strict if f is not injective.** With $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ and $E = \{1\}$:
$$f(E) = \{1\}, \qquad f^{-1}(f(E)) = \{-1, 1\} \supsetneq E.$$
- **(4) may be strict if f is not surjective.** With $f : \mathbb{R} \to \{a, b\}$ given by $f(x) = a$ for all $x \in \mathbb{R}$, and $F = \{a, b\} \subseteq Y$:
$$f^{-1}(F) = \mathbb{R}, \qquad f\big(f^{-1}(F)\big) = f(\mathbb{R}) = \{a\} \subsetneq \{a, b\} = F.$$

Remark. Here $f^{-1}(F)$ is the *inverse image* (preimage) of a set $F \subseteq Y$; it is **not** the "inverse function." Only when f is bijective does the inverse function $f^{-1} : Y \to X$ exist, and in that case $f^{-1}(F)$ coincides with the preimage via that inverse function.

Composition of Functions

Let $f : X \to Y$ and $g : Y \to Z$. We define the **composition**

$$g \circ f : X \to Z, \qquad (g \circ f)(x) = g\big(f(x)\big).$$

Type check: the codomain of f must coincide with the domain of g. Otherwise $f \circ g$ or $g \circ f$ "does not make sense."

Example (Order matters)**.** Let $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x_1, x_2) = 3x_1 + 2x_1x_2$, and $g : \mathbb{R} \to \mathbb{R}$, $g(t) = 3t$. Then $g \circ f : \mathbb{R}^2 \to \mathbb{R}$ is well defined and

$$(g \circ f)(x_1, x_2) = g\big(3x_1 + 2x_1x_2\big) = 3\,[\,3x_1 + 2x_1x_2\,].$$

In contrast, $f \circ g$ does *not* make sense because g returns a real while f expects a pair in $\mathbb{R}^2$.

Useful properties.

- **Associativity:** $h \circ (g \circ f) = (h \circ g) \circ f$ (when compositions are well typed).
- **Identity:** $\mathrm{id}_Y \circ f = f$ and $f \circ \mathrm{id}_X = f$.
- **Monotonicity of properties:**
  - If f and g are *injective*, then $g \circ f$ is injective.
  - If f and g are *surjective*, then $g \circ f$ is surjective.
  - If f and g are *bijective*, then $g \circ f$ is bijective; its inverse is $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Correspondences (Multivalued Maps)

Definition (Correspondence)**.** A *correspondence* (or multivalued map) from X to Y is a rule Γ that assigns to each $x \in X$ a nonempty set $\Gamma(x) \subseteq Y$. We can view it as a function

$$\Gamma : X \longrightarrow 2^Y \setminus \{\varnothing\}, \qquad x \mapsto \Gamma(x).$$

Its **graph** is $\mathrm{Gr}(\Gamma) = \{(x, y) \in X \times Y : \; y \in \Gamma(x)\}$.

Remark. A correspondence **is** a function, but its *codomain* is not Y but the *power set* $2^Y \setminus \{\varnothing\}$. It does not require uniqueness of the value: it allows multiple "outputs" for each x.

Example (Your discrete sets)**.** Let $X = \{a, b\}$ and $Y = \{x, y, z, q, r\}$. Define

$$\Gamma(a) = \{x, y, z\}, \qquad \Gamma(b) = \{y, q, r\}.$$

Then $\Gamma : X \to 2^Y \setminus \{\varnothing\}$ is a correspondence with

$$\mathrm{Gr}(\Gamma) = \{(a, x), (a, y), (a, z), (b, y), (b, q), (b, r)\}.$$

Image and preimage for correspondences. For $E \subseteq X$ and $F \subseteq Y$:

$$\Gamma(E) = \bigcup_{x \in E} \Gamma(x) \subseteq Y, \qquad \Gamma^{-1}(F) = \{\, x \in X : \; \Gamma(x) \cap F \neq \varnothing \,\} \subseteq X.$$

(Note: in the preimage we require "*nonempty intersection*" with F.)

Definition (Selections)**.** A *selection* of Γ is a function $s : X \to Y$ such that $s(x) \in \Gamma(x)$ for all x. (Useful when one wants to "pick" an element from each set of the correspondence.)

Real Numbers and Bounds

Upper and Lower Bounds

Let $X \subseteq \mathbb{R}$ be a nonempty set.

Definition. Lower and upper bounds:

- $u \in \mathbb{R}$ is an **upper bound** of X if $\forall x \in X$, $x \leq u$.
- $\ell \in \mathbb{R}$ is a **lower bound** of X if $\forall x \in X$, $\ell \leq x$.

Definition (Bounded)**.** We say that X is **bounded above** if it has some upper bound; **bounded below** if it has some lower bound; and **bounded** if it has both.

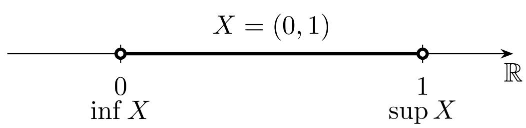
Remark. A set may have neither an upper nor a lower bound (for example, $\mathbb{Z} \subset \mathbb{R}$ has no upper bound).

Supremum and Infimum

Definition (Supremum and infimum)**.** Let $X \subseteq \mathbb{R}$ be nonempty.

- The **supremum** of X, denoted $\sup X$, is the *least* upper bound of X: it is an upper bound and satisfies $\sup X \leq u$ for every upper bound u.
- The **infimum** of X, denoted $\inf X$, is the *greatest* lower bound of X: it is a lower bound and satisfies $\ell \leq \inf X$ for every lower bound ℓ.

Remark. $\sup X$ and $\inf X$ need *not* belong to X. For example, if $X = (0, 1)$, then $\inf X = 0 \notin X$ and $\sup X = 1 \notin X$.



Axiom of Completeness of $\mathbb{R}$

Definition (Axiom of completeness)**.** Every nonempty subset $X \subseteq \mathbb{R}$ that is bounded above has a **supremum** in $\mathbb{R}$; and every nonempty subset that is bounded below has an **infimum** in $\mathbb{R}$.

Remark. This *fails* in $\mathbb{Q}$. For example, $X = \{q \in \mathbb{Q} : q^2 < 2\}$ is nonempty and bounded above, but $\sup X = \sqrt{2} \notin \mathbb{Q}$. This is why we work in $\mathbb{R}$ when using suprema/infima without further hypotheses.

Remark (Convention)**.** If X has no upper bound, we write $\sup X = +\infty$; if it has no lower bound, $\inf X = -\infty$. (These are *conventional* values useful for stating results.)

Basic Properties of the Supremum

Proposition. *Let $X \subseteq \mathbb{R}$ be nonempty and bounded above. Then:*

(i) $\sup X$ *is* ***unique****.*

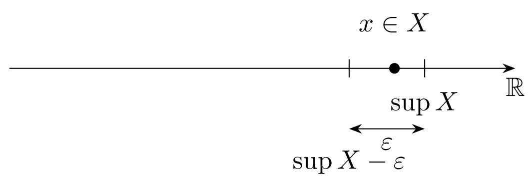
(ii) *(*Approximation by elements of X*) For every $\varepsilon > 0$ there exists $x \in X$ such that*

$$\sup X - \varepsilon < x \leq \sup X.$$

Dually, if X is bounded below, then for every $\varepsilon > 0$ there exists $x \in X$ with $\inf X \leq x < \inf X + \varepsilon$*.*

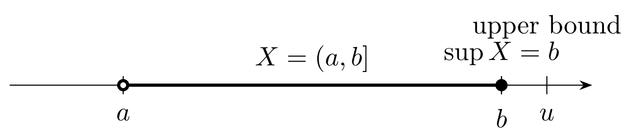
Proof. (i) If u_1 and u_2 are suprema, then $u_1 \le u_2$ (since u_2 is an upper bound) and $u_2 \le u_1$ (since u_1 is an upper bound). Hence $u_1 = u_2$.

(ii) Let $\varepsilon > 0$. Suppose there is *no* $x \in X$ with $\sup X - \varepsilon < x$. Equivalently, $\forall x \in X$, $x \le \sup X - \varepsilon$. But then $\sup X - \varepsilon$ would be an upper bound of X *smaller* than $\sup X$, contradicting minimality of $\sup X$. Therefore, there exists $x \in X$ with $\sup X - \varepsilon < x$. Since $\sup X$ is an upper bound, we also have $x \le \sup X$. ☐

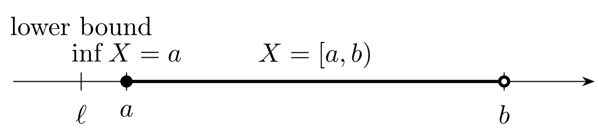


Two Conceptual Diagrams

Upper bounds and supremum.



Lower bounds and infimum (dual).



Alternative Proofs for Prop. (ii)

Recall: if $X \neq \varnothing$ is bounded above, then for every $\varepsilon > 0$ there exists $x \in X$ such that $\sup X - \varepsilon < x \le \sup X$.

Proof. (a) Proof by contradiction. Suppose there exists $\varepsilon > 0$ such that $\forall x \in X$, $x \le \sup X - \varepsilon$. Then $\sup X - \varepsilon$ is an *upper bound* of X and is strictly smaller than $\sup X$, contradicting the minimality of $\sup X$. ☐

Proof. (b) Proof by contrapositive. Let u be an upper bound of X. If there exists $\varepsilon > 0$ such that $u - \varepsilon \ge x$ for all $x \in X$, then $u - \varepsilon$ would also be an upper bound, hence u could not be the *least* upper bound. By contrapositive, if $u = \sup X$, then for every $\varepsilon > 0$ there must exist $x \in X$ with $x > u - \varepsilon$. ☐

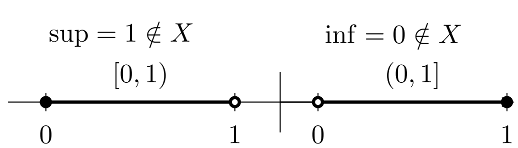
Maximum and Minimum

Definition (Maximum and minimum)**.** Let $X \subseteq \mathbb{R}$ be nonempty.

- $\bar{x}$ is a **maximum** of X if $\bar{x} \in X$ and $x \le \bar{x}$ for every $x \in X$.
- $\underline{x}$ is a **minimum** of X if $\underline{x} \in X$ and $\underline{x} \le x$ for every $x \in X$.

Remark. There may be bounds without a maximum/minimum. Examples:

$$X = [0, 1) \quad \text{(no maximum)}, \qquad X = (0, 1] \quad \text{(no minimum)}.$$



Proposition (Maximum/minimum vs. supremum/infimum)**.** *Let $X \subseteq \mathbb{R}$ be nonempty.*

*(i) $\bar{x}$ is a **maximum** of $X \iff \bar{x} = \sup X$ and $\bar{x} \in X$.*

*(ii) $\underline{x}$ is a **minimum** of $X \iff \underline{x} = \inf X$ and $\underline{x} \in X$.*

Proof. (i) If $\bar{x}$ is a maximum, then it is an upper bound and belongs to X; by minimality of the supremum, $\sup X \leq \bar{x}$, and since $\bar{x}$ is already a bound, $\sup X = \bar{x}$. Conversely, if $\bar{x} = \sup X \in X$, then no $x \in X$ exceeds $\bar{x}$; hence it is a maximum. The case (ii) is dual. □

Corollary. X has a **maximum** $\iff \sup X \in X$ (in that case, it is unique). Similarly, X has a **minimum** $\iff \inf X \in X$.

Metric Spaces

Definition (Metric)**.** Let X be a set. A **metric** on X is a function $d : X \times X \to \mathbb{R}$ such that, for all $x, y, z \in X$:

(a) $d(x, y) \geq 0$ (nonnegativity);

(b) $d(x, y) = 0 \iff x = y$ (identity of indiscernibles);

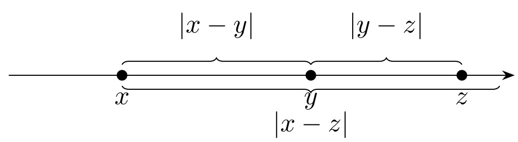
(c) $d(x, y) = d(y, x)$ (symmetry);

(d) $d(x, z) \leq d(x, y) + d(y, z)$ (triangle inequality).

Definition (Metric space)**.** A **metric space** is a pair (X, d) with $X \neq \varnothing$ and d a metric on X.

Example 1: usual metric on $\mathbb{R}$. For $X \subseteq \mathbb{R}$, $d_u(x, y) = |x - y|$. Properties (a)–(c) follow from properties of the absolute value. For (d):

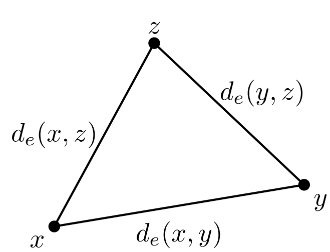
$$|x - z| = |(x - y) + (y - z)| \leq |x - y| + |y - z|.$$



Example 2: Euclidean metric on $\mathbb{R}^n$. For $X \subseteq \mathbb{R}^n$, $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$,

$$d_e(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

When $n = 1$, $d_e = d_u$.



Example 3: Discrete Metric

Let X be a nonempty set. Define $d : X \times X \to \mathbb{R}$ by

$$d(x,y) = \begin{cases} 1, & x \neq y, \\ 0, & x = y. \end{cases}$$

It is immediate to verify (nonnegativity, identity, symmetry). For the triangle inequality: if $x \neq z$ then $d(x,z) = 1 \leq d(x,y) + d(y,z)$ since the right-hand side is $0, 1$, or 2; if $x = z$, both sides equal 0. Therefore, d is a metric (the *discrete metric*).

Bounded Real Functions

Definition (Bounded functions)**.** Fix $X \neq \varnothing$. A real function $f : X \to \mathbb{R}$ is **bounded** if $f(X) \subseteq \mathbb{R}$ is a bounded set. We denote

$$\mathcal{F}^B(X,\mathbb{R}) = \{\, f : X \to \mathbb{R} \,:\, f \text{ is bounded} \,\}.$$

Remark. If f or g are not bounded, the set $\{|f(x) - g(x)| : x \in X\}$ may fail to have a finite upper bound (and its "supremum" would be $+\infty$), which would prevent defining a distance with values in $\mathbb{R}$. That is why we restrict the domain to $\mathcal{F}^B(X,\mathbb{R})$.

Supremum (Uniform) Metric on $\mathcal{F}^B(X,\mathbb{R})$

Definition (Supremum metric)**.** Let $X \neq \varnothing$. Define $d_\infty : \mathcal{F}^B(X,\mathbb{R}) \times \mathcal{F}^B(X,\mathbb{R}) \to \mathbb{R}$ by

$$d_\infty(f,g) = \sup_{x \in X} |f(x) - g(x)|.$$

Note that $d_\infty(f,g) = \|f - g\|_\infty$, where $\|h\|_\infty := \sup_{x \in X} |h(x)|$.

Proposition. *d_∞ is a metric on $\mathcal{F}^B(X,\mathbb{R})$.*

Proof. Let $f, g, h \in \mathcal{F}^B(X,\mathbb{R})$.

(a) **Nonnegativity:** $|f(x) - g(x)| \geq 0$ for all x, so the supremum is ≥ 0.

(b) **Identity of indiscernibles:** $d_\infty(f,g) = 0$ implies $\sup_x |f(x) - g(x)| = 0$, hence $|f(x) - g(x)| = 0$ for all x, i.e. $f = g$. The converse is clear.

(c) **Symmetry:** $|f(x) - g(x)| = |g(x) - f(x)|$ for all x, so the suprema coincide.

(d) **Triangle inequality:** by the usual triangle inequality in $\mathbb{R}$, $|f(x)-h(x)| \leq |f(x)-g(x)|+|g(x)-h(x)|$ for all $x \in X$. Taking suprema on both sides and using the fact

$$\big(\forall x,\ a_x \leq b_x\big) \implies \sup_x a_x \leq \sup_x b_x,$$

we obtain

$$\sup_x |f(x) - h(x)| \leq \sup_x |f(x) - g(x)| + \sup_x |g(x) - h(x)|.$$

That is, $d_\infty(f,h) \leq d_\infty(f,g) + d_\infty(g,h)$.

$\square$

Lemma. *If for every x one has $A(x) \leq B(x)$, then* $\sup_x A(x) \leq \sup_x B(x)$. Reason: *every supremum is the* ***least upper bound****; since* $\sup_x B(x)$ *bounds each $A(x)$, it also bounds their set of values and, by minimality, dominates the supremum of A.*

Remark (When the supremum in d_∞ exists)**.** If X is compact and f, g are continuous, then $x \mapsto |f(x)-g(x)|$ attains a maximum (extreme value theorem), and $d_\infty(f,g)$ exists without needing to assume that f, g are globally bounded in $\mathbb{R}$. In general, it suffices that $f - g$ be bounded.

Sequences: Boundedness, Monotonicity, and Basic Theorems

Definition (Bounded sequence in a metric space)**.** Let (X, d) be a metric space and $(x_n)_{n\in\mathbb{N}}$ a sequence in X. We say that (x_n) is *bounded* if there exists some $y \in X$ and $M > 0$ such that

$$d(x_n, y) \le M \quad \text{for all } n \in \mathbb{N}.$$

In particular, in $(\mathbb{R}, d_u)$ with $d_u(x, y) = |x - y|$, a real sequence (x_n) is bounded if and only if

$$\exists\, M > 0 \text{ such that } |x_n| \le M \quad \forall n.$$

Indeed, if $|x_n - y| \le M$ for all n (with some $y \in \mathbb{R}$), then

$$|x_n| \le |x_n - y| + |y| \le M + |y| \quad \forall n,$$

and therefore it is bounded taking $\widehat{M} := M + |y|$.

Remark (Boundedness of the set of values)**.** For a real sequence (x_n), the set of its values $E := \{x_n : n \ge 1\}$ is bounded if there exist $a, b \in \mathbb{R}$ with

$$a \le x_n \le b \quad \forall n.$$

Then $|x_n| \le \max\{|a|, |b|\}$ for all n.

Proposition. *Let (X, d) be a metric space. If (x_n) converges in X, then (x_n) is bounded.*

Proof. Let $x \in X$ be the limit of (x_n). Taking $\varepsilon = 1$, there exists N such that $d(x_n, x) \le 1$ for all $n \ge N$. Define

$$M := \max\big\{\, 1,\ d(x_1, x), \dots, d(x_{N-1}, x) \,\big\}.$$

Then $d(x_n, x) \le M$ for all n, and by definition (x_n) is bounded. □

Example. The sequence $((-1)^n)_{n\in\mathbb{N}} = (-1, 1, -1, 1, \dots)$ is bounded in $(\mathbb{R}, d_u)$ (since $|(-1)^n| \le 1$) but does not converge.

Definition (Monotone sequences)**.** A real sequence (x_n) is

- *increasing* if $x_{n+1} \ge x_n$ for all n;
- *decreasing* if $x_{n+1} \le x_n$ for all n;
- *monotone* if it is either increasing or decreasing.

Proposition (Monotone convergence theorem)**.** *Let (x_n) be a monotone and bounded real sequence. Then (x_n) converges. Moreover:*

$$\textit{if } (x_n) \textit{ is increasing,} \qquad \lim_{n\to\infty} x_n = \sup\{x_n : n \in \mathbb{N}\};$$
$$\textit{if } (x_n) \textit{ is decreasing,} \qquad \lim_{n\to\infty} x_n = \inf\{x_n : n \in \mathbb{N}\}.$$

Proof (increasing case). Let $S := \{x_n : n \in \mathbb{N}\}$; by boundedness, S has an upper bound, and by completeness of $\mathbb{R}$ there exists $x^* := \sup S$. Let $\varepsilon > 0$. By the definition of supremum, there exists N such that $x_N > x^* - \varepsilon$. Since (x_n) is increasing, for every $n \ge N$ we have $x_n \ge x_N > x^* - \varepsilon$ and, since x^* is an upper bound, $x_n \le x^*$. Hence $0 \le x^* - x_n < \varepsilon$, i.e. $|x_n - x^*| < \varepsilon$ for all $n \ge N$. Therefore $x_n \to x^*$. The decreasing case is analogous replacing sup by inf. □

Remark. Proposition , together with the boundedness criterion of Proposition , is useful to *verify that a sequence does not converge*: if it is monotone but not bounded, it cannot converge.

Lemma (Monotone subsequence)**.** *Every sequence (x_n) in $(\mathbb{R}, d_u)$ has a monotone subsequence.*

Corollary (Bolzano–Weierstrass)**.** Every bounded sequence in $(\mathbb{R}, d_u)$ has a convergent subsequence.

Proposition (Squeeze theorem)**.** *Let* (x_n)*,* (y_n)*,* (z_n) *be real sequences such that, for all* n*,*

$$x_n \le y_n \le z_n, \qquad \lim_{n\to\infty} x_n = \lim_{n\to\infty} z_n = a \in \mathbb{R}.$$

Then $\lim_{n\to\infty} y_n = a$*.*

Proposition (Algebra of limits)**.** *Let* (x_n) *and* (y_n) *be real sequences with* $\lim_{n\to\infty} x_n = x \in \mathbb{R}$ *and* $\lim_{n\to\infty} y_n = y \in \mathbb{R}$*. Then:*

1. $\lim_{n\to\infty} |x_n| = |x|$*;*
2. $\lim_{n\to\infty} (x_n + y_n) = x + y$*;*
3. $\lim_{n\to\infty} (x_n\, y_n) = x\, y$*;*
4. $\lim_{n\to\infty} (x_n / y_n) = x/y$ *when* $y \neq 0$ *and* $y_n \neq 0$*.*

Example (Example 4)**.** Consider the sequence

$$x_n \;=\; \frac{3\big(1-\frac{1}{n}\big)^2}{\dfrac{2(n+1)}{n}} \;+\; \Big(\lambda + \frac{1}{n}\Big).$$

Since $\big(1-\frac{1}{n}\big)^2 \to 1$ and $\frac{2(n+1)}{n} \to 2$, the first term tends to $\frac{3}{2}$; moreover $\lambda + \frac{1}{n} \to \lambda$. Therefore

$$\lim_{n\to\infty} x_n = \lambda + \tfrac{3}{2}.$$

Application of Limits and Bounds: O and o Notation

Let (x_n) and (y_n) be two real sequences with $y_n \neq 0$ for all n.

Definition (Same order / Big-O)**.** We say that (x_n) is of the *order* of (y_n) if the quotient $\left(\frac{x_n}{y_n}\right)$ is bounded, that is, there exists $\Pi > 0$ such that

$$\left|\frac{x_n}{y_n}\right| \le \Pi \quad \text{for all } n,$$

equivalently,

$$|x_n| \le \Pi\, |y_n| \quad \forall n.$$

In this case we write $x_n = O(y_n)$. As a particular case, if $y_n \equiv 1$, then $x_n = O(1)$ means that (x_n) is bounded.

Definition (Lower order / little-o)**.** We say that (x_n) is of *lower order* than (y_n) if

$$\lim_{n\to\infty} \frac{x_n}{y_n} = 0,$$

and we write $x_n = o(y_n)$. In particular, if $y_n \equiv 1$, then $x_n = o(1)$ is equivalent to $x_n \to 0$.

Proposition (Basic properties of O and o)**.** *Let* $(x_n), (\widehat{x}_n), (y_n), (\widehat{y}_n)$ *be real sequences.*

1. *If* $x_n = O(y_n)$ *and* $\widehat{x}_n = O(\widehat{y}_n)$*, then*
$$x_n\, \widehat{x}_n = O\big(y_n\, \widehat{y}_n\big).$$
   Proof: $|x_n| \le \Pi |y_n|$ *and* $|\widehat{x}_n| \le \widehat{\Pi} |\widehat{y}_n| \Rightarrow |x_n \widehat{x}_n| \le (\Pi\widehat{\Pi}) |y_n \widehat{y}_n|$.
2. *If* $k \neq 0$ *and* $x_n = O(y_n)$*, then* $k\, x_n = O(|k|\, y_n)$*.*

3. *If $x_n = O(y_n)$ and $\widehat{x}_n = O(\widehat{y}_n)$, then*

$$x_n + \widehat{x}_n \;=\; O\big(\max\{|y_n|, |\widehat{y}_n|\}\big).$$

In particular, $|x_n + \widehat{x}_n| \le (\Pi + \widehat{\Pi}) \max\{|y_n|, |\widehat{y}_n|\} \le 2\max\{\Pi, \widehat{\Pi}\} \max\{|y_n|, |\widehat{y}_n|\}$.

4. *If $x_n = O(y_n)$ and also $\widehat{x}_n = O(y_n)$, then $x_n + \widehat{x}_n = O(y_n)$. (This is a particular case of (3).)*
5. *If $k \neq 0$ and $x_n = o(y_n)$, then $k\,x_n = o(y_n)$.*
6. *If $x_n = o(y_n)$ and $\widehat{x}_n = o(\widehat{y}_n)$, then*

$$x_n\,\widehat{x}_n \;=\; o\big(y_n\,\widehat{y}_n\big),$$

since $\dfrac{x_n\widehat{x}_n}{y_n\widehat{y}_n} = \dfrac{x_n}{y_n} \cdot \dfrac{\widehat{x}_n}{\widehat{y}_n} \to 0 \cdot 0 = 0.$

7. *If $x_n = o(y_n)$ and $\widehat{x}_n = o(y_n)$, then $x_n + \widehat{x}_n = o(y_n)$, because* $\dfrac{x_n + \widehat{x}_n}{y_n} = \dfrac{x_n}{y_n} + \dfrac{\widehat{x}_n}{y_n} \to 0 + 0 = 0.$

Cauchy and a Clarification on Boundedness

Initial clarification

Given a metric space (X, d) and a sequence $(x_n)_{n\in\mathbb{N}} \subset X$:

- $(x_1, x_2, x_3, \dots)$ is a *sequence*; in contrast, $\{x_1, x_2, x_3, \dots\}$ is the *set* of values the sequence may produce.
- We say that (x_n) is *bounded* if there exists $y \in X$ and $M > 0$ such that $d(x_n, y) \le M$ for all $n \in \mathbb{N}$. For a fixed $y \in X$, let

  $$E_y = \{\, d(x_n, y) : n \in \mathbb{N} \,\} \subset \mathbb{R}.$$

  Then (x_n) is bounded $\iff$ E_y is bounded above. Moreover, if $d(x_n, y) \le M$ for all n and $y' \in X$ is arbitrary, by the triangle inequality

  $$d(x_n, y') \le d(x_n, y) + d(y, y') \le M + d(y, y'),$$

  so the notion of boundedness does not depend on the chosen center.
- Trivial case: if the set of values $\{x_n : n \in \mathbb{N}\}$ is finite, then (x_n) is bounded. The case that requires attention is when *infinitely many* distinct values appear.
- **Observation.** An unbounded sequence does *not necessarily* contain a bounded subsequence (e.g. $x_n = n$ in $\mathbb{R}$ has none).

Remark. We will say that a subsequence is *proper* when it does not coincide with the original sequence.

Cauchy Sequences

Definition (Cauchy sequence)**.** Let (X, d) be a metric space and $(x_n)_{n\in\mathbb{N}} \subset X$. We say that (x_n) is *Cauchy* if for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $m, n \ge N$ we have

$$d(x_n, x_m) < \varepsilon.$$

Example (1b)**.** In $(\mathbb{R}, d_u)$ (the usual metric), the sequence $x_n = \frac{1}{n}$ is Cauchy. Indeed, given $\varepsilon > 0$, choose N with $\frac{1}{N} < \varepsilon$. If $m, n \ge N$, then

$$|x_n - x_m| = \Big|\frac{1}{n} - \frac{1}{m}\Big| \le \frac{1}{N} < \varepsilon.$$

(Note that in $(0, 1)$ the same sequence does *not* converge because its limit $0 \notin (0, 1)$.)

Example (3)**.** The sequence $x_n = (-1)^n$ in $(\mathbb{R}, d_u)$ is not Cauchy. Take $\varepsilon \in (0,1)$. For every N we have

$$|x_{N+1} - x_N| = \left|(-1)^{N+1} - (-1)^N\right| = 2 > \varepsilon,$$

hence the Cauchy condition fails.

Proposition (7)**.** *Every convergent sequence is Cauchy.*

Proof. Let (x_n) be a sequence in (X,d) that converges to $x \in X$. Given $\varepsilon > 0$, by convergence there exists N such that $d(x_n, x) < \varepsilon/2$ for all $n \geq N$. If $m, n \geq N$, by the triangle inequality,

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

□

Remark. The converse does not always hold in an arbitrary metric space: a sequence may be Cauchy and not converge if the space is not complete (e.g. $1/n$ in $(0,1)$).

Proposition (Convergent ⇒ Cauchy)**.** *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a metric space (X,d) that converges to $x \in X$. Then (x_n) is Cauchy.*

Proof. Fix $\varepsilon > 0$. By convergence, there exists N such that $d(x_n, x) < \varepsilon/2$ for all $n \geq N$. If $m, n \geq N$, then by the triangle inequality,

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

□

Proposition (Every Cauchy sequence is bounded)**.** *Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in (X,d). Then there exist $y \in X$ and $M > 0$ such that $d(x_n, y) \leq M$ for all n (that is, (x_n) is bounded).*

Proof. Take $\varepsilon = 1$. Since (x_n) is Cauchy, there exists N such that $d(x_n, x_m) \leq 1$ for all $m, n \geq N$. Fix $y := x_N$. Then $d(x_n, y) \leq 1$ for all $n \geq N$. For the finitely many indices $1 \leq n < N$, define

$$K := \max\{\, d(x_1, y),\ d(x_2, y),\ \ldots,\ d(x_{N-1}, y)\,\}.$$

Let $M := \max\{K, 1\}$. Then $d(x_n, y) \leq M$ for all $n \in \mathbb{N}$. □

Definition (Complete space)**.** We say that a metric space (X,d) is *complete* if every Cauchy sequence in (X,d) converges (to a point in X).

Example.

- $(\mathbb{R}^n, d_{\text{eucl}})$ is complete.
- Every finite set with the discrete metric is complete.
- If $\mathcal{F}^B(X,\mathbb{R})$ denotes the space of *bounded* functions $f : X \to \mathbb{R}$ and $d_\infty(f,g) := \sup_{x\in X} |f(x) - g(x)|$, then $(\mathcal{F}^B(X,\mathbb{R}), d_\infty)$ is complete.

Example (Incomplete spaces)**.**

- $((0,1), d_u)$ is not complete (e.g. $x_n = \frac{1}{n}$ is Cauchy but does not converge in $(0,1)$).
- $(\mathbb{Q}, d_u)$ is not complete (there are Cauchy sequences that "aim" at irrational numbers, e.g. $\sqrt{2}$).

Divergent Sequences and Cluster Points

Example. In $(\mathbb{R}, d_u)$ consider

$$x_n = (-1)^n\Big(1 - \frac{1}{n}\Big) = \begin{cases} 1 - \frac{1}{n}, & n \text{ even}, \\ -1 + \frac{1}{n}, & n \text{ odd}. \end{cases}$$

The sequence does not converge, but it approaches 1 and -1 infinitely many times.

Definition (Cluster point of a sequence)**.** Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in (X, d). A point $x \in X$ is a *cluster point* of (x_n) if for every $\varepsilon > 0$ and every N there exists $n \geq N$ such that $d(x_n, x) < \varepsilon$.

Remark. Equivalently, x is a cluster point of (x_n) if and only if there exists a *subsequence* (x_{n_k}) that converges to x.

Proposition (Cluster points of **??**)**.** *The sequence $x_n = (-1)^n(1 - 1/n)$ has exactly two cluster points: $\{1, -1\}$.*

Proof. That 1 is a cluster point: given $\varepsilon > 0$ choose N such that $1/N < \varepsilon$ and take $n \geq N$ *even.* Then

$$|x_n - 1| = \Big|1 - \frac{1}{n} - 1\Big| = \frac{1}{n} < \varepsilon.$$

Similarly, for -1: given $\varepsilon > 0$, choose N with $1/N < \varepsilon$ and take $n \geq N$ *odd.* Then

$$|x_n - (-1)| = \Big|-1 + \frac{1}{n} + 1\Big| = \frac{1}{n} < \varepsilon.$$

To see that there are no other cluster points, let (x_{n_k}) be a convergent subsequence. Either it contains infinitely many even indices or infinitely many odd indices. In the first case, the subsequence of even indices $x_{2m} = 1 - \frac{1}{2m}$ converges to 1; in the second, the subsequence of odd indices $x_{2m-1} = -1 + \frac{1}{2m-1}$ converges to -1. Therefore, every subsequential limit belongs to $\{1, -1\}$. □

Cluster Points, $\limsup/\liminf$, and Open Sets

Cluster Points of a Sequence

Definition (Cluster point of (x_n))**.** Let (X, d) be a metric space and $(x_n)_{n\in\mathbb{N}} \subset X$. We say that $x \in X$ is a *cluster point* of (x_n) if for every $\varepsilon > 0$ and every $N \in \mathbb{N}$ there exists $n \geq N$ such that $d(x_n, x) < \varepsilon$.

Proposition. *Let (x_n) in (X, d).*

1. *If a subsequence (x_{n_k}) converges to x^*, then x^* is a cluster point of (x_n).*
2. *If (x_n) converges to x, then x is a cluster point (in fact, the only one).*
3. *An unbounded sequence can have cluster points. For example,*

$$x_n = \begin{cases} 1, & n \text{ odd}, \\ n, & n \text{ even}, \end{cases} \quad \text{in } (\mathbb{R}, d_{\mathrm{u}}),$$

is unbounded and has 1 as a cluster point.

Theorem 1 (Bolzano–Weierstrass in $\mathbb{R}$)**.** *Every bounded sequence in $(\mathbb{R}, d_{\mathrm{u}})$ has a convergent subsequence; in particular, it has (at least) one cluster point.*

Remark (Caution with the ambient space)**.** In the subspace $\big((0, 1], d_{\mathrm{u}}\big)$ the sequence $x_n = \frac{1}{n}$ is bounded but does *not* have a cluster point in $(0, 1]$ (its only candidate would be $0 \notin (0, 1]$).

$\limsup$ and $\liminf$

Definition. For a real sequence (x_n) we define

$$\limsup_{n\to\infty} x_n = \inf_{n\in\mathbb{N}} \sup_{k\geq n} x_k, \qquad \liminf_{n\to\infty} x_n = \sup_{n\in\mathbb{N}} \inf_{k\geq n} x_k.$$

Proposition. *For any real sequence (x_n):*

1. $\limsup_{n\to\infty} x_n$ *is the* largest *cluster point of* (x_n).
2. $\liminf_{n\to\infty} x_n$ *is the* smallest *cluster point of* (x_n).
3. *It always holds that* $\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n$, *with equality if and only if* (x_n) *converges.*

Example. Let $x_n = (-1)^n\big(1-\frac{1}{n}\big)$. Then

$$\sup_{k\geq n} x_k = 1-\tfrac{1}{n} \quad\Rightarrow\quad \limsup_{n\to\infty} x_n = \inf_n\big(1-\tfrac{1}{n}\big) = 1,$$

and

$$\inf_{k\geq n} x_k = -1+\tfrac{1}{n} \quad\Rightarrow\quad \liminf_{n\to\infty} x_n = \sup_n\big(-1+\tfrac{1}{n}\big) = -1.$$

Open Balls, Interior Points, and Open Sets

Definition (Open ball)**.** In a metric space (X,d) and for $x\in X$, $\varepsilon>0$, the *open ball* centered at x with radius ε is

$$B(x,\varepsilon) = \{\, y\in X :\ d(x,y)<\varepsilon \,\}.$$

Example.

1. In $(\mathbb{R}, d_{\mathrm{u}})$, $B(x,\varepsilon) = (x-\varepsilon,\ x+\varepsilon)$.
2. In the subspace $\big([0,1], d_{\mathrm{u}}\big)$,

$$B(1/2, 3/4) = \big(-\tfrac{1}{4},\ \tfrac{5}{4}\big)\cap[0,1] = [0,1].$$

3. In $(\mathbb{R}^2, \text{Euclidean distance})$, $B((a,b),\varepsilon)$ is the usual open disk.

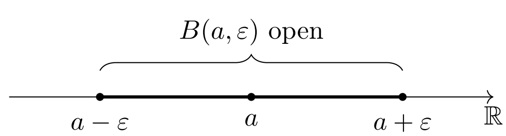
Definition (Interior point and interior)**.** Let $E\subset X$. A point $x\in E$ is *interior* to E if there exists $\varepsilon>0$ such that $B(x,\varepsilon)\subset E$. The *interior* of E is

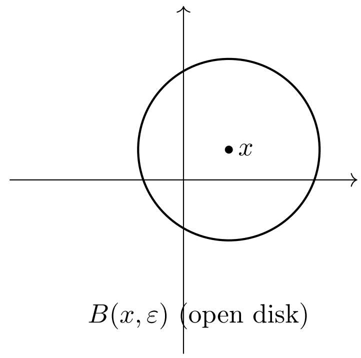
$$\operatorname{int}(E) = \{\, x\in E :\ x \text{ is interior to } E \,\}.$$

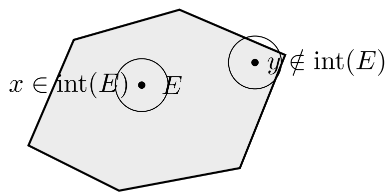
Definition (Open set)**.** A set $E\subset X$ is *open* if $E=\operatorname{int}(E)$, i.e., if every point of E is interior.

Remark.

1. For every E, always $\operatorname{int}(E)\subset E$.
2. E is open $\iff E\subset\operatorname{int}(E)$.
3. Openness depends on the ambient space: in $(\mathbb{R}, d_{\mathrm{u}})$ the set $[0,1)$ is not open (because of 0), whereas in the subspace $\big([0,1], d_{\mathrm{u}}\big)$ the set $(0,1]$ *is* relatively open.







Cluster points and subsequences

Definition (Cluster point)**.** Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a metric space (X, d). We say that $x \in X$ is a *cluster point* of (x_n) if for every $\varepsilon > 0$ and every $N \in \mathbb{N}$ there exists $n \geq N$ such that $d(x_n, x) < \varepsilon$.

- If a *subsequence* (x_{n_k}) converges to x^*, then x^* is a cluster point of the original sequence (x_n).
- If (x_n) converges to x, then x is a cluster point of (x_n).
- An unbounded sequence can have a cluster point. For example, $x_{2k} = k$ and $x_{2k+1} = 1$ is unbounded and has 1 as a cluster point.
- In $(\mathbb{R}, d_\mu)$ (the usual metric), every bounded sequence has a cluster point (Bolzano–Weierstrass): it has a convergent subsequence and its limit is a cluster point.
- In the subspace $((0, 1], d_\mu)$, the sequence $x_n = \frac{1}{n}$ is bounded but does *not* have a cluster point in X (its only natural limit would be $0 \notin X$).

Upper and lower limits

Definition (lim sup and lim inf)**.** For a real sequence (x_n) define

$$\limsup_{n\to\infty} x_n := \inf_{n\in\mathbb{N}} \sup_{k\geq n} x_k, \qquad \liminf_{n\to\infty} x_n := \sup_{n\in\mathbb{N}} \inf_{k\geq n} x_k.$$

The lim sup (resp. lim inf) is the largest (resp. smallest) cluster point.

Example. For $x_n = (-1)^n\bigl(1 - \frac{1}{n}\bigr)$ we have

$$\limsup_{n\to\infty} x_n = 1, \qquad \liminf_{n\to\infty} x_n = -1.$$

Indeed, fixing n, the suprema (resp. infima) of the tails $\{x_k : k \geq n\}$ approach 1 (resp. -1).

Open balls, interior points, and open sets

Definition (Open ball)**.** Given (X, d), $x \in X$, and $\varepsilon > 0$, the *open ball* centered at x with radius ε is

$$B(x, \varepsilon) = \{\, y \in X : \; d(x, y) < \varepsilon \,\}.$$

Example.

1. In $(\mathbb{R}, d_\mu)$: $B(x,\varepsilon) = (x-\varepsilon, x+\varepsilon)$.
2. In the subspace $([0,10], d_\mu)$:

$$B\big(\tfrac{1}{2}, \tfrac{3}{4}\big) = (-\tfrac{1}{4},\ \tfrac{5}{4}) \cap [0,10] = [0,\ \tfrac{5}{4}).$$

3. In $(\mathbb{R}^2, d_{\text{euc}})$: $B(x,\varepsilon)$ is the open disk of radius ε.

Definition (Interior point and interior)**.** Let $E \subset X$. A point $x \in E$ is *interior* to E if there exists $\varepsilon > 0$ such that $B(x,\varepsilon) \subset E$. The *interior* of E is $\text{int}(E) = \{x \in E :\ x \text{ is interior}\}$.

Definition (Open set)**.** A set $E \subset X$ is *open* if $E = \text{int}(E)$, equivalently, if every $x \in E$ is an interior point.

Remark. Openness depends on the underlying space. For example, in $(\mathbb{R}, d_\mu)$, $[0,1)$ is not open; however, in the subspace $((0,100), d_\mu)$ the set $[0,1]$ *is* open.

Example.

1. In the discrete metric (X, d_{disc}), $\{x\}$ is open because $B(x, \frac{1}{2}) = \{x\}$. In fact, *every* subset of X is open.
2. In $(\mathbb{R}, d_\mu)$, no finite subset (nor, more generally, any countable set) is open: every open ball in $\mathbb{R}$ contains infinitely many (indeed uncountably many) points.

Basic properties of open sets

Proposition. *In a metric space (X,d):*

1. *$\varnothing$ and X are open.*
2. *Arbitrary unions of open sets are open.*
3. *Finite intersections of open sets are open.*

Example. A countable intersection of open sets need not be open in $(\mathbb{R}, d_\mu)$:

$$\bigcap_{n\geq 1} \big(-\tfrac{1}{n}, \tfrac{1}{n}\big) = \{0\},$$

which is not open.

Proposition. *For every $x \in X$ and $\varepsilon > 0$, the open ball $B(x,\varepsilon)$ is an open set. In particular,* $\text{int}(E)$ *is always an open set and* $\text{int}(E) \subset E$.

Closed balls and closed sets

Definition (Closed ball and closed set)**.** The *closed ball* centered at x with radius $\varepsilon > 0$ is

$$\overline{B}(x,\varepsilon) = \{\, y \in X :\ d(x,y) \leq \varepsilon \,\}.$$

A set $F \subset X$ is *closed* if $X \setminus F$ is open.

Proposition. *Every closed ball $\overline{B}(x,\varepsilon)$ is a closed set.*

Remark (The real line $\mathbb{R}$)**.** In $(\mathbb{R}, d_\mu)$, the complement of a closed interval $[x-\varepsilon, x+\varepsilon]$ is open: if $y \notin [x-\varepsilon, x+\varepsilon]$, then taking $\varepsilon' = \frac{1}{2}\,\text{dist}(y, [x-\varepsilon, x+\varepsilon]) > 0$ we have $B(y,\varepsilon') \subset \mathbb{R} \setminus [x-\varepsilon, x+\varepsilon]$.

Lemma. *Let (x_n) be a sequence in a metric space (X,d). A point $x^* \in X$ is a* cluster point *of (x_n) if and only if there exists a subsequence (x_{n_k}) such that $x_{n_k} \to x^*$.*

Example. Consider in $(\mathbb{R}, d_u)$ the sequence

$$x_1 = 1,\ x_2 = 2,\ x_3 = 1,\ x_4 = 3,\ x_5 = 1,\ x_6 = 4, \dots$$

For each n we have $\sup\{x_k : k \geq n\} = +\infty$, so

$$\limsup_{n\to\infty} x_n = \inf_n \sup_{k\geq n} x_k = +\infty,$$

which is *not* a cluster point in $\mathbb{R}$. On the other hand, $\inf\{x_k : k \geq n\} = 1$ for all n, and thus

$$\liminf_{n\to\infty} x_n = \sup_n \inf_{k\geq n} x_k = 1,$$

and 1 is indeed a cluster point (taking the subsequence of odd terms).

Proposition. *For any real sequence (x_n):*

1. *If x^* is a cluster point, then $\liminf_{n\to\infty} x_n \leq x^* \leq \limsup_{n\to\infty} x_n$.*
2. *If $\limsup_{n\to\infty} x_n \in \mathbb{R}$, then this value is a cluster point (analogously for the* lim inf*).*
3. *The limit $\lim_{n\to\infty} x_n = x \in \mathbb{R}$ exists if and only if $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n = x$.*

Closed sets

Proposition. *In a metric space (X, d):*

1. *$\varnothing$ and X are closed.*
2. *The finite union of closed sets is closed.*
3. *The arbitrary intersection of closed sets is closed.*

Example. In $(\mathbb{R}, d_u)$, for $n \in \mathbb{N}$ let $F_n = [-1 + 1/n,\ 1 - 1/n]$. Each F_n is closed, but

$$\bigcup_{n\in\mathbb{N}} F_n = (-1, 1),$$

which is not closed (its complement is not open; at $x = 1$ there is no $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq (-\infty, -1] \cup [1, \infty)$).

Proposition (Sequential characterization of closed sets)**.** *A set $E \subseteq X$ is closed if and only if for every convergent sequence (x_n) with $x_n \in E$ for all n, we have $\lim_{n\to\infty} x_n \in E$.*

Idea. If E is closed and $x_n \to x$ with $x_n \in E$, assuming $x \notin E$ implies $x \in X \setminus E$, which is open; hence there exists $\varepsilon > 0$ with $B(x, \varepsilon) \subseteq X \setminus E$. But then, for n large, $x_n \in B(x, \varepsilon) \subseteq X \setminus E$, a contradiction. The converse follows by considering complements. □

Proposition (Sequential characterization of closed sets)**.** *Let (X, d) be a metric space and $E \subseteq X$. The following are equivalent:*

1. *E is closed.*
2. *For every sequence $\{x_n\}_{n\in\mathbb{N}}$ with $x_n \in E$ and $x_n \xrightarrow[n\to\infty]{} x \in X$, it follows that $x \in E$.*

Proof. (1) ⇒ (2): If E is closed, then $X \setminus E$ is open. If there were $x \in X \setminus E$ and $x_n \in E$ with $x_n \to x$, choose $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq X \setminus E$, contradicting the fact that eventually $x_n \in B(x, \varepsilon)$.

(2) ⇒ (1) (contrapositive): Suppose E is not closed. Then $X \setminus E$ is not open: there exists $x \in X \setminus E$ that is not an interior point of $X \setminus E$. Hence, for every $n \in \mathbb{N}$,

$$B(x, 1/n) \cap E \neq \varnothing.$$

Choose $x_n \in B(x, 1/n) \cap E$; then $x_n \to x$ and $x \notin E$. Thus we obtain a sequence of points in E converging to a point outside E, contradicting (2). Therefore, E is closed. $\square$

Remark (Practical guide)**.** To show that E is closed, it suffices to:

1. Prove that $X \setminus E$ is open; **or**
2. Prove that every convergent sequence in E has its limit in E.

To show that E is *not* closed, it suffices to:

1. Exhibit that $X \setminus E$ is not open; **or**
2. Find a sequence $\{x_n\}_{n\in\mathbb{N}}$ with $x_n \in E$ and $x_n \to x \notin E$.

Proposition (Operations with closed sets)**.** *Let* (X, d) *be a metric space. Then:*

1. $\varnothing$ *and* X *are closed.*
2. *A finite union of closed sets is closed.*
3. *An (arbitrary) intersection of closed sets is closed.*

Example (Countable union of closed sets that is not closed)**.** In $(\mathbb{R}, d_u)$ with the usual metric, for $n \in \mathbb{N}$ define

$$F_n = \left[-1 + \tfrac{1}{n},\ 1 - \tfrac{1}{n}\right].$$

Each F_n is closed. However

$$\bigcup_{n\in\mathbb{N}} F_n = (-1, 1),$$

which is not closed (for instance, 1 is an adherent point but $1 \notin (-1, 1)$).

Interior, adherence, and closure

Definition (Interior point and interior)**.** Let $E \subseteq X$. We say that $x \in E$ is an *interior point* of E if there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq E$. The set of all interior points is denoted

$$\mathrm{int}(()E) = \{x \in E : \exists\, \varepsilon > 0 \text{ with } B(x, \varepsilon) \subseteq E\}$$

and is called the *interior* of E.

Proposition (Basic properties of the interior)**.** *For every* $E \subseteq X$*:*

1. $\mathrm{int}(()E)$ *is open and* $\mathrm{int}(()E) \subseteq E$.
2. *If* U *is open and* $U \subseteq E$*, then* $U \subseteq \mathrm{int}(()E)$ *(largest open set contained in* E*).*
3. $\mathrm{int}(()\mathrm{int}(()E)) = \mathrm{int}(()E)$.

Definition (Adherent point and closure)**.** Let $E \subseteq X$. We say that $x \in X$ is an *adherent point* of E if

$$\forall\, \varepsilon > 0, \quad B(x, \varepsilon) \cap E \neq \varnothing.$$

The *closure* of E is the set of its adherent points:

$$\overline{E} = \{x \in X : \forall\, \varepsilon > 0,\ B(x, \varepsilon) \cap E \neq \varnothing\}.$$

Proposition (Basic properties of the closure)**.** *For every* $E \subseteq X$*:*

1. $\overline{E}$ *is closed and* $E \subseteq \overline{E}$.
2. *If* F *is closed and* $E \subseteq F$*, then* $\overline{E} \subseteq F$ *(smallest closed set containing* E*).*

3. *We have the open–closed sandwich:*

$$\operatorname{int}(()E) \ \subseteq\ E \ \subseteq\ \overline{E}.$$

Example (Set neither open nor closed in $\mathbb{R}$)**.** In $(\mathbb{R}, d_u)$, the set $E = [0, 1)$ is not open (no ball around 0 fits inside E) and not closed (its complement $\mathbb{R} \setminus E = (-\infty, 0) \cup [1, \infty)$ is not open).

Example (Open ball plus isolated points)**.** In $(\mathbb{R}^2, d_{\text{eucl}})$ fix $x \in \mathbb{R}^2$ and $\varepsilon > 0$. Let

$$F = B(x, \varepsilon) \ \cup\ \{p_1, \dots, p_m\},$$

where $p_i \in \partial B(x, \varepsilon)$ are isolated boundary points. Then:

- F is not open (no p_i has a ball $B(p_i, \delta)$ contained in F).
- F is not closed: for instance, a point q of the circle $\partial B(x, \varepsilon)$ different from the p_i is an adherent point of F but $q \notin F$. In particular,

$$\operatorname{int}(()F) = B(x, \varepsilon) \quad \text{and} \quad \overline{F} = \overline{B}(x, \varepsilon).$$

Remark (Closure of open balls)**.** In normed spaces (in particular in $\mathbb{R}^n$ with the Euclidean norm), $\overline{B(x, \varepsilon)} = \overline{B}(x, \varepsilon)$.

Definition (Adherent point, closure, and boundary)**.** Let (X, d) be a metric space and $E \subseteq X$. We say that $x \in X$ is an *adherent point* of E if for every $\varepsilon > 0$ we have

$$B(x, \varepsilon) \cap E \neq \varnothing.$$

The *closure* of E is

$$\overline{E} = \operatorname{cl}(E) = \{x \in X :\ x \text{ is an adherent point of } E\}.$$

The *boundary* of E is

$$\partial E = \overline{E} \setminus \operatorname{int}(E).$$

Proposition. *Let $E \subseteq X$. Then:*

1. $\operatorname{int}(E)$ *is the largest open set contained in E.*
2. $\overline{E}$ *is the smallest closed set containing E.*

Proof. (1) By definition, $x \in \operatorname{int}(E)$ if there exists $\varepsilon > 0$ with $B(x, \varepsilon) \subseteq E$, hence $\operatorname{int}(E)$ is open and $\operatorname{int}(E) \subseteq E$. If U is open and $U \subseteq E$, then for each $x \in U$ we have $B(x, \varepsilon) \subseteq U \subseteq E$, so $x \in \operatorname{int}(E)$ and $U \subseteq \operatorname{int}(E)$.

(2) If $x \notin \overline{E}$, there exists $\varepsilon > 0$ with $B(x, \varepsilon) \cap E = \varnothing$, hence $B(x, \varepsilon) \subseteq X \setminus \overline{E}$, so $X \setminus \overline{E}$ is open and $\overline{E}$ is closed. If F is closed and $E \subseteq F$, every adherent point of E is also an adherent point of F, thus $\overline{E} \subseteq F$. ☐

Definition (Cover and open cover)**.** Let $E \subseteq X$. A family $\{F_i\}_{i \in I}$ of subsets of X is a *cover* of E if $E \subseteq \bigcup_{i \in I} F_i$. It is an *open cover* if, moreover, each F_i is open in X.

Definition (Compact set)**.** We say that $E \subseteq X$ is *compact* if every open cover of E admits a finite subcover; that is, if $\{F_i\}_{i \in I}$ is an open cover of E, then there exists a finite set $J \subseteq I$ such that $E \subseteq \bigcup_{i \in J} F_i$. A metric space (X, d) is *compact* if X is compact as a subset of itself.

Example ($(0, 1)$ is not compact in $(\mathbb{R}, d_{\text{us}})$)**.** Consider the family of open sets

$$\mathcal{U} = \{\, (1/n,\, 1) : n \geq 2 \,\}.$$

We have $\bigcup_{n \geq 2} (1/n, 1) = (0, 1)$, so $\mathcal{U}$ is an open cover of $(0, 1)$. If we take a finite subfamily $\{(1/n_k, 1)\}_{k=1}^{m}$, then

$$\bigcup_{k=1}^{m} (1/n_k, 1) = \big(1/\max\{n_k\}_{k=1}^{m},\, 1\big) \neq (0, 1),$$

since, for example, $x = \frac{1}{\max\{n_k\}+1} \in (0,1)$ is not covered. Therefore, no finite subcover exists and $(0,1)$ is not compact.

Remark. Being compact does not mean "having some" finite open cover, but that *every* open cover admits a finite subcover.

Example ($(0,1)$ is not compact in $(\mathbb{R}, d_{\text{euc}})$)**.** Let $E = (0,1)$ and consider the open cover

$$\mathcal{U} = \{ U_n = (\tfrac{1}{n}, 1) : n \geq 2 \}.$$

We have $\bigcup_{n\geq 2} U_n = (0,1)$. If we take a finite subcover $\{U_{n_1}, \ldots, U_{n_k}\}$, then $\bigcup_{j=1}^{k} U_{n_j} = (\frac{1}{N}, 1)$ with $N = \max\{n_1, \ldots, n_k\}$, and thus points near 0 are not covered (for instance $x \in (0, \frac{1}{N})$). Therefore E does not admit a finite subcover and is not compact.

Example (A closed unbounded set that is not compact)**.** In $(\mathbb{R}, d_{\text{euc}})$, let $E = [0, \infty)$. For $n \in \mathbb{Z}$ with $n \geq -1$, define $U_n = (n, n+2)$. Then $\bigcup_{n\geq -1} U_n = (-1, \infty) \supset E$, i.e. this is an open cover of E. No finite subcover can cover E: if we choose finitely many $U_{n_1}, \ldots, U_{n_k}$ and $N = \max\{n_1, \ldots, n_k\}$, then $\bigcup_{j=1}^{k} U_{n_j} \subset (-\infty, N+2)$, so $E \cap (N+3, \infty)$ remains uncovered. Hence E is not compact.

Definition (Bounded set)**.** Let (X, d) be a metric space. We say that $E \subset X$ is *bounded* if there exist $x \in X$ and $\Pi > 0$ such that $d(x, y) \leq \Pi$ for all $y \in E$.

Proposition. *If $E \subset X$ is compact, then E is closed and bounded.*

Proof. *Closed:* If $(x_n) \subset E$ and $x_n \to x$ in X, since E is compact there exists a subsequence $x_{n_k} \to y$ with $y \in E$. But convergence in metric spaces is unique, hence $x = y \in E$. Thus E contains the limits of its sequences and is closed.

Bounded: Fix $x_0 \in X$. The balls $\{B(x_0, n)\}_{n\in\mathbb{N}}$ cover X, hence they cover E. By compactness, there exists N such that $E \subset B(x_0, N)$. This shows that E is bounded. □

Example (Closed and bounded but not compact in the discrete metric)**.** Consider $X = \{1/n : n \in \mathbb{N}\}$ with the discrete metric d_{disc}. Then every subset (in particular each singleton $\{1/n\}$) is open. The family $\mathcal{U} = \{\{1/n\} : n \in \mathbb{N}\}$ is an open cover of X with no finite subcover, so X is not compact. Nevertheless, X is bounded and closed in itself.

Theorem 2 (Heine–Borel)**.** *In $(\mathbb{R}^j, d_{\text{euc}})$ a set $E \subset \mathbb{R}^j$ is compact if and only if it is closed and bounded.*

Proposition (Maximum in compact subsets of $\mathbb{R}$)**.** *If $E \subset \mathbb{R}$ is compact in $(\mathbb{R}, d_{\text{euc}})$, then E attains its maximum (and analogously its minimum).*

Proof. By Proposition , E is closed and bounded; in particular it has a supremum $\bar{x} = \sup E \in \mathbb{R}$. For each n there is $x_n \in E$ with $\bar{x} - \frac{1}{n} < x_n \leq \bar{x}$. Then $x_n \to \bar{x}$; since E is closed, $\bar{x} \in E$, and therefore $\max E = \bar{x}$. □

Proposition. *Let (X, d) be a compact metric space and $F \subset X$ a closed subset. Then F is compact.*

Proof. If $\{U_i\}_{i\in I}$ is an open cover of F, then $\{U_i\}_{i\in I} \cup \{X \setminus F\}$ is an open cover of X. By compactness of X it admits a finite subcover; removing (if it appears) $X \setminus F$, we obtain a finite subcover of F. □

Proposition. *Every compact metric space is complete.*

Proof. Let (x_n) be a Cauchy sequence in compact X. Then it has a convergent subsequence $x_{n_k} \to x \in X$. Since (x_n) is Cauchy, every subsequence has the same Cauchy bound and necessarily $x_n \to x$. □

Continuity in metric spaces

Fix metric spaces (X, d_X) and (Y, d_Y) and a function $f : X \to Y$.

Definition (Continuity at a point)**.** We say that f is continuous at $x_0 \in X$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

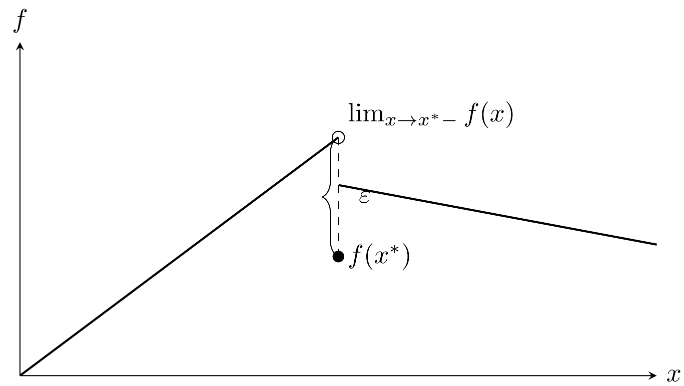
$$d_X(x, x_0) < \delta \quad \Longrightarrow \quad d_Y\big(f(x), f(x_0)\big) < \varepsilon.$$

We say that f is continuous if it is continuous at every point of X.

Proposition (Sequential characterization)**.** *f is continuous at x_0 if and only if for every sequence $(x_n) \subset X$ with $x_n \to x_0$ we have $f(x_n) \to f(x_0)$ in Y.*

Continuity

Let (X, d_X) and (Y, d_Y) be two metric spaces and $f : X \to Y$ a function. Intuitively, a *discontinuity* at x^* occurs if there are points x' near x^* such that $f(x')$ is not near $f(x^*)$.



Definition (Pointwise continuity)**.** We say that $f : X \to Y$ is *continuous at* $x \in X$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d_X(x, x') < \delta \implies d_Y\big(f(x), f(x')\big) < \varepsilon.$$

The function is *continuous* if it is continuous at every $x \in X$.

Proposition (Sequential characterization)**.** *f is continuous at $x \in X$ if and only if for every sequence $(x_n)_{n\in\mathbb{N}}$ in X with $x_n \to x$ we have*

$$f(x_n) \longrightarrow f(x) \quad \text{in } Y.$$

Proposition (Topological characterization)**.** *A function $f : X \to Y$ is continuous if and only if for every open set $U \subseteq Y$ the preimage $f^{-1}(U) = \{x \in X : \ f(x) \in U\}$ is open in X. The same is true for closed sets.*

Example (Basic examples of continuous functions)**.**

1) **Identity.** If $X = Y$ and $f(x) = x$, then f is continuous (indeed, $f^{-1}(U) = U$).

2) **Constant function.** If $f(x) = y_0$ for all $x \in X$, then f is continuous (the preimage of any open $U \subseteq Y$ is X if $y_0 \in U$ and $\varnothing$ otherwise).

3) **Finite domain.** If X is finite (with any metric), every subset of X is open; therefore, *every* function $f : X \to Y$ is continuous.

4) **Linear functions in $\mathbb{R}$.** With the usual metric $d(u, v) = |u - v|$, the function $f(x) = ax + b$ $(a \neq 0)$ is continuous. Given $\varepsilon > 0$, choose $\delta = \varepsilon/|a|$, since

$$|x - x'| < \delta \ \Rightarrow \ |f(x) - f(x')| = |a||x - x'| < |a|\delta = \varepsilon.$$

Proposition (Algebra of continuous functions)**.** *Let $f, g : X \to \mathbb{R}$ be continuous (with the usual metric on $\mathbb{R}$). Then the following are also continuous:*

(i) $f+g$;

(ii) $f \cdot g$;

(iii) $\dfrac{f}{g}$, *provided* $g(x) \neq 0$ *for all* $x \in X$;

(iv) $\max\{f,g\}$;

(v) $\min\{f,g\}$;

(vi) $|f| : x \mapsto |f(x)|$.

Definition (Uniform continuity)**.** We say that $f : X \to Y$ is *uniformly continuous* if for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$d_X(x,x') < \delta \implies d_Y\big(f(x), f(x')\big) < \varepsilon \qquad \text{for } \textit{all } x, x' \in X.$$

The key difference with pointwise continuity is that here δ does not depend on the point x.

Example (Not uniformly continuous in $\mathbb{R}$)**.** Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, with the usual metric. f is continuous on $\mathbb{R}$, but *not* uniformly continuous.

Proof. Take $\varepsilon = 1$. Let $\delta > 0$ be arbitrary and define

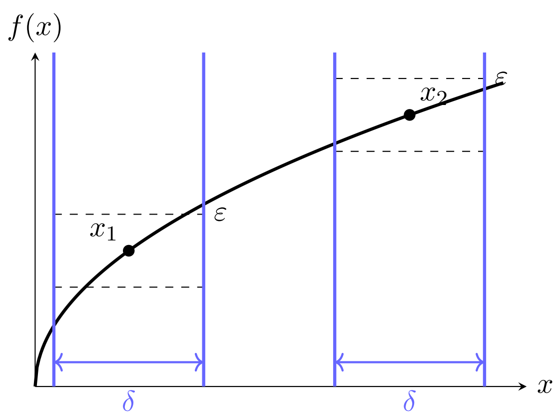
$$x := \frac{1}{\delta} + \frac{\delta}{4}, \qquad x' := x - \frac{\delta}{2}.$$

Then $|x - x'| = \delta/2 < \delta$, but

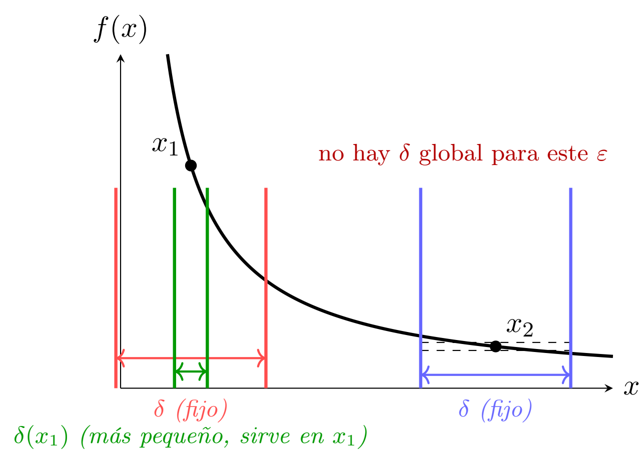
$$\big|x^2 - (x')^2\big| = \big|x - x'\big| \cdot \big|x + x'\big| = \frac{\delta}{2}\Big(2x - \frac{\delta}{2}\Big) = \delta x - \frac{\delta^2}{4} = \delta\Big(\frac{1}{\delta} + \frac{\delta}{4}\Big) - \frac{\delta^2}{4} = 1.$$

In particular, for that pair x, x' we have $|x^2 - (x')^2| \geq \varepsilon$ even though $|x - x'| < \delta$. Since this holds for every $\delta > 0$, f is not uniformly continuous. ☐

(a) **Uniform continuity:** $f(x) = \sqrt{x}$ en $[0,1]$. Mismo δ funciona para cualquier x_0 (dados ε).



(b) **Pointwise continuity:** $f(x) = 1/x$ en $(0,1]$. El δ requerido depende fuertemente de x_0.



Lets build intuition. We have to different but related concepts:

- **Uniform continuity.** Given a function $f : A \to \mathbb{R}$, we say it is uniformly continuous if for every $\varepsilon > 0$ there exists *a single* $\delta > 0$ (which does not depend on the point in the domain) such that, for any $x, y \in A$,

$$|x - y| < \delta \quad \Longrightarrow \quad |f(x) - f(y)| < \varepsilon.$$

  The "uniformity" means precisely that *one* $\delta(\varepsilon)$ works *across the whole domain.* Geometrically: imagine a "caliper" that opens δ along the x-axis. If, while sliding it over the graph, the images of any pair of points separated by less than δ always remain within a vertical band of height ε, then f is uniformly continuous. Choosing δ amounts to ensuring that this caliper works in the *worst corner* of the domain; if it works there, it will surely work elsewhere.

- **Pointwise (or "conventional") continuity.** In contrast, f is continuous at x_0 if for every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon, x_0)$ such that

$$|x - x_0| < \delta \quad \Longrightarrow \quad |f(x) - f(x_0)| < \varepsilon.$$

  Here δ *may depend on the point* x_0. That is why a function may be continuous at every point and yet *fail* to be uniformly continuous: the required δ keeps shrinking "more and more" as we move across certain regions of the domain, and there is *no single* δ that works for all points simultaneously.