A TOPOLOGICAL INTRODUCTION TO REAL ANALYSIS: REFERENCE MANUAL

FRÉDÉRIC LATRÉMOLIÈRE

Abstract. We present a topological approach to real analysis. This set of notes assume the notion of an ordered field, and focuses on the definitions, theorems and proofs, with no additional comment, as a sort of reference manual.

1. The Continuum

1.1. Review: special elements related to orders.

Definition 1.1. Let $(E, \leqslant)$ be an ordered set and $A \subseteq E$, $A \neq \emptyset$. An element $M \in E$ is a *lower bound* of A when $\forall a \in A \quad M \leqslant a$. An element $M \in E$ is an *upper bound* of A when $\forall a \in A \quad a \leqslant M$.

Definition 1.2. Let $(E, \leqslant)$ be an ordered set, and let $A \subseteq E$ be nonempty. The set A is *bounded above* when it has an upper bound; A is *bounded below* when it has a lower bound. Last, A is *bounded* when it has both an upper bound and a lower bound.

Definition 1.3. Let $(E, \leqslant)$ be an ordered set and $A \subseteq E$, $A \neq \emptyset$. An element $x \in A$ is a *smallest element* of A when x is lower bound of A. An element $x \in A$ is a *largest element* of A when x is an upper bound of A.

If $x, y \in A$ are two smallest elements of A then $x \leqslant y$ and $y \leqslant x$ so $x = y$ — smallest elements, if they exist, are unique. Similarly, largest elements are unique as well. We denote the smallest element of A, if it exists, by $\min A$, and the largest element of A, if it exists, by $\max A$.

Definition 1.4. Let $(E, \leqslant)$ be an ordered set and $A \subseteq E$ a nonempty subset of E. The *infimum* $\inf A$, if it exists, is the largest element of the set of lower bounds of A. The *supremum* $\sup A$, if it exists, is the smallest element of the set of upper bounds of A.

Suprema and infima are unique, if they exist.

Let $(E, \leqslant)$ be an ordered set, $A \subseteq E$ not empty, and $x \in E$. The *key idea* is the following equivalence:

- $x = \inf A$,
- x is a lower bound of A, and $\forall y \in E \quad y > x \implies \exists a \in A \quad a < y$.

Similarly,

Date: October 26, 2024.

- $x = \sup A$,
- x is an upper bound of A, and $\forall y \in E \quad y < x \implies \exists a \in A \quad y < a$.

1.2. **Archimedean Fields.**

Definition 1.5. An ordered field $\mathbb{F}$ is *Archimedean* when $\inf\left\{\frac{1}{n} : n \in \mathbb{N}, n > 0\right\} = 0$.

Lemma 1.6. *$\mathbb{Q}$ is an Archimedean field.*

Proof. For all $n \in \mathbb{N}$, $n > 0$, we observe that $\frac{1}{n} \geqslant 0$. Let $q > 0$. Then $q = \frac{p}{n} > \frac{1}{n}$ for some $p, n \in \mathbb{N} \setminus \{0\}$. So $\inf\{\frac{1}{n} : n \in \mathbb{N}, n > 0\} = 0$. This concludes our proof. $\square$

Lemma 1.7. *A field $\mathbb{F}$ is Archimedean if, and only if, for all $x, y \in \mathbb{F}$, if $x > 0$ and $y > 0$, then there exists $n \in \mathbb{N}$ such that $y < nx$.*

Proof. Assume that $\mathbb{F}$ is Archimedean. There exists $n \in \mathbb{N}$ such that $\frac{1}{n} < \frac{x}{y}$ since $\frac{x}{y} > 0$; since $y > 0$ and $n > 0$, we conclude that $y < nx$, as claimed.

Assume that for all $x, y \in \mathbb{F}$, with $x > 0$ and $y > 0$, there exists $n \in \mathbb{N}$ such that $x < ny$. First, note that $\frac{1}{n} > 0$ for all $n \in \mathbb{N}$, $n > 0$. On the other hand, let $x \in \mathbb{F}$, with $x > 0$. There exists $n \in \mathbb{N}$ such that $nx > 1$ — in particular, $n > 0$. Thus $x > \frac{1}{n}$. So $\inf\left\{\frac{1}{n} : n \in \mathbb{N}, n > 0\right\} = 0$ and $\mathbb{F}$ is Archimedean. $\square$

While ordered fields always contain a copy of $\mathbb{Q}$, Archimedean fields always contain a *dense* copy of the field of rational numbers.

Theorem 1.8. *If $\mathbb{F}$ is an Archimedean ordered field then:*

$$\begin{aligned}\forall x \in \mathbb{F} \quad x &= \inf\{q \in \mathbb{Q} : x \leqslant q\} \\ &= \inf\{q \in \mathbb{Q} : x < q\} \\ &= \sup\{q \in \mathbb{Q} : q \leqslant x\} \\ &= \sup\{q \in \mathbb{Q} : q < x\}.\end{aligned}$$

In particular, if $x, y \in \mathbb{F}$ and $x < y$ then there exists $q \in \mathbb{Q}$ such that $x < q < y$.

Proof. Let $x \in \mathbb{F}$ with $x > 0$.

Let $S = \{q \in \mathbb{Q} : x < q\}$. Since $\mathbb{F}$ is Archimedean, there exists $k \in \mathbb{N}$ such that $x \leqslant k$, so $k + 1 \in S$ and thus $S \neq \emptyset$.

By construction, x is a lower bound of S. Let now $y \in \mathbb{F}$ with $x < y$. Since $0 = \inf\left\{\frac{1}{n} : n \in \mathbb{N}\right\}$, there exists $n \in \mathbb{N}$ such that $\frac{1}{n} < y - x$.

Let $N = \left\{p \in \mathbb{N} : x < \frac{p}{n}\right\}$. Since $\mathbb{F}$ is Archimedean, there exists $p \in \mathbb{N}$ such that $p\frac{1}{nx} > 1$ i.e. $\frac{p}{n} > x$ so N is not empty. As a nonempty subset of $\mathbb{N}$, the set N has a smallest element q. Note that since $x > 0$, we have $q \geqslant 1$.

Now, $\frac{q-1}{n} \leqslant x$ since $q - 1 < q$ and thus $q - 1 \notin N$. Therefore:

$$x < \frac{q}{n} = \frac{q-1}{n} + \frac{1}{n} \leqslant x + \frac{1}{n} < x + (y - x) = y.$$

Therefore, $\frac{q}{n} \in S$ and $\frac{q}{n} < y$. Thus y is not a lower bound for S.

Therefore, we have proven that $x = \inf S$ as desired. Moreover, as claimed, there exists $\frac{q}{n} \in \mathbb{Q}$ with $x < \frac{q}{n} < y$. We note that if $x' < y' < 0$ then there exists $q \in \mathbb{Q}$ with $-y' < q < -x'$ and thus $x' < -q < y'$. The case $x' < 0 < y'$ being trivial here, we note that indeed, for all $x' < y'$ there exists $q \in \mathbb{Q}$ such that $x' < q < y'$.

Let now $U = \{q \in \mathbb{Q} : q < x\}$. As $x > 0$, we have $0 \in U$ so U is not empty. By construction, x is an upper bound for U.

Let $y < x$. As $\mathbb{F}$ is Archimedean, as above, there exists $n \in \mathbb{N}$ such that $\frac{1}{n} < x - y$ since $0 < x - y$. Let $P = \{p \in \mathbb{N} : \frac{p}{n} \geqslant x\}$; for the same reason as with the set N, the set P is not empty; let $q = \min P$. By definition of P, we thus have $\frac{q-1}{n} \leqslant x$.

We then note that:

$$y = x + y - x < x - \frac{1}{n} \leqslant \frac{q}{n} - \frac{1}{n} = \frac{q-1}{n}.$$

Thus $\frac{q-1}{n} \in U$, and $y < \frac{q-1}{n}$, so y is not an upper bound of U. Consequently, $\sup U = x$ as desired.

If $x = 0$ then, since $\left\{\frac{1}{n} : n \in \mathbb{N}\right\} \subseteq S$, we have $0 \leqslant \inf S \leqslant \inf \left\{\frac{1}{n} : n \in \mathbb{N}\right\} = 0$ so $\inf S = 0$.

Now, if $x < 0$ then:

$$\begin{aligned} x &= -(-x) \\ &= -\inf\{q \in \mathbb{Q} : -x \leqslant q\} \\ &= \sup\{-q \in \mathbb{Q} : -x \leqslant -q\} \\ &= \sup\{q \in \mathbb{Q} : q \leqslant x\}. \end{aligned}$$

Similarly, $x = \inf\{q \in \mathbb{Q} : x \leqslant q\}$.

Last, let $x \in \mathbb{F}$. As x is also a lower bound for $S' = \{q \in \mathbb{Q} : x \leqslant q\}$ and $S \subseteq S'$ so $x \leqslant \inf S' \leqslant \inf S = x$, so $x = \inf S'$. Alternatively, if $x \in \mathbb{Q}$ then of course x is the smallest element of S' so $\inf S' = x$, and if $x \notin Q$ then $S = S'$ and thus $\inf S' = \inf S = x$.

Similarly, $x = \sup\{q \in \mathbb{Q} : q \leqslant x\}$. $\square$

1.3. **The Field $\mathbb{R}$ of real numbers.** We build a field $\mathbb{R}$ from the field of rationals, using Dedekind cut construction.

Definition 1.9. A subset $C \subseteq \mathbb{Q}$ is a *cut* when:

(1) $C \neq \emptyset$,
(2) $\forall x, y \in \mathbb{Q} \quad x < y$ and $x \in C \implies y \in C$,
(3) C is bounded below,
(4) C has no smallest element.

Notation 1.10. For any $q \in \mathbb{Q}$, we write $q_{\mathbb{R}} = \{x \in Q : x > q\}$; the set $q_{\mathbb{R}}$ is a cut.

Inclusion defines an order on the set of cuts.

Notation 1.11. If x, y are cuts, we write $x \leqslant y$ whenever $y \subseteq x$.

Theorem 1.12. *The ordering on cuts is linear, and every nonempty bounded-below set of cuts has an infimum.*

Proof. Let x, y be two cuts. Assume that $x \not\subseteq y$. Let $t \in y$. If for all $s \in x$, we have $s \geqslant t$ then $\forall s \in x\ s \in y$ as y is a cut. Since $x \not\subseteq y$, there exists $s \in x$ such that $s < t$. Since x is a cut, $t \in x$. So $y \subseteq x$. So either $x \leqslant y$ or $y \leqslant x$.

Now, let S be a nonempty set of cuts, bounded below. Let $x = \bigcup S$. Of course, x is the least upper bound of S for $\subseteq$ and thus, if x is indeed a cut, then x is the infimum of S. Since cuts are not empty, x is not empty. Since S is bounded, there exists a cut m such that $m \leqslant t$ for all $t \in S$. Thus $t \subseteq m$ for all $t \in S$, so $x \subseteq m$. So x is bounded below, since m is. Let $t, s \in \mathbb{Q}$ with $t < s$ and $t \in x$. By definition, there exists $y \in S$ such that $t \in y$. Since y is a cut, $s \in y$ so $s \in x$. Last, assume x has a smallest element $m \in \mathbb{Q}$. Then $m \in y$ for some $y \in S$. Yet m is not the smallest element of y since cuts do not have smallest elements. Therefore, there exists $u \in y$ with $u < m$ — but then, $u \in x$, which contradicts that $m = \min x$. So x is a cut, as desired. □

Theorem 1.13. *If x, y are cuts, then $x +_{\mathbb{R}} y$, defined by*

$$x +_{\mathbb{R}} y = \{t + s : t \in x, s \in y\}$$

is a cut.

Proof. By construction, $x +_{\mathbb{R}} y$ is not empty. Let $M_x, M_y \in \mathbb{Q}$ be lower bounds for, respectively, x and y. Let $t \in x$ and $s \in y$. Then $M_x < t$ and $M_y < s$ so $M_x + M_y < s + t$. So $x +_{\mathbb{R}} y$ is bounded below. Let $t, s \in \mathbb{Q}$ such that $t < s$ and $t \in x +_{\mathbb{R}} y$. By definition, $t = u + v$ with $u \in x$ and $v \in y$. Thus $s - u > t - u = v$ so $s - u \in y$ as y is a cut. Thus $s = u + (s - u) \in x +_{\mathbb{R}} y$ as needed. Last, assume $s \in x +_{\mathbb{R}} y$ is the smallest element of $x +_{\mathbb{R}} y$. By definition, $s = u + v$ for some $u \in x$ and $v \in y$. Let now $a \in x$. Since $a + v \in x +_{\mathbb{R}} y$, we have $u + v = s \leqslant a + v$ so $u \leqslant a$. Hence u is the smallest element of x, which does not exist. So $x +_{\mathbb{R}} y$ does not have a smallest element. □

Theorem 1.14. *The set of all cuts with the operation $+_{\mathbb{R}}$ is an Abelian group with neutral element $0_{\mathbb{R}}$.*

Proof. Let $x \subseteq \mathbb{Q}$ be a cut. Let $t \in x +_{\mathbb{R}} 0_{\mathbb{R}}$, so $t = u + v$ with $u \in x$ and $v > 0$. So $t > u$ and $u \in x$, so $t \in x$ as x is a cut. So $x + 0_{\mathbb{R}} \subseteq x$. Now, let $t \in x$. Since x has no smallest element, there exists $u \in x$ with $u < t$. Let $\varepsilon = t - u > 0$ — by definition, $\varepsilon \in 0_{\mathbb{R}}$. Moreover, $t = u + \varepsilon$, so $t \in x +_{\mathbb{R}} 0_{\mathbb{R}}$. So $x +_{\mathbb{R}} 0_{\mathbb{R}} = x$ for all cuts x.

The operation $+_{\mathbb{R}}$ is trivially commutative by definition. It is also easy to check that it is associative. Let x, y, z be three cuts. Let $t \in (x +_{\mathbb{R}} y) +_{\mathbb{R}} z$. Thus, $t = u + v$ for some $u \in x +_{\mathbb{R}} y$ and $v \in z$. Further, there exist $s \in x$ and $w \in y$ such that $u = s + w$. So $t = (s + w) + v = s + (w + v)$, i.e. $t \in x +_{\mathbb{R}} (y +_{\mathbb{R}} z)$. We thus proved that $(x +_{\mathbb{R}} y) +_{\mathbb{R}} z \subseteq x +_{\mathbb{R}} (y +_{\mathbb{R}} z)$. The same argument will show that $x +_{\mathbb{R}} (y +_{\mathbb{R}} z) \subseteq (x +_{\mathbb{R}} y) +_{\mathbb{R}} z$, i.e. $(x +_{\mathbb{R}} y) +_{\mathbb{R}} z = x +_{\mathbb{R}} (y +_{\mathbb{R}} z)$.

It thus suffices to prove that every cut has an opposite. Let x be a cut. We define:

$$-x = \{s \in \mathbb{Q} : \exists u \in \mathbb{Q} \quad \forall t \in x \quad s > u \text{ and } u + t > 0\}.$$

We first check that $-x$ is a cut. Let $M \in \mathbb{Q}$ such that for all $t \in x$, we have $M < t$. Thus for all $t \in x$, we have $t - M > 0$, and thus if $s > -M$ then $s \in -x$. So $x \neq \emptyset$. On the other hand, let $t \in x$ and $s < -t$. Then $s + t < 0$ so $s \notin -x$, and thus $-x$ is bounded below.

By construction, if $t \in -x$ and $s > t$ then $s \in -x$. Last, if $t \in -x$ then there exists $t \in \mathbb{Q}$ such that $t > s$ and $\forall u \in x \quad t + u > 0$. So $\frac{s+t}{2} \in -x$ and thus, $-x$ does not have a smallest element. Hence, $-x$ is indeed a cut.

By construction, $-x + x \subseteq 0_{\mathbb{R}}$. It is thus sufficient to prove the reverse inclusion to conclude that $(\mathbb{R}, +_{\mathbb{R}})$ is a group. Let $s \in 0_{\mathbb{R}}$, i.e. $s > 0$. There exists $t \in x$ such that $t - s \notin x$: indeed, pick $a \in x$; the set $\{n \in \mathbb{N} : a - ns \notin x\}$ is not empty since x is a cut (if it was empty, then $a - ns \in x$ for all $n \in \mathbb{N}$, and since for any $q \in \mathbb{Q}$, there exists $n \in \mathbb{N}$ such that $a - ns < q$ as $\mathbb{Q}$ is Archimedean, we would conclude $q \in x$ i.e. $x = \mathbb{Q}$, which is not a cut). Let $m = \min\{n \in \mathbb{N} : a - ns \notin x\}$ and set $t = a - (n-1)s$ (note: $n \geqslant 1$ since $a \in x$). By definition, $t \in x$ yet $t - s \notin x$.

Now, let $u \in x$. Since x is a cut and $t - s \notin x$, we have $u > t - s$ (as cuts are closed upward). Then $s - t + u > s - t + (t - s) = 0$. Hence, $s - t \in -x$. Since $s = \underbrace{s-t}_{\in -x} + \underbrace{t}_{\in x}$, we conclude $s \in -x +_{\mathbb{R}} x$ as desired. So $0_{\mathbb{R}} \subseteq -x +_{\mathbb{R}} x$, and thus $-x + x = 0_{\mathbb{R}}$. □

Theorem 1.15. *If x, y, z are cuts with $x \leqslant y$, then $x +_{\mathbb{R}} z \leqslant y +_{\mathbb{R}} z$.*

Proof. Let $t \in y +_{\mathbb{R}} z$. There exist $u \in y$ and $v \in z$ such that $t = u + v$. Since $x \leqslant y$, i.e. $y \subseteq x$, we conclude $u \in x$ and thus $t \in x +_{\mathbb{R}} z$. So $y +_{\mathbb{R}} z \subseteq x +_{\mathbb{R}} z$, i.e. $x +_{\mathbb{R}} z \leqslant y +_{\mathbb{R}} z$. □

Notation 1.16. Let $\mathbb{R}_+ = \{x \in \mathbb{R} : 0_{\mathbb{R}} \leqslant x\}$.

Theorem 1.17. *If, for $x, y \in \mathbb{R}_+$, we define:*

$$xy = \{st : s \in x, t \in y\}$$

then $xy \in \mathbb{R}_+$. Moreover, this operation is associative and commutative on $\mathbb{R}_+$, and $(\mathbb{R}_+ \setminus \{0\}, \cdot)$ is an Abelian group.

Proof. Let $x, y \in \mathbb{R}_+$. It is immediate that $xy = yx$.

It is also clear that $xy \neq \emptyset$, and that xy has 0 as a lower bound. Moreover, if $t, q \in \mathbb{Q}$, $t \in xy$ and $q > t$, then by construction, $t = uv$ with $u \in x$ and $v \in y$. By assumption, $u > 0$, and thus $\frac{t}{u} = v$. Thus $\frac{q}{u} > v$. Therefore, $\frac{q}{u} \in y$ as y is a cut. Therefore, $q = u\frac{q}{u} \in xy$.

Last, if $s \in xy$ then $s = uv$ for $u \in x$ and $v \in y$. Since x does not have a smallest element, there exists $q \in x$ with $q < u$, so $qv < uv$, and by definition, $qv \in xy$ and $qv < s$. So xy has no smallest element.

So xy is indeed a cut.

Associativity is proven as with the sum.

Let $x \in \mathbb{R}_+$. Let $s \in 1_{\mathbb{R}}$, i.e. $s > 1$. Let $t \in x$. Then $ts > t$, and since $t \in x$ and x is cut, $ts \in x$. So $x1_{\mathbb{R}} \subseteq x$. On the other hand, let $t \in x$. Since x is a cut, there exists $u \in x$ with $u < t$. Now, $t = \frac{t}{u}u$, with $\frac{t}{u} > 1$ i.e. $\frac{t}{u} \in 1_{\mathbb{R}}$, and $u \in x$, so $t \in x1_{\mathbb{R}}$. Thus $x = x1_{\mathbb{R}}$.

Last, let $x > 0_{\mathbb{R}}$. First, note that $x > 0_{\mathbb{R}}$ implies that there exists $u \in \mathbb{Q}, u > 0$ with $\forall t \in x \quad t \geqslant u$. Indeed otherwise, for any $\varepsilon > 0$, there would exist $t \in x$ with $t < \varepsilon$; in turn this implies, as x is a cut, that $\varepsilon \in x$. We see then that $x = 0_{\mathbb{R}}$, which contradicts $x > 0_{\mathbb{R}}$.

Let

$$x^{-1} = \{s \in \mathbb{Q} : \exists t \in \mathbb{Q} \quad \forall v \in x \quad s > t \text{ and } sv > 1\}.$$

Of course, x^{-1} is not empty since $x > 0$. Moreover, x^{-1} is bounded below by 0. If $s \in x$ and $t > s$, then by construction, $t \in x^{-1}$. If $t \in x^{-1}$, then there exists $s \in x^{-1}$ with $t > s$ (indeed, there exists $u \in \mathbb{Q}$ such that $\forall a \in x \quad au > 1$; let $s = \frac{t+u}{2}$). So x^{-1} does not have a smallest element.

Thus x^{-1} is indeed a cut.

Thus by construction, $xx^{-1} \subseteq 1_{\mathbb{R}}$. We now prove the converse inclusion. Let $s > 1$. There exists $n \in \mathbb{N}$ such that, if $m \geqslant n$, then $s > \frac{m}{m-1}$ by the Archimedean property for $\mathbb{Q}$. Let now $t \in x$. Let $q \in \mathbb{Q}$ such that $0 < q < \frac{t}{n}$.

Let $m = \min\{k \geqslant n : kq \in x\}$ — this set is not empty, as if it where, so would x be, as x is a cut (as soon as it contains an element, which it does, it must contain all elements greater). By construction, $(m-1)q \notin x$. Therefore, for any $t \in x$, we have $t > (m-1)q$. Therefore:

$$t\frac{s}{mq} > (m-1)q\frac{s}{mq} = \frac{m-1}{m}s > \frac{m-1}{m}\frac{m}{m-1} = 1.$$

Therefore, $\frac{s}{mq} \in x^{-1}$, and thus

$$s = \underbrace{mq}_{\in x}\ \underbrace{\frac{s}{mq}}_{\in x^{-1}}$$

so $s \in xx^{-1}$. In all, we have showed that $xx^{-1} = 1_{\mathbb{R}}$.

This concludes our proof. $\square$

Theorem 1.18. *If $x, y, z \in \mathbb{R}_+$ then $x(y +_{\mathbb{R}} z) = xy +_{\mathbb{R}} xz$.*

Proof. This follows from the distributivity of the multiplication with respect to the addition on $\mathbb{Q}$, in a manner identical to the proof of associativity. $\square$

We extent the multiplicative structure to $\mathbb{R}$ as follows.

Definition 1.19. Form all $x, y \in \mathbb{R}$,

(1) if $x > 0$ and $y \leqslant 0$ then $xy = x(-y)$,
(2) if $x \leqslant 0$ and $y > 0$ then $xy = (-x)y$,
(3) if $x \leqslant 0$ and $y \leqslant 0$ then $xy = (-x)(-y)$.

Theorem-Definition 1.20. *The field $\mathbb{R}$ is an ordered field which is* Dedekind complete*, i.e with the property that every nonempty subset of $\mathbb{R}$ with a lower bound has an infimum.*

Proof. We have shown that $(\mathbb{R}, +)$ is an Abelian group, that the multiplication on $\mathbb{R}_+$ is associative, commutative, and distributive with respect to the addition, and that $(\mathbb{R}_+ \setminus \{0\}, \cdot)$ is an Abelian group.

Of course, if $x \geqslant 0$ and $y \geqslant 0$ then $xy \geqslant 0$. Moreover, we saw the order $\leqslant$ is linear on $\mathbb{R}$, and every nonempty bounded-below subset of $\mathbb{R}$ has an infimum.

Last, it is a routine exercise to check that the extended multiplication to $\mathbb{R}$ is still associative, commutative and distributive with respect to the addition. Moreover, it is easy to check that if $x < 0$ then $(-x)^{-1}$ is the inverse of x for the multiplication.

As an example, let $x, y, z \in \mathbb{R}$ with $x \geqslant 0$, $y \leqslant 0$, $z \geqslant 0$ such that $y +_{\mathbb{R}} z \geqslant 0$. Then

$$\begin{aligned} xz &= x(z +_{\mathbb{R}} 0_{\mathbb{R}}) \\ &= x\left(\underbrace{z +_{\mathbb{R}} y}_{\geqslant 0} + \underbrace{(-y)}_{\geqslant 0}\right) \\ &= x(z +_{\mathbb{R}} y) + x(-y) \end{aligned}$$

and thus $xz + xy = x(z +_{\mathbb{R}} y)$, as desired. The other cases are handled similarly. $\square$

We conclude by showing that $\mathbb{R}$ is Archimedean.

Theorem 1.21. *Let $A \subseteq \mathbb{R}$ be not empty and bounded below. The number $x \in \mathbb{R}$ is the infimum of A if, and only if x is a lower bound of A, and*

$$\forall \varepsilon > 0 \quad \exists y \in A \quad y < x + \varepsilon.$$

Proof. Assume $x = \inf A$. By definition, x is the largest lower bound of A. Let $\varepsilon > 0$. Since $x + \varepsilon > x$, we conclude that $x + \varepsilon$ is *not* a lower bound of A. Thus, there exists $y \in A$ such that $y < x + \varepsilon$.

Assume that x is a lower bound of A, and that, for all $\varepsilon > 0$, there exists $y \in A$ such that $y < x + \varepsilon$. Let $z > x$. We set $\varepsilon = z - x > 0$. By assumption, there exists $y \in A$ such that $y < x + \varepsilon = z$. Thus, z is not a lower bound for A. Hence, x is the largest lower bound of A. □

Dedekind completeness implies that $\mathbb{R}$ is an *Archimedean field*.

Theorem 1.22. *The infimum of $\left\{\frac{1}{n} : n \in \mathbb{N} \setminus \{0\}\right\}$ in $\mathbb{R}$ is 0.*

Proof. By assumption, $S = \left\{\frac{1}{n} : n \in \mathbb{N} \setminus \{0\}\right\}$ is bounded below by 0 and not empty, so it has an infimum $x = \inf S \geqslant 0$. Note that $x \leqslant \frac{1}{2n}$ for all $n \in \mathbb{N}$, $n > 0$, and thus $x + x \leqslant \frac{1}{n}$ for all $n \in \mathbb{N} \setminus \{0\}$. Since x is the largest lower bound for S, we have $x + x \leqslant x$, i.e. $x \leqslant 0$. We therefore conclude $x = 0$, as claimed. □

We note that given any Archimedean field $\mathbb{F}$, there exists an injective increasing field morphism from $\mathbb{F}$ into $\mathbb{R}$, using the density of $\mathbb{Q}$. Thus $\mathbb{R}$ is the largest Archimedean field as well.

2. Basic topology of $\mathbb{R}$: sets

2.1. Distance from a point to a set.

Definition 2.1. Let $A \subseteq \mathbb{R}$ be any nonempty set. The *distance* $\operatorname{dist}(x, A)$ from $x \in \mathbb{R}$ to A is:

$$\operatorname{dist}(x, A) = \inf \{|x - y| : y \in A\}.$$

Theorem 2.2. *The following properties hold.*

1. $\forall A \subseteq \mathbb{R} \quad A \neq \emptyset \implies 0 \leqslant \operatorname{dist}(x, A)$,
2. $\forall x, y \in \mathbb{R} \quad \operatorname{dist}(x, \{y\}) = |x - y|$,
3. $\forall x \in \mathbb{R} \quad \forall A \subseteq \mathbb{R} \quad A$ *finite, not empty* $\implies \operatorname{dist}(x, A) = \min\{|x - y| : y \in A\}$,
4. $\forall A \subseteq \mathbb{R} \quad \forall x \in \mathbb{R} \quad x \in A \implies \operatorname{dist}(x, A) = 0$,
5. $\forall A, B \subseteq \mathbb{R} \quad A \subseteq B$ *and* $A \neq \emptyset \implies \forall x \in \mathbb{R} \quad \operatorname{dist}(x, B) \leqslant \operatorname{dist}(x, A)$,
6. $\forall A \subseteq \mathbb{R} \quad A \neq \emptyset \implies \forall x, y \in \mathbb{R} \quad |\operatorname{dist}(x, A) - \operatorname{dist}(y, A)| \leqslant |x - y|$.

Proof. By construction, 0 is a lower bound of $\{|x - y| : y \in A\}$ for all $A \subseteq \mathbb{R}$ with $A \neq \emptyset$ and $x \in \mathbb{R}$. So $0 \leqslant \operatorname{dist}(x, A)$, i.e. Assertion (1) holds.

A nonempty finite subset of $\mathbb{R}$ has a smallest element, which is also its infimum, so Assertion (3) holds. In particular, let $x, y \in \mathbb{R}$. By definition, $\inf\{|x - z| : z \in \{y\}\} = \min\{|x - y|\} = |x - y|$. This shows that Assertion (2) holds as well.

If $x \in A$ then $0 = |x - x| \in \{|x - y| : y \in A\}$ so $\operatorname{dist}(x, A) \leqslant 0$ by definition. As $0 \leqslant \operatorname{dist}(x, A)$, we conclude that $\operatorname{dist}(x, A) = 0$. Thus, Assertion (4) holds.

Let now $A, B \subseteq \mathbb{R}$ with $A \subseteq B$ and $A \neq \emptyset$. Let $x \in \mathbb{R}$. Since $A \subseteq B$, we conclude:

$$\{|x-y| : y \in A\} \subseteq \{|x-y| : y \in B\}$$

and thus:

$$\operatorname{dist}(x, A) = \inf \{|x-y| : y \in A\} \geqslant \{|x-y| : y \in B\} = \operatorname{dist}(x, B).$$

Last, let $A \subseteq \mathbb{R}$, $A \neq \emptyset$, and let $x, y \in \mathbb{R}$. Let now $z \in A$. Since:

$$\operatorname{dist}(x, A) \leqslant |x-z| \leqslant |x-y| + |y-z|,$$

and since $z \in A$ is arbitrary in A, we conclude that $\operatorname{dist}(x, A) - |x-y| \leqslant \operatorname{dist}(y, A)$. Therefore:

$$\operatorname{dist}(x, A) - \operatorname{dist}(y, A) \leqslant |x-y|.$$

Similarly, $\operatorname{dist}(y, A) - \operatorname{dist}(x, A) \leqslant |x-y|$, and thus:

$$|\operatorname{dist}(x, A) - \operatorname{dist}(y, A)| \leqslant |x-y|,$$

which completes our proof. $\square$

2.2. **Closure of a set.**

In general, $\operatorname{dist}(x, A) = 0$ does *not* imply that $x \in A$. We introduce the following concept.

Definition 2.3. The *closure* $\operatorname{cl}(A)$ of a nonempty subset $A \subseteq \mathbb{R}$ is

$$\operatorname{cl}(A) = \{x \in \mathbb{R} : \operatorname{dist}(x, A) = 0\}.$$

We set $\operatorname{cl}(\emptyset) = \emptyset$.

We thus have defined an operator on the power set of $\mathbb{R}$, called the closure operator. To study this operator, we present a standard technique to prove inequalities, which we will use very often.

Theorem 2.4. *Let $a, b \in \mathbb{R}$. If there exists $M > 0$ such that, for all $\varepsilon \in (0, M)$, we have $a \leqslant b + \varepsilon$, then $a \leqslant b$.*

Proof. If $a > b$, let $\varepsilon = \min\{M, \frac{a-b}{2}\} > 0$. Then $b + \varepsilon \leqslant \frac{b+a}{2} < a$. Our theorem holds by contraposition. $\square$

We remark that, as a corollary, if $\forall \varepsilon > 0 \quad a \leqslant b + \varepsilon$, then $a \leqslant b$.

We now state the core properties of the closure operator.

Theorem 2.5. *The closure operator has the following properties.*

(1) $\operatorname{cl}(\emptyset) = \emptyset$,
(2) $\forall A \subseteq \mathbb{R} \quad A \subseteq \operatorname{cl}(A)$,
(3) $\forall A \subseteq \mathbb{R} \quad \operatorname{cl}(\operatorname{cl}(A)) = \operatorname{cl}(A)$,
(4) $\forall A, B \subseteq \mathbb{R} \quad A \subseteq B \implies \operatorname{cl}(A) \subseteq \operatorname{cl}(B)$,
(5) $\forall A, B \subseteq \mathbb{R} \quad \operatorname{cl}(A \cup B) = \operatorname{cl}(A) \cup \operatorname{cl}(B)$.

Proof. Assertion (1) is by definition.

If $A = \emptyset$, then all assertions are trivial (and Assertion (5) is trivial if $B = \emptyset$ as well). We assume henceforth that $A \neq \emptyset$.

If $x \in A$, then $\operatorname{dist}(x, A) = 0$ by Theorem (2.2), and thus $x \in \operatorname{cl}(A)$. So $A \subseteq \operatorname{cl}(A)$.

In particular, $\mathrm{cl}(A) \subseteq \mathrm{cl}(\mathrm{cl}(A))$. If $x \in \mathrm{cl}(\mathrm{cl}(A))$ then $\mathrm{dist}(x, \mathrm{cl}(A)) = 0$. Let $\varepsilon > 0$. There exists $y \in \mathrm{cl}(A)$ such that $|x - y| < \frac{\varepsilon}{2}$. By definition of $\mathrm{cl}(A)$, we have $\mathrm{dist}(y, A) = 0$, so there exists $z \in A$ such that $|y - z| < \frac{\varepsilon}{2}$. Therefore

$$\mathrm{dist}(x, A) \leqslant |x - z| \leqslant |x - y| + |y - z| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Therefore $\mathrm{dist}(x, A) < \varepsilon$ for all $\varepsilon > 0$ and therefore, $\mathrm{dist}(x, A) = 0$. So $x \in \mathrm{cl}(A)$. Therefore, we have proven that $\mathrm{cl}(\mathrm{cl}(A)) \subseteq \mathrm{cl}(A)$, and thus $\mathrm{cl}(A) = \mathrm{cl}(\mathrm{cl}(A))$.

Now, if $A \subseteq B$ and $x \in \mathrm{cl}(A)$ then $\mathrm{dist}(x, A) = 0$ by definition of the closure, and since $A \subseteq B$, by Theorem (2.2), we also have $0 \leqslant \mathrm{dist}(x, B) \leqslant \mathrm{dist}(x, A) = 0$ so $x \in \mathrm{cl}(B)$. Thus $\mathrm{cl}(A) \subseteq \mathrm{cl}(B)$.

Let now $A, B \subseteq E$. First, note that since $A \subseteq A \cup B$ so $\mathrm{cl}(A) \subseteq \mathrm{cl}(A \cup B)$. Similarly, $B \subseteq A \cup B$ so $\mathrm{cl}(B) \subseteq \mathrm{cl}(A \cup B)$. Therefore, $\mathrm{cl}(A) \cup \mathrm{cl}(B) \subseteq \mathrm{cl}(A \cup B)$.

On the other hand, let $x \in \mathrm{cl}(A \cup B)$. Thus $\mathrm{dist}(x, A \cup B) = 0$. Assume that $x \notin \mathrm{cl}(A)$, so that $\mathrm{dist}(x, A) > 0$.

Let $\varepsilon \in (0, \mathrm{dist}(x, A))$. There exists $y \in A \cup B$ such that $\mathrm{dist}(x, y) < \varepsilon$. Since $\mathrm{dist}(x, y) < \mathrm{dist}(x, A)$, we conclude that $y \in B$. Thus $\mathrm{dist}(x, B) < \varepsilon$ for all $\varepsilon \in (0, \mathrm{dist}(x, A))$ and thus $0 \leqslant \mathrm{dist}(x, B) \leqslant \inf(0, \mathrm{dist}(x, A)) = 0$, so $x \in \mathrm{cl}(B)$. Hence, if $x \in \mathrm{cl}(A \cup B)$, then $x \in \mathrm{cl}(A) \cup \mathrm{cl}(B)$.

In conclusion, $\mathrm{cl}(A) \cup \mathrm{cl}(B) = \mathrm{cl}(A \cup B)$. This completes our theorem. □

Corollary 2.6. *If $A, B \subseteq \mathbb{R}$, then*

$$\mathrm{cl}(A \cap B) \subseteq \mathrm{cl}(A) \cap \mathrm{cl}(B).$$

Proof. Since $A \cap B \subseteq A$, we conclude by Theorem (2.5) that $\mathrm{cl}(A \cap B) \subseteq \mathrm{cl}(A)$. Similarly, $\mathrm{cl}(A \cap B) \subseteq \mathrm{cl}(B)$. Thus our theorem is proven. □

Remark 2.7. We note that if an operator on the power set of some set satisfies (5) in Theorem (2.5), then it automatically satisfies (4) as well (since $A \subseteq B$ implies $A \cup B = B$, we have $\mathrm{cl}(A) \cup \mathrm{cl}(B) = \mathrm{cl}(A \cup B) = \mathrm{cl}(B)$ so $\mathrm{cl}(A) \subseteq \mathrm{cl}(B)$); it is however more natural for us to follow the path of proof of Theorem (2.5). Properties (1),(2),(3) and (5)) together define what is called a Kuratoswky closure operator.

We now compute the closure of some basic important sets.

Theorem 2.8. *If $F \subseteq \mathbb{R}$ is finite, then* $\mathrm{cl}(F) = F$.

Proof. We have $F \subseteq \mathrm{cl}(F)$. On the other hand, if $\mathrm{dist}(x, F) = 0$ then $x \in F$, so $\mathrm{cl}(F) \subseteq F$. □

Theorem 2.9. $\mathrm{cl}(\emptyset) = \emptyset$ *and* $\mathrm{cl}(\mathbb{R}) = \mathbb{R}$.

Proof. $\mathrm{cl}(\emptyset) = \emptyset$ by definition. On the other hand, $\mathbb{R} \subseteq \mathrm{cl}(\mathbb{R}) \subseteq \mathbb{R}$ so $\mathrm{cl}(\mathbb{R}) = \mathbb{R}$. □

Definition 2.10. A subset $D \subseteq \mathbb{R}$ of $\mathbb{R}$ is dense in $\mathbb{R}$ when $\mathrm{cl}(D) = \mathbb{R}$.

Theorem 2.11. *The closure of $\mathbb{Q}$ is $\mathbb{R}$. The closure of $\mathbb{R} \setminus \mathbb{Q}$ is $\mathbb{R}$.*

Proof. First, $\mathbb{Q} \subseteq \mathrm{cl}(\mathbb{Q})$. Let $x \in \mathbb{R}$. Let $\varepsilon > 0$. There exists $q \in \mathbb{Q}$ such that $x < q < x + \varepsilon$ by Theorem (1.8). Therefore, $\mathrm{dist}(x, \mathbb{Q}) \leqslant |x - q| < \varepsilon$. As $\varepsilon > 0$ is arbitrary, we conclude that $\mathrm{dist}(x, \mathbb{Q}) = 0$, i.e. $x \in \mathrm{cl}(\mathbb{Q})$.

The same reasoning applies to $\mathbb{R} \setminus \mathbb{Q}$. □

Theorem 2.12. *If $a \in \mathbb{R}$ then*

$$\mathrm{cl}\,((-\infty,a)) = \mathrm{cl}\,((-\infty,a]) = (-\infty,a] \text{ and } \mathrm{cl}\,((a,\infty)) = \mathrm{cl}\,([a,\infty)) = [a,\infty).$$

If $a \leqslant b \in \mathbb{R}$ then

$$\mathrm{cl}\,((a,b)) = \mathrm{cl}\,([a,b)) = \mathrm{cl}\,((a,b]) = \mathrm{cl}\,([a,b]) = [a,b].$$

Proof. Let $a \in \mathbb{R}$. If $b > a$, then $\mathrm{dist}(b,(-\infty,a]) \geqslant |b-a| > 0$, and therefore $b \notin \mathrm{cl}\,((-\infty,a])$; henceforth $\mathrm{cl}\,((-\infty,a]) \subseteq (-\infty,a]$. Since $(-\infty,a] \subseteq \mathrm{cl}\,((-\infty,a])$, we conclude $\mathrm{cl}\,((-\infty,a]) = (-\infty,a]$. A similar argument shows that $\mathrm{cl}\,([a,\infty)) = [a,\infty)$.

Let now $b > a$. Of course, $(a,b) \subseteq (-\infty,b]$, so $\mathrm{cl}\,((a,b)) \subseteq \mathrm{cl}\,((-\infty,b]) = (-\infty,b]$. Similarly, $\mathrm{cl}\,((a,b)) \subseteq [a,\infty)$. Hence $\mathrm{cl}\,((a,b)) \subseteq [a,b]$. Of course, $(a,b) \subseteq \mathrm{cl}\,((a,b))$.

Let now $\varepsilon > 0$. Let $c = \min\left\{b-\varepsilon, \frac{a+b}{2}\right\}$. By construction, $c \in (a,b)$, and $\mathrm{dist}(b,(a,b)) \leqslant |b-c| \leqslant \varepsilon$. Therefore, $\mathrm{dist}(b,(a,b)) = 0$. Similarly, $\mathrm{dist}(a,(a,b)) = 0$. Hence $\mathrm{cl}\,((a,b)) = [a,b]$.

Now, $(-\infty,a) = (-\infty,a-1) \cup (a-2,a)$ so

$$\mathrm{cl}\,((-\infty,a)) = \mathrm{cl}\,((-\infty,a-1)) \cup \mathrm{cl}\,((a-2,a)) = (-\infty,a-1] \cup [a-2,a] = (-\infty,a].$$

Similarly, $\mathrm{cl}\,((a,\infty)) = [a,\infty)$.

Last, if $(a,b) \subseteq I \subseteq [a,b]$, then $[a,b] = \mathrm{cl}\,((a,b)) \subseteq \mathrm{cl}\,(I) \subseteq \mathrm{cl}\,([a,b]) = [a,b]$, as claimed. □

We note that taking the closure of a bounded interval does not increase its diameter; this is a general observation.

Definition 2.13. If $A \subseteq \mathbb{R}$ is bounded and not empty, then the diameter $\mathrm{diam}\,(A)$ of A is defined as $\mathrm{diam}\,(A) = \sup\{|x-y| : x,y \in A\}$.

Theorem 2.14. *If $A \subseteq \mathbb{R}$ is a bounded subset of $\mathbb{R}$, so is $\mathrm{cl}\,(A)$, and moreover:*

$$\mathrm{diam}\,(\mathrm{cl}\,(A)) = \mathrm{diam}\,(A).$$

Proof. Let $x,y \in \mathrm{cl}\,(A)$. Let $\varepsilon > 0$. There exists $x_\varepsilon, y_\varepsilon \in A$ such that $|x-x_\varepsilon| < \frac{\varepsilon}{2}$ and $|y-y_\varepsilon| < \frac{\varepsilon}{2}$. Thus:

$$|x-y| \leqslant |x-x_\varepsilon| + |x_\varepsilon - y_\varepsilon| + |y_\varepsilon - y| \leqslant \mathrm{diam}\,(A) + \varepsilon.$$

As $\varepsilon > 0$ is arbitrary, we conclude that $|x-y| \leqslant \mathrm{diam}\,(A)$. Therefore, $\mathrm{diam}\,(\mathrm{cl}\,(A)) \leqslant \mathrm{diam}\,(A)$. In particular, $\mathrm{cl}\,(A)$ is bounded in $\mathbb{R}$.

Since $A \subseteq \mathrm{cl}\,(A)$, it is immediate by Definition (2.13) that $\mathrm{diam}\,(A) \leqslant \mathrm{diam}\,(\mathrm{cl}\,(A))$, thus concluding our proof. □

2.3. **Closed Sets.** The fixed points for the closure operator are called *closed sets*.

Definition 2.15. A subset $A \subseteq \mathbb{R}$ is *closed* in $\mathbb{R}$ when $\mathrm{cl}\,(A) = A$.

A subset A of $\mathbb{R}$ is thus closed if and only if:

$$\forall x \in \mathbb{R} \quad \mathrm{dist}(x,A) = 0 \iff x \in A.$$

Example 2.16. By Theorem (2.9), both $\emptyset$ and $\mathbb{R}$ are closed.

Example 2.17. By Theorem (2.12), for all $a \leqslant b$, the intervals $(-\infty,a]$, $[a,\infty)$ and $[a,b]$ are closed — these are, in fact, all closed intervals.

Example 2.18. Every finite subset of $\mathbb{R}$ is closed by Theorem (2.8).

Theorem 2.19. *A subset $F \subseteq \mathbb{R}$ of $\mathbb{R}$ is closed if and only if there exists a subset $A \subseteq \mathbb{R}$ of $\mathbb{R}$ such that $F = \operatorname{cl}(A)$.*

Proof. Let $A \subseteq \mathbb{R}$. By Theorem (2.5), $\operatorname{cl}(\operatorname{cl}(A)) = \operatorname{cl}(A)$ so $\operatorname{cl}(A)$ is closed by Definition (2.15).

By Definition (2.15), if $F \subseteq \mathbb{R}$ is closed then $F = \operatorname{cl}(F)$. This completes our proof. □

Closed sets form a lattice for the inclusion.

Theorem 2.20. *The following assertions hold.*

1. *$\emptyset$ and $\mathbb{R}$ are closed subsets of $\mathbb{R}$.*
2. *If $F, G \subseteq \mathbb{R}$ are closed then $F \cup G$ is closed.*
3. *If $\mathscr{F} \subseteq 2^{\mathbb{R}}$ is a nonempty set of closed subsets of $\mathbb{R}$, then:*
$$\bigcap \mathscr{F} := \{x \in \mathbb{R} : \forall F \in \mathscr{F} \quad x \in F\}$$
is closed.

Proof. By definition, $\operatorname{cl}(\emptyset) = \emptyset$, and by Theorem (2.9), $\operatorname{cl}(\mathbb{R}) = \mathbb{R}$.

Let $F, G \subseteq E$ be closed subsets of $\mathbb{R}$. Then
$$\operatorname{cl}(F \cup G) = \operatorname{cl}(F) \cup \operatorname{cl}(G) = F \cup G,$$
and thus $F \cup G$ is closed.

Let $\mathscr{F}$ be a subset of the power set $2^{\mathbb{R}}$ whose elements are all closed subsets of $\mathbb{R}$. We first note that $\bigcap \mathscr{F} \subseteq \operatorname{cl}(\bigcap \mathscr{F})$. On the other hand, if $F \in \mathscr{F}$ then $\bigcap \mathscr{F} \subseteq F$. Therefore, $\operatorname{cl}(\bigcap \mathscr{F}) \subseteq \operatorname{cl}(F) = F$. Consequently, as $F \in \mathscr{F}$ was arbitrary, we conclude that $\operatorname{cl}(\bigcap \mathscr{F}) \subseteq \bigcap \mathscr{F}$.

Therefore, $\operatorname{cl}(\bigcap \mathscr{F}) = \bigcap \mathscr{F}$, i.e. by definition, $\bigcap \mathscr{F}$ is closed in $\mathbb{R}$. □

Remark 2.21. If $(F_j)_{j \in J}$ is a family of closed subsets of $\mathbb{R}$ then $\bigcap_{j \in J} F_j = \bigcap \{F_j : j \in J\}$ is closed by Theorem (2.20).

Theorem 2.22. *If $A \subseteq \mathbb{R}$ is a subset of $\mathbb{R}$ then:*
$$\operatorname{cl}(A) = \bigcap \{F \subseteq \mathbb{R} : F \textit{ is closed and } A \subseteq F\}.$$
In particular, $\operatorname{cl}(A)$ is the smallest closed subset of $\mathbb{R}$ containing A.

Proof. If $A \subseteq \mathbb{R}$ then $\operatorname{cl}(\operatorname{cl}(A)) = \operatorname{cl}(A)$ and thus $\operatorname{cl}(A)$ is closed. Moreover $A \subseteq \operatorname{cl}(A)$. So in particular:
$$\bigcap \{F \subseteq \mathbb{R} : F \text{ is closed and } A \subseteq F\} \subseteq \operatorname{cl}(A).$$

Let $F \subseteq \mathbb{R}$ be a closed subset of $\mathbb{R}$ containing A. Since $A \subseteq F$, we have $\operatorname{cl}(A) \subseteq \operatorname{cl}(F) = F$. So $\operatorname{cl}(A)$ is the smallest closed subset of $\mathbb{R}$ containing A.

In particular:
$$\operatorname{cl}(A) \subseteq \bigcap \{F \subseteq \mathbb{R} : F \text{ is closed and } A \subseteq F\}.$$
This concludes our proof. □

Theorem 2.23. *If $F \subseteq \mathbb{R}$ is closed, bounded above and not empty then* $\sup F \in F$. *Similarly, if F is closed, bounded below and not empty, then* $\inf F \in F$.

Proof. Assume F is not empty, closed and bounded above. For all $\varepsilon > 0$, there exists $x \in F$ such that $\sup F - \varepsilon < x \leqslant \sup F$, so $\operatorname{dist}(\sup F, F) \leqslant |\sup F - x| < \varepsilon$. Hence $\operatorname{dist}(\sup F, F) = 0$, so $\sup F \in \operatorname{cl}(F) = F$.

The same reasoning applies to show $\inf F \in F$ whenever F is closed, not empty and bounded below. □

While closed sets form a sub-lattice of $2^{\mathbb{R}}$, they do not form a Boolean subalgebra: indeed, $\mathbb{R} \setminus [a,b] = (-\infty, a) \cup (b, \infty)$ is not closed since $\operatorname{cl}((-\infty,a) \cup (b,\infty)) = \operatorname{cl}((-\infty,a)) \cup \operatorname{cl}((b,\infty)) = (-\infty,a] \cup [b,\infty)$. We now study the geometry of complements of closed sets.

2.4. **Open Sets.**

Definition 2.24. A subset U of $\mathbb{R}$ is *open* when $\mathbb{R} \setminus U$ is closed in $\mathbb{R}$.

The set of all open sets of a $\mathbb{R}$ satisfies properties inherited from the structure of the closed sets.

Theorem 2.25. *The following assertions hold.*

1. *$\emptyset$ and $\mathbb{R}$ are open subsets of $\mathbb{R}$.*
2. *If $U, V \subseteq \mathbb{R}$ are open then $U \cap V$ is open.*
3. *If $\mathscr{U} \subseteq 2^{\mathbb{R}}$ is a nonempty set of open subsets of $\mathbb{R}$, then:*
$$\bigcup \mathscr{U} := \{x \in \mathbb{R} : \exists U \in \mathscr{U} \quad x \in U\}$$
is open.

Proof. This follows from Theorem (2.20) by taking complements. □

Definition 2.26. The collection of all open subsets of $\mathbb{R}$ is called the *topology* of $\mathbb{R}$.

A fundamental example of open subset of $\mathbb{R}$ is given by open intervals.

Theorem 2.27. *For all $a \leqslant b$, the open interval (a,b) is open in $\mathbb{R}$.*

Proof. Simply note that $\mathbb{R} \setminus (a,b) = (-\infty, a] \cup [b, \infty)$ is the union of two closed subsets, which is closed by Theorem (2.20). □

A simple but useful consequence of Theorem (2.27) is the *Hausdorff property* of the topology of $\mathbb{R}$.

Theorem 2.28. *If $x, y \in \mathbb{R}$ and $x \neq y$, then there exist two* disjoint *open subsets U and V such that $x \in U$ and $y \in V$.*

Proof. Let $\delta = \frac{|x-y|}{2} > 0$. Set $U = (x-\delta, x+\delta)$ and $V = (y-\delta, y+\delta)$. □

The main characterization of open sets is the following theorem, which illustrates the important role, in $\mathbb{R}$, of open intervals.

Theorem 2.29. *A subset $U \subseteq \mathbb{R}$ of $\mathbb{R}$ is open if and only if:*

$$\forall x \in U \quad \exists \delta_x > 0 \quad (x - \delta_x, x + \delta_x) \subseteq U. \tag{2.1}$$

Proof. Let $x \in U$. Thus $x \notin \mathbb{R} \setminus U$. Now, by Definition (2.24), $\mathbb{R} \setminus U = \operatorname{cl}(\mathbb{R} \setminus U)$. Therefore, $\operatorname{dist}(x, \mathbb{R} \setminus U) > 0$, by Definition (2.15). Let $\delta_x = \operatorname{dist}(x, \mathbb{R} \setminus U)$.

Let $y \in (x - \delta_x, x + \delta_x)$. Then $|x - y| < \delta_x$, so $y \notin \mathbb{R} \setminus U$, therefore $y \in U$. Therefore, if U is open, then Assertion (2.1) holds.

Assume now that Assertion (2.1) holds for some subset U of $\mathbb{R}$. Thus, for all $x \in U$, there exists $\delta_x > 0$ such that if $y \in E$ and $\operatorname{dist}(x, y) < \delta_x$, then $y \in U$. In other words, for all $x \in U$, there exists $\delta_x > 0$ such that $(x - \delta_x, x + \delta_X) \subseteq U$. Then $U = \bigcup_{x \in U}(x - \delta_x, x + \delta_x)$. As $(x - \delta_x, x + \delta_x)$ is open for all $x \in U$, we conclude that U is open in $\mathbb{R}$, by Theorem (2.25). □

Corollary 2.30. *A subset U of $\mathbb{R}$ is open if, and only if,*

$$U = \bigcup \{I \subseteq U : I \text{ is an open interval}\}.$$

Proof. By construction, $\bigcup \{I \subseteq U : I \text{ is an open interval }\} \subseteq U$. The converse inclusion follows from Theorem (2.29). □

We can characterize closure yet again using open sets.

Theorem 2.31. *If $A \subseteq \mathbb{R}$ is a subset $\mathbb{R}$ then:*

$$\operatorname{cl}(A) = \{x \in \mathbb{R} : \forall U \subseteq \mathbb{R} \quad U \text{ open and } x \in U \implies U \cap A \neq \emptyset\}.$$

Metric Method. Let $x \in \operatorname{cl}(A)$. If $V \subseteq \mathbb{R}$ is open in $\mathbb{R}$ and $x \in V$, then there exists $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq V$. Since $\operatorname{dist}(x, A) = 0$, there exists $y \in A$ such that $\operatorname{dist}(x, y) < \delta$, and thus $y \in V \cap A$. So $V \cap A \neq \emptyset$.

If, for all open subset V of $\mathbb{R}$ such that $x \in V$, we have $A \cap V \neq \emptyset$, then for all $\varepsilon > 0$, there exists $y \in A \cap (x - \varepsilon, x + \varepsilon)$, i.e.

$$\operatorname{dist}(x, A) \leqslant |x - y| < \varepsilon.$$

Thus $\operatorname{dist}(x, A) = 0$, i.e. $x \in \operatorname{cl}(A)$. □

Topological Method. Let $x \notin \operatorname{cl}(A)$. Thus $x \in \mathbb{R} \setminus \operatorname{cl}(A)$. If $V = \mathbb{R} \setminus \operatorname{cl}(A)$ then V is an open neighborhood of x in $\mathbb{R}$. On the other hand, $A \subseteq \operatorname{cl}(A)$ so $\mathbb{R} \setminus \operatorname{cl}(A) \subseteq \mathbb{R} \setminus A$ and thus $V \cap A = \mathbb{R} \setminus \operatorname{cl}(A) \cap A \subseteq \mathbb{R} \setminus A \cap A = \emptyset$.

If there exists an open subset $V \subseteq \mathbb{R}$ of $\mathbb{R}$ such that $A \cap V = \emptyset$, then $A \subseteq \mathbb{R} \setminus V$ and $\mathbb{R} \setminus V$ is closed, so $\operatorname{cl}(A) \subseteq \mathbb{R} \setminus V$. In particular, $x \notin \operatorname{cl}(A)$. □

We note the following useful result about closure.

Theorem 2.32. *If $A \subseteq \mathbb{R}$, and if $U \subseteq \mathbb{R}$ is an open set, then* $\operatorname{cl}(A) \cap U \subseteq \operatorname{cl}(A \cap U)$.

Metric proof. Let $a \in \operatorname{cl}(A) \cap U$. Since U is open, there exists $\delta > 0$ such that $(a - \delta, a + \delta) \subseteq U$. Since $a \in \operatorname{cl}(A)$, for all $\varepsilon \in (0, \delta)$, there exists $b \in A$ such that $|b - a| < \varepsilon$, i.e. $b \in U \cap A$. Thus $\operatorname{dist}(a, A \cap U) \leqslant |b - a| < \varepsilon$. We conclude that $\operatorname{dist}(a, A \cap U) = 0$, i.e. $a \in \operatorname{cl}(A \cap U)$. □

Topological proof. Let $x \in \operatorname{cl}(A) \cap U$. Let $V \subseteq \mathbb{R}$ be an open subset of $\mathbb{R}$ with $x \in V$. Since $x \in U \cap V$, and since $x \in \operatorname{cl}(A)$, by Theorem (2.31), we conclude that $(V \cap U) \cap A \neq \emptyset$. Thus, again by Theorem (2.31), we thus have $V \cap (A \cap U) \neq \emptyset$, so $x \in \operatorname{cl}(A \cap U)$. □

2.5. Optional: Isolated points, Accumulation points.

Definition 2.33. Let $A \subseteq \mathbb{R}$. A point $x \in A$ is *isolated* when there exists an open set U of $\mathbb{R}$ such that $U \cap A = \{x\}$.

Definition 2.34. Let $A \subseteq \mathbb{R}$. A point $x \in \mathbb{R}$ is an *accumulation point* of A when for all open subset V of $\mathbb{R}$, if $x \in V$ then $V \cap (A \setminus \{x\}) \neq \emptyset$.

Accumulation points behave, in a sense, in the opposite manner to isolated points.

Theorem 2.35. *Let $A \subseteq \mathbb{R}$. A point x is an accumulation point of A if and only if for all open neighborhood V of A, the set $V \cap A$ is infinite.*

Proof. Suppose $V \cap A$ is finite for some open subset V of $\mathbb{R}$ containing x. Let

$$V \cap A \setminus \{x\} = \{x_1, \ldots, x_n\}.$$

Let $\varepsilon = \operatorname{dist}(x, \{x_1, \ldots, x_m\}) > 0$. We then note that the set W defined by $W = V \cap (x-\varepsilon, x+\varepsilon)$ is open, as the intersection of two open subsets of $\mathbb{R}$. By construction, $x \in W$; moreover, $x_1 \notin W$,…,$x_n \notin W$, so $(A \setminus \{x\}) \cap W = \emptyset$. Thus x is not an accumulation point of A. By contraposition, if x is an accumulation point of A and V is an open neighborhood of x in $\mathbb{R}$, then $(A \setminus \{x\}) \cap V$ is infinite. Therefore, $A \cap V$ is also infinite.

If $A \cap V$ is infinite then $A \setminus \{x\} \cap V$ is infinite as well, and in particular, not empty. So x is an accumulation point of A. □

Theorem 2.36. *If $A \subseteq \mathbb{R}$, if $\mathscr{I}(A)$ is the set of all isolated points of A, and if A' is the set of all accumulation points of A in $\mathbb{R}$ then:*

$$\operatorname{cl}(A) = \mathscr{I}(A) \cup A' \text{ and } \mathscr{I}(A) \cap A' = \emptyset.$$

Proof. If $x \in \mathscr{I}(A)$ then there exists an open neighborhood V of x in $\mathbb{R}$ such that $V \cap A = \{x\}$, so $V \cap (A \setminus \{x\}) = \emptyset$ and thus $x \notin A'$. Thus $\mathscr{I}(A) \cap A' = \emptyset$.

Of course, $\mathscr{I}(A) \subseteq A \subseteq \operatorname{cl}(A)$. On the other hand, let $x \in A'$ be an accumulation point of A. Let V be an open neighborhood of x in $\mathbb{R}$. By definition of A', we have $V \cap A \neq \emptyset$. Thus $x \in \operatorname{cl}(A)$. So $\mathscr{I}(A) \cup A' \subseteq \operatorname{cl}(A)$.

Let now $x \in \operatorname{cl}(A)$. If there exists an open subset V of $\mathbb{R}$ containing x such that $(A \setminus \{x\}) \cap V = \emptyset$ then $A \cap V = \{x\}$ since $A \cap V \neq \emptyset$. Thus $x \in \mathscr{I}(A)$; otherwise $x \in A'$. So $x \in \mathscr{I}(A) \cup A'$. We have shown that $\operatorname{cl}(A) \subseteq \mathscr{I}(A) \cup A'$ and therefore our theorem is proven. □

Theorem 2.37. *If $A \subseteq \mathbb{R}$, then the set $\mathscr{I}(A)$ of isolated points in A is at most countable.*

Proof. If $x \in \mathscr{I}(A)$, then there exists $\delta > 0$ such that $(x-\delta, x+\delta) \cap A = \{x\}$. By density of $\mathbb{Q}$ in $\mathbb{R}$, there exists $(q(x), r(x)) \in \mathbb{Q}^2$ such that $(q(x), r(x)) \subseteq (x-\delta, x+\delta)$ so $A \cap (q(x), r(x)) = \{x\}$. Let now $x, y \in \mathscr{I}(A)$ with $(q(x), r(x)) = (q(y), r(y))$. Then $\{x\} = A \cap (q(x), r(x)) = A \cap (q(y), r(y)) = \{y\}$. So $x \in \mathscr{I}(A) \mapsto (q(x), r(x)) \in \mathbb{Q}^2$ is an injection. Since $\mathbb{Q}^2$ is countable, the set $\mathscr{I}(A)$ is at most countable. □

The set of isolated points need not be closed (for instance, consider $S = \left\{\frac{1}{n+1} : n \in \mathbb{N}\right\}$). On the other hand, the set of accumulation points is closed.

Theorem 2.38. *If $A \subseteq \mathbb{R}$ then the set A' of the accumulation points of A is closed.*

Proof. Let $x \in \mathrm{cl}\,(A')$. Let V be an open subset of $\mathbb{R}$ containing x. By Theorem (2.31), there exists $y \in V \cap A'$. Thus, $y \in A'$ so $V \cap (A \setminus \{y\})$ is infinite. Therefore $V \cap (A \setminus \{x, y\})$ is also infinite, and thus $V \cap (A \setminus \{x\})$ is infinite as well. So $x \in A'$. So $\mathrm{cl}\,(A') = A'$. □

Definition 2.39. A subset $P \subseteq \mathbb{R}$ is *perfect* when $P = P'$, i.e. P is the set of all its accumulation points.

Definition 2.40. A subset $A \subseteq \mathbb{R}$ is discrete when $A = \mathscr{I}(A)$.

By Theorem (2.38), if a set is perfect, it is closed. By Theorem (2.36), a set is perfect if it is closed and has no isolated points.

2.6. **Optional: Interior of a set.** Complementation implements by definition a duality between closed and open subsets. It is natural to ask what the complement of the closure operator is, in terms of open sets. Theorem (2.22) describes the closure of a set A as the smallest closed subset of a metric space containing A, owing to the observation in Theorem (2.20) that arbitrary intersections of closed sets are closed. Noting the duality between Theorem (2.25) and Theorem (2.20), we thus introduce the notion of the *interior* of a set.

We begin with the definition of an interior point, by observing that it can be equivalently described in metric terms or in topological terms.

Theorem 2.41. *Let $A \subseteq \mathbb{R}$ and $x \in A$. The following assertions are equivalent:*

(1) there exists an open subset U of $\mathbb{R}$ with $x \in U$ and $U \subseteq A$,
(2) there exists $\varepsilon > 0$ such that if $y \in \mathbb{R}$ and $|x - y| < \varepsilon$ then $y \in A$.

Proof. If there exists an open set U of $\mathbb{R}$ such that $x \in U$ and $U \subseteq A$, then by Theorem (2.29), there exists $\varepsilon > 0$ such that if $y \in \mathbb{R}$ and $|x - y| < \varepsilon$ then $y \in U$ and thus $y \in A$. Thus Assertion (1) implies Assertion (2).

If Assertion (2) holds, then let $U = (x - \varepsilon, x + \varepsilon)$. By Theorem (2.29), U is open in $\mathbb{R}$. By Assertion (2), $U \subseteq A$, and thus Assertion (1) holds. □

As the topological perspective is more general, we choose it for our definition.

Definition 2.42. Let $A \subseteq \mathbb{R}$. A point $x \in A$ is an interior point of A if there exists an open subset U of $\mathbb{R}$ such that $x \in U$ and $U \subseteq A$.

Definition 2.43. A *neighborhood* A of a point $x \in \mathbb{R}$ is a subset $A \subseteq \mathbb{R}$ of $\mathbb{R}$ such that x is an interior point of A.

Definition 2.44. The *interior* $\mathrm{int}(A)$ of a subset $A \subseteq \mathbb{R}$ of $\mathbb{R}$ is the set of all the interior points of A.

The interior of a set is a sort of dual concept to the closure of a set — seen for instance with the following result, which we invite the reader to compare with Theorem (2.22).

Theorem 2.45. *Let $A \subseteq \mathbb{R}$. The interior $\mathrm{int}(A)$ of A is the largest open subset of $\mathbb{R}$ contained in A.*

Proof. Let:
$$U = \bigcup \{V \subseteq A : V \text{ is open in } \mathbb{R}\} .$$
By Theorem (2.25), the set U is open in $\mathbb{R}$. Moreover, $U \subseteq A$ by construction. If $V \subseteq \mathbb{R}$ is open in $\mathbb{R}$ and $V \subseteq A$ then $V \subseteq U$ by construction. Thus, U is the largest open subset of $\mathbb{R}$ containing A.

If $x \in U$ then x is an interior point of A since $U \subseteq A$. So $U \subseteq \text{int}(A)$.

If $x \in A$ is an interior point of A, then there exists an open subset V of $\mathbb{R}$ such that $x \in V$ and $V \subseteq A$. Thus $V \subseteq U$ and thus $x \in U$. Thus $\text{int}(A) \subseteq U$.

Therefore $U = \text{int}(A)$ as claimed. □

Just as a set is closed if, and only if it equals its closure, we obtain the following characterization of open sets in terms of their interior.

Corollary 2.46. *A subset U of $\mathbb{R}$ is open if, and only if, $U = \text{int}(U)$.*

Proof. Since $\text{int}(U)$ is an open set, if $U = \text{int}(U)$ then U is of course open.

If U is open, then U is an open set contained in itself, and thus $U \subseteq \text{int}(U)$ by Theorem (2.45). Since $\text{int}(U) \subseteq U$ by Definition (2.42), we indeed conclude that $U = \text{int}(U)$. □

The duality between interior and closure is clarified with the following theorem.

Theorem 2.47. *For all $A \subseteq \mathbb{R}$, the following assertion holds:*
$$\text{cl}\,(\mathbb{R} \setminus A) = \mathbb{R} \setminus \text{int}(A) \text{ and } \text{int}(\mathbb{R} \setminus A) = \mathbb{R} \setminus \text{cl}\,(A).$$

Proof. Since $\text{int}(A)$ is open, the set $\mathbb{R}\setminus\text{int}(A)$ is closed. Moreover, since $\text{int}(A) \subseteq A$, we also conclude that $\mathbb{R} \setminus A \subseteq \mathbb{R} \setminus \text{int}(A)$. Therefore, $\text{cl}\,(\mathbb{R} \setminus A) \subseteq \mathbb{R} \setminus \text{int}(A)$ by Theorem (2.22).

Let now $x \in \text{cl}\,(\mathbb{R} \setminus A)$. If $x \in \text{int}(A)$, since $\text{int}(A)$ is open, then by Theorem (2.31), we have $\mathbb{R}\setminus A \cap \text{int}(A) \neq \emptyset$, which contradicts $\text{int}(A) \subseteq A$. So $x \in \mathbb{R}\setminus\text{int}(A)$. Thus, we have proven that $\text{cl}\,(\mathbb{R} \setminus A) = \mathbb{R} \setminus \text{int}(A)$.

We then note that, if $B = \mathbb{R} \setminus A$, then:
$$\mathbb{R} \setminus \text{cl}\,(\mathbb{R} \setminus B) = \mathbb{R} \setminus (\mathbb{R} \setminus \text{int}(B)) = \text{int}(\mathbb{R} \setminus A).$$
This completes our proof. □

Hence, Theorem (2.5) together with Theorem (2.47) gives us:

Theorem 2.48. *The following assertions hold:*

(1) $\text{int}(\emptyset) = \emptyset$,
(2) $\forall A \subseteq \mathbb{R} \quad \text{int}(A) \subseteq A$,
(3) $\forall A \subseteq \mathbb{R} \quad \text{int}(\text{int}(A)) = \text{int}(A)$,
(4) $\forall A \subseteq B \subseteq \mathbb{R} \quad \text{int}(A) \subseteq \text{int(B)}$,
(5) $\forall A, B \subseteq \mathbb{R} \quad \text{int}(A \cap B) = \text{int}(A) \cap \text{int}(B)$,
(6) $\forall A, B \subseteq \mathbb{R} \quad \text{int}(A) \cup \text{int}(B) \subseteq \text{int}(A \cup B)$.

Proof. These assertions are obtained from Theorem (2.5) by taking the complement and using Theorem (2.47). □

If $A \subseteq \mathbb{R}$, then it is quite possible that A is *neither* open *nor* closed.

Example 2.49. The set $[0,1)$ is neither open nor closed — it is a strict subset of its closure $[0,1]$ which is the smallest possible closed set containing $[0,1)$ and it is a strict superset of its interior $[0,1)$, the largest open set contained in $[0,1)$.

The closure and interior operators both provide a mean to replace a set by, respectively, a closed set or an open set, which are best approximations in the sense of the order given on sets by inclusion. We also note that some sets, such as $\emptyset$, are *both* closed and open — such sets are sometimes called *clopen sets*. So in no way does the power set of $\mathbb{R}$ gets partitioned in open sets and closed sets.

3. Basic Topology of $\mathbb{R}$: functions — Continuity and Limit of Functions

3.1. **Continuity on a set.** Informally, a function $f : D \to \mathbb{R}$ is continuous when, for any set A, it maps a point in D which is "almost" in A to a point "almost" in $f(A \cap D)$, which indeed captures the motivation behind the notion of continuity — a function is continuous when it does not jump and create gaps.

Definition 3.1. A function $f : D \to \mathbb{R}$ is continuous on D when, for all subset $A \subseteq D$,

$$f(\mathrm{cl}\,(A) \cap D) \subseteq \mathrm{cl}\,(f(A)).$$

Theorem 3.2. *Let $f : D \to \mathbb{R}$ and $g : F \to \mathbb{R}$ with $f(D) \subseteq F$. If f is continuous on D and g is continuous on F, then $g \circ f$ is continuous over D.*

Proof. Let $A \subseteq D$. Since f is continuous on D, we conclude that $f(\mathrm{cl}\,(A) \cap D) \subseteq \mathrm{cl}\,(f(A))$. Since $f(D) \subseteq F$, we conclude that $f(A) \subseteq F$. Since g is continuous over F, $g(F \cap \mathrm{cl}\,(f(A))) \subseteq \mathrm{cl}\,(g(f(A)))$. Thus $g \circ f(\mathrm{cl}\,(A) \cap D) \subseteq \mathrm{cl}\,(g(f(A)))$. Therefore, $g \circ f$ is continuous over D. $\square$

Theorem 3.3. *If $f : D \to \mathbb{R}$ is continuous over D, and if $B \subseteq D$, then the restriction $f_{|B}$ of f to B is continuous over B.*

Proof. We simply note that for all $A \subseteq \mathbb{R}$,

$$f_{|B}(\mathrm{cl}\,(A) \cap B) = f(\mathrm{cl}\,(A) \cap B) \subseteq f(\mathrm{cl}\,(A) \cap D) \subseteq \mathrm{cl}\,(f(A)).$$

This concludes our proof. $\square$

An interesting proper subset of continuous functions is given by the *Lipschitz functions*:

Definition 3.4. Let $D \subseteq \mathbb{R}$ be a nonempty set. A function $f : D \to \mathbb{R}$ is k-Lipschitz, for some $k \geqslant 0$, when

$$\forall x, y \in D \quad |f(x) - f(y)| \leqslant k|x - y|.$$

Theorem 3.5. *If $f : D \subseteq \mathbb{R}$ is Lipschitz over some $D \subseteq \mathbb{R}$, then f is continuous over D.*

Proof. Let $k > 0$ such that, for all $x, y \in D$, we have $|f(x) - f(y)| \leqslant k|x - y|$. Let $A \subseteq D$ and let $x \in \mathrm{cl}\,(A) \cap D$. Since $\mathrm{dist}(x, A) = 0$, for all $\varepsilon > 0$, there exists $y \in A$ such that $|x - y| < \frac{\varepsilon}{k}$, and thus $\mathrm{dist}(f(x), f(A)) \leqslant |f(x) - f(y)| \leqslant k|x - y| < \varepsilon$. As $\varepsilon > 0$ is arbitrary, we conclude that $\mathrm{dist}(f(x), f(A)) = 0$, i.e. $f(x) \in \mathrm{cl}\,(f(A))$. So f is continuous on D. $\square$

Corollary 3.6. *For any nonempty set $D \subseteq \mathbb{R}$, every constant function over D, the canonical injection functions $x \in D \mapsto x$, and the function $x \in \mathbb{R} \mapsto \mathrm{dist}(x, D)$, are all continuous; thus in particular, the absolute value is continuous as well over $\mathbb{R}$.*

Proof. All the listed functions are 1-Lipschitz (see Theorem (2.2) for dist). □

Continuity can be understood as a type of morphism between classes of closed sets or between topologies. Note that the following result is topological in nature: it only depends on the properties of closure, closed sets and open sets.

Theorem 3.7. *Let $f : D \to \mathbb{R}$. The following assertions are equivalent:*

(1) f is continuous over D.
(2) For all closed subset $F \subseteq \mathbb{R}$ of $\mathbb{R}$, there exists a closed subset G of $\mathbb{R}$ such that $f^{-1}(F) = G \cap D$.
(3) For all open subset $V \subseteq \mathbb{R}$ of $\mathbb{R}$, there exists an open subset U of $\mathbb{R}$ such that $f^{-1}(V) = U \cap D$.

Proof. Assume (1), so $f : D \to \mathbb{R}$ is continuous on D. Let $F \subseteq \mathbb{R}$ be a closed subset of $\mathbb{R}$. Of course, $f^{-1}(F) \subseteq \mathrm{cl}\left(f^{-1}(F)\right)$ and since $f^{-1}(F) \subseteq D$ by definition, we end up with:

$$f^{-1}(F) \cap D \subseteq \mathrm{cl}\left(f^{-1}(F)\right) \cap D.$$

We now prove the converge inclusion. By continuity:

$$\begin{aligned} f(\underbrace{\mathrm{cl}\left(f^{-1}(F)\right) \cap D}_{=:A}) &\subseteq \mathrm{cl}\left(f(A)\right) = \mathrm{cl}\left(f(f^{-1}(F))\right) \\ &\subseteq \underbrace{\mathrm{cl}\left(F\right)}_{f(f^{-1}(F)) \subseteq F} \\ &= F \text{ since } F \text{ is closed.} \end{aligned}$$

So $\mathrm{cl}\left(f^{-1}(F)\right) \subseteq f^{-1}(F)$. In summary, $\mathrm{cl}\left(f^{-1}(F)\right) = f^{-1}(F)$, we obtain (2) by setting $G \coloneqq \mathrm{cl}\left(f^{-1}(F)\right)$.

Assume (2), so for all closed subset $F \subseteq \mathbb{R}$ of $\mathbb{R}$, there exists a closed subset $G \subseteq \mathbb{R}$ of $\mathbb{R}$ such that $f^{-1}(F) = G \cap D$. Let $A \subseteq \mathbb{R}$. Let $G \subseteq \mathbb{R}$ be a closed subset of $\mathbb{R}$ such that $G \cap D = f^{-1}(\mathrm{cl}\left(f(A)\right))$. Then:

$$A \subseteq f^{-1}(f(A)) \subseteq f^{-1}(\mathrm{cl}\left(f(A)\right)) = G \cap D \subseteq G.$$

Since G is closed, we thus have $\mathrm{cl}\left(A\right) \subseteq \mathrm{cl}\left(G\right) = G$. Thus $\mathrm{cl}\left(A\right) \cap D \subseteq G \cap D = f^{-1}(\mathrm{cl}\left(f(A)\right))$. So

$$f(\mathrm{cl}\left(A\right) \cap D) \subseteq \mathrm{cl}\left(f(A)\right).$$

Therefore, (1) holds.

(2) and (3) are equivalent simply by taking complements. □

We now get a metric characterization of continuity: the ε-δ characterization of continuity.

Theorem 3.8. *A function $f : D \to \mathbb{R}$ is continuous on D if, and only if, for all $x \in D$:*

$$\forall \varepsilon > 0 \quad \exists y \in D \quad |y - x| < \delta \implies |f(y) - f(x)| < \varepsilon. \tag{3.1}$$

Proof. Assume Expression (3.1). Let $A \subseteq \mathbb{R}$. Let $y \in f(\operatorname{cl}(A) \cap D)$. Let $x \in \operatorname{cl}(A) \cap D$ such that $y = f(x)$. Let $\varepsilon > 0$. By assumption, there exists $\delta > 0$ such that, if $t \in D$, $|t - x| < \delta$, then $|f(t) - f(x)| < \varepsilon$. Since $x \in \operatorname{cl}(A) \cap D$, we conclude that $\operatorname{dist}(x, A) = 0$. Thus, there exists $t \in A$ with $|t - x| < \delta$. Therefore $f(t) \in f(A)$, and $\operatorname{dist}(f(x), f(A)) \leqslant |f(x) - f(t)| < \varepsilon$. Consequently, $\operatorname{dist}(f(x), f(A)) = 0$. Therefore, $f(x) \in \operatorname{cl}(f(A))$. We thus have shown that $f(\operatorname{cl}(A) \cap D) \subseteq \operatorname{cl}(f(A))$.

Suppose now that for some $x \in D$ and some $\varepsilon > 0$, for all $\delta > 0$, there exists $t \in D$ such that $|t - x| < \delta$ and $|f(t) - f(x)| \geqslant \varepsilon$. Let $A = \{t \in D : |f(t) - f(x)| \geqslant \varepsilon\}$, and note that by our assumption, $A \neq \emptyset$. Moreover, for all $\delta > 0$, there exists $t \in A$ with $|t - x| < \delta$. Thus, $\operatorname{dist}(x, A) < |x - t| < \delta$, and thus $\operatorname{dist}(x, A) = 0$, so $x \in \operatorname{cl}(A)$. Of course, $x \in D$ so $x \in \operatorname{cl}(A) \cap D$.

On the other hand, if $t \in A$, then $|f(t) - f(x)| \geqslant \varepsilon$, so $\operatorname{dist}(f(t), f(A)) \geqslant \varepsilon$. Hence $f(x) \notin \operatorname{cl}(f(A))$. Our proof follows by contraposition. □

We note that Theorem (3.8) gives an easy proof that Lipschitz functions are in fact continuous:

Alternate proof of Theorem (3.5). Let $k > 0$ such that $\forall x, y \in D \quad |f(x) - f(y)| \leqslant k|x - y|$. Let $\varepsilon > 0$. Set $\delta = \frac{\varepsilon}{k} > 0$. If $x, y \in D$ then $|x - y| < \delta$ then

$$|f(x) - f(y)| \leqslant k\delta = k\frac{\varepsilon}{k} = \varepsilon.$$

This concludes our proof by Theorem (3.8). □

Remark 3.9. The choice of δ in the proof of Theorem (3.5) does not depend on the choice of $x \in D$, which is stronger than what we need, and is a manifestation of the fact that Lipschitz functions are in fact uniformly continuous, a concept we will return to later on.

We can use the ε-δ characterization of continuity to recover Theorem (3.7) via metric methods.

Alternate proof of Theorem (3.7). Assume (1). Let $F \subseteq \mathbb{R}$ and let $G = \operatorname{cl}\left(f^{-1}(F)\right)$. By construction, $f^{-1}(F) \subseteq G \cap D$. Let now $x \in G \cap D$. Let $\varepsilon > 0$. Since f is continuous at x, there exists $\delta > 0$ such that, if $t \in D$ and $|t - x| < \delta$, then $|f(t) - f(x)| < \varepsilon$. Since $G = \operatorname{cl}\left(f^{-1}(F)\right)$ and $x \in G$, we conclude $\operatorname{dist}(x, f^{-1}(F)) = 0$, so there exists $t \in f^{-1}(F)$ such that $|t - x| < \delta$. Since $f(t) \in F$, we conclude that $\operatorname{dist}(f(x), F) \leqslant |f(x) - f(t)| < \varepsilon$. As $\varepsilon > 0$ is arbitrary, $f(x) \in \operatorname{cl}(F) = F$ — as F is closed. Thus $x \in f^{-1}(F)$. Thus $f^{-1}(F) = G \cap D$, i.e. (2) holds.

Assume (2). Let $U \subseteq \mathbb{R}$ be an open subset. By (2), there exists a closed subset $G \subseteq \mathbb{R}$ such that

$$D \setminus f^{-1}(U) = f^{-1}(\mathbb{R} \setminus U) = G \cap D$$

since $\mathbb{R} \setminus U$ is closed. Therefore, $f^{-1}(U) = (\mathbb{R} \setminus G) \cap D$, and $V = \mathbb{R} \setminus G$ is open in $\mathbb{R}$. We thus have proven (3).

Assume (3). Let $\varepsilon > 0$ and $x \in D$. Let $U = (f(x) - \varepsilon, f(x) + \varepsilon)$. Since U is open, we conclude that there exists an open subset V of $\mathbb{R}$ such that $f^{-1}(U) = V \cap D$. Of course, $x \in V$. Since V is open, there exists $\delta > 0$ such that $(x - \delta, x + \delta) \subseteq V$. By construction, $f((x - \delta, x + \delta) \cap D) \subseteq f(V \cap D) \subseteq U$. So, for all $t \in D$ such that $|t - x| < \delta$, we conclude that $|f(t) - f(x)| < \varepsilon$. This proves (1).

Our proof is now complete. ☐

Remark 3.10. As a corollary of Theorem (3.7), if $f : U \to \mathbb{R}$, and if U is open, then f is continuous on U if, and only if, the preimage $f^{-1}(V)$ of any open subset of $\mathbb{R}$ is open; if $f : F \to \mathbb{R}$, and if F is closed, then f is continuous over F if, and only if, the preimage $f^{-1}(F)$ of F is closed.

3.2. **Limit at a point.** We start from the characterization of continuity over a set given by Theorem (3.8), and we make it *local* by removing the quantification over the entire set, while allowing for more general limits. We are led to the following definition.

Definition 3.11. Let $f : D \to \mathbb{R}$ be a function over a set D. Let $a \in \mathrm{cl}(D)$ and $l \in \mathbb{R}$. We say that f converges to l at a along D, when

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall t \in D \quad |t - a| < \delta \implies |f(t) - l| < \varepsilon.$$

Lemma 3.12. *Let $l, l' \in \mathbb{R}$. If $|l - l'| < \varepsilon$ for all $\varepsilon > 0$, then $l = l'$.*

Proof. Our assumption implies that $|l - l'| = 0$, so $l = l'$. ☐

Theorem 3.13. *Let $f : D \to \mathbb{R}$ and $a \in \mathrm{cl}(D)$. If f converges to l and l' at a along D, then $l = l'$.*

Proof. Let $\varepsilon > 0$. There exists $\delta_1 > 0$ such that, for all $t \in D$, if $|t - a| < \delta_1$, then $|f(t) - l| < \frac{\varepsilon}{2}$. There exists $\delta_2 > 0$ such that, for all $t \in D$, if $|t - a| < \delta_2$, then $|f(t) - l'| < \frac{\varepsilon}{2}$.

Therefore, if $t \in D$ such that $|t - a| < \min\{\delta_1, \delta_2\}$ (which exists since $a \in \mathrm{cl}(D)$), then

$$|l - l'| \leqslant |l - f(t)| + |f(t) - l'| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Thus $l = l'$. ☐

Notation 3.14. If $f : D \to \mathbb{R}$ converges to l at $a \in \mathrm{cl}(D)$, then we write

$$l = \lim_{\substack{x \to a \\ x \in D}} f(x).$$

If $D = \{x \in \mathbb{R} : P(x)\}$ then we can also write

$$l = \lim_{\substack{x \to a \\ P(x)}} f(x).$$

Convention 3.15. Let $f : I \to \mathbb{R}$ where I is some interval, and let $a \in I$. We use the following notations (whenever they are well-defined): $\lim_{x \to a^+} f(x)$ is meant for $\lim_{\substack{x \to a \\ x \in D, x > a}} f(x)$; $\lim_{x \to a^-} f(x)$ is meant for $\lim_{\substack{x \to a \\ x \in D, x < a}} f(x)$; last, $\lim_{x \to a} f(x)$ is meant for $\lim_{\substack{x \to a \\ x \in D, x \neq a}} f(x)$. In each of these notations, one must check that indeed, a is in the closure of the set along which we take our limit. Also note that in all these cases, a is removed from the domain prior to taking the limit; this is important as one can check that if $a \in D$ and f has a limit at a along D then the only possible limit is $f(a)$.

Limits live in the closure of the range of the function under consideration.

Theorem 3.16. *If $f : D \to \mathbb{R}$ converges to l at a along D then $l \in \mathrm{cl}(f(D))$.*

Proof. Let $\varepsilon > 0$. Since f converges to l at a along D, there exists $\delta > 0$ such that if $x \in D$ and $|x-a| < \delta$ then $|f(x) - l| < \varepsilon$. Since $a \in \mathrm{cl}(D)$, there exists $x \in D$ with $|x-a| < \delta$. Therefore, $\mathrm{dist}(l, f(D)) \leqslant |l - f(x)| < \varepsilon$. So $\mathrm{dist}(l, f(D)) = 0$ i.e. $l \in \mathrm{cl}(f(D))$. □

Remark 3.17. It is usually difficult to determine the exact range of a function. However, if $f : D \to \mathbb{R}$ and if we know that $f(D) \subseteq B$, then we can conclude that $\lim_{\substack{x\to a\\ x\in D}} f(x) \in \mathrm{cl}(B)$ by Theorem (3.16), and using its notation.

Limits are indeed a local concept.

Theorem 3.18. *Let $D, A \subseteq \mathbb{R}$. Let $a \in \mathrm{cl}(D \cap A)$. If $f : D \to \mathbb{R}$ has converges to l at a along D, then f converges to l at a along $D \cap A$.*

Proof. Let $\varepsilon > 0$. Since f has limit l at a along D, there exists $\delta > 0$ such that, for all $t \in D$, if $|t-a| < \delta$, then $|f(t) - l| < \varepsilon$. Therefore, for all $t \in D \cap A$, if $|t-a| < \delta$, then $|f(t) - l| < \varepsilon$. □

Theorem 3.19. *Let $D \subseteq \mathbb{R}$. Let $f : D \to \mathbb{R}$, $a \in \mathrm{cl}(D)$ and $l \in \mathbb{R}$. The following assertions are equivalent:*

1. *f has limit l at a along D.*
2. *For all open subset $U \subseteq \mathbb{R}$ with $a \in U$, f has limit l at a along $D \cap U$.*
3. *For all $\delta > 0$, f has limit l at a along $D \cap (a-\delta, a+\delta)$.*
4. *There exists an open subset $U \subseteq \mathbb{R}$ with $a \in U$, f has limit l at a along $D \cap U$.*
5. *There exists $\delta > 0$ such that f has limit l at a along $D \cap (a-\delta, a+\delta)$.*

Proof. Assume (1). Let $\delta > 0$. By (1), for all $\varepsilon > 0$, there exists $\alpha > 0$ such that for all $t \in D$, if $|t-a| < \alpha$, then $|f(x) - l| < \varepsilon$; of course, it follows that for all $t \in D \cap (a-\delta, a+\delta)$, we have $|f(x) - l| < \varepsilon$. So (2) holds.

Assume (1). Let $U \subseteq \mathbb{R}$ be open with $a \in U$. We first note that $a \in \mathrm{cl}(D \cap U)$ by Theorem (2.32). Then (2) follows from Theorem (3.18).

Assume (2). Let $\delta > 0$. Set $U = (a-\delta, a+\delta)$. Applying (2) to the open set U, we have established (3).

Of course, (2) implies (4) and (3) implies (5). Moreover, just as above, (4) implies (5).

Assume (5). Let $\varepsilon > 0$. There exists $\alpha > 0$ such that, for all $x \in D \cap (a-\delta, a+\delta)$, if $|x-a| < \alpha$ then $|f(x) - l| < \varepsilon$. Let $\alpha' = \min\{\delta, \alpha\} > 0$. For all $x \in D$, if $|x-a| < \alpha'$ then $x \in D \cap (a-\delta, a+\delta)$ and $|x-a| < \alpha$, hence $|f(x) - l| < \varepsilon$. So (1) holds, and our proof is complete. □

Definition 3.20. We say that f is *continuous at $x \in D$* when $\lim_{\substack{t\to x\\ t\in D}} f(t) = f(x)$.

Therefore, we note that $f : D \to \mathbb{R}$ *is continuous on D exactly when* $\forall x \in D \quad \lim_{t\to x} f(t) = f(x)$*, by definition.*

3.3. **Limits and Order.**

Theorem 3.21. *Let $D \subseteq \mathbb{R}$. Let $f : D \to \mathbb{R}$, and let $x \in \operatorname{cl}(D)$. If f has a limit at x along D, then there exists an open neighborhood U of x such that $\{f(t) : t \in U \cap D\}$ is bounded, with diameter at most 1.*

Proof. Let $l = \lim_{\substack{t\to x\\ t\in D}} f(t)$. There exists $\delta > 0$ such that if $t \in (x-\delta, x+\delta) \cap D$, then $|l - f(t)| < \frac{1}{2}$. Thus, if $t, t' \in (x-\delta, x+\delta) \cap D$ then

$$|f(t) - f(t')| \leqslant |f(t) - l| + |y - f(t')| < 1.$$

Thus $\operatorname{diam}(f((x-\delta, x+\delta) \cap D))d_F \leqslant 1$. □

Theorem 3.22. *Let $D \subseteq \mathbb{R}$. Let $f : D \to \mathbb{R}$, $g : D \to \mathbb{R}$, and $a \in \operatorname{cl}(D)$. If f and g have limits at a along D, and if*

$$\exists \delta > 0 \quad \forall x \in (a-\delta, a+\delta) \quad f(x) \leqslant g(x)$$

then

$$\lim_{\substack{x\to a\\ x\in D}} f(x) \leqslant \lim_{\substack{x\to a\\ x\in D}} g(x).$$

Proof. Let $\varepsilon > 0$. Writing $l = \lim_{\substack{x\to a\\ x\in D}} f(x)$, there exists $\delta_1 > 0$ such that, if $x \in D$ and $|a - x| < \delta_1$ then $|f(x) - l| < \frac{\varepsilon}{2}$. Writing $l' = \lim_{\substack{x\to a\\ x\in D}} g(x)$, there exists $\delta_2 > 0$ such that, if $x \in D$ and $|a - x| < \delta_2$ then $|g(x) - l'| < \frac{\varepsilon}{2}$.

Consequently, for $x \in D$ with $|x - a| < \min\{\delta_1, \delta_2\}$, we conclude:

$$l - \frac{\varepsilon}{2} \leqslant f(x) \leqslant g(x) \leqslant l' + \frac{\varepsilon}{2},$$

so $l \leqslant l' + \varepsilon$. Since $\varepsilon > 0$ is arbitrary, we conclude $l \leqslant l'$, as claimed. □

Theorem 3.23. *Let $D \subseteq \mathbb{R}$.*

Let $f : D \to \mathbb{R}$, $g : D \to \mathbb{R}$, and $h : D \to \mathbb{R}$. Let $x \in \operatorname{cl}(D)$.

If:

$$\forall t \in D \quad f(t) \leqslant g(t) \leqslant h(t)$$

if f and h have limits at x along D, and if

$$\lim_{\substack{t\to x\\ t\in D}} f(t) = \lim_{\substack{t\to x\\ t\in D}} h(t)$$

then g has a limit at x along D and

$$\lim_{\substack{t\to x\\ t\in D}} g(t) = \lim_{\substack{t\to x\\ t\in D}} f(t).$$

Proof. Let $\varepsilon > 0$. By assumption, there exists $\delta_f > 0$ such that if $t \in D$ and $\operatorname{dist}(t, x) < \delta_f$ then $|l - f(t)| < \varepsilon$. In particular, if $t \in D$ and $\operatorname{dist}(t, x) < \delta_f$ then $f(t) < l + \varepsilon$. Similarly, there exists $\delta_g > 0$ such that if $t \in D$ and $\operatorname{dist}(x, t) < \delta_g$ then $|l - h(t)| < \varepsilon$, so in particular $l - \varepsilon < h(t)$.

Therefore, if $t \in D$ and $\operatorname{dist}(t, x) < \min\{\delta_f, \delta_f\}$ then:

$$l - \varepsilon < f(t) - l < g(t) - l < h(t) - l < l + \varepsilon$$

i.e. $|g(t) - l| < \varepsilon$. Note that $\min\{\delta_f, \delta_f\} > 0$, therefore, our claimed is proven. □

A common way to apply Theorem (4.7) is given in the following corollary. In its proof, we use the obvious, but noteworthy, observation that for any function $f : D \to \mathbb{R}$ from a subset $D \subseteq \mathbb{R}$, if $x \in \mathrm{cl}(D)$ and $y \in \mathbb{R}$, then:

$$\lim_{\substack{t\to x\\ t\in D}} f(t) = y \iff \lim_{\substack{t\to x\\ t\in D}} |f(t) - y| = 0.$$

Corollary 3.24. *Let $D \subseteq \mathbb{R}$ and $x \in \mathrm{cl}(D)$. Let $y \in \mathbb{R}$. If $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$ are two functions such that:*

$$\forall t \in D \quad |f(t) - y| \leqslant g(t)$$

and if $\lim_{\substack{t\to x\\ t\in D}} g(t) = 0$, then

$$\lim_{\substack{t\to x\\ t\in D}} f(t) = y.$$

Proof. By assumption, for all $t \in D$, the following assertion holds:

$$0 \leqslant |f(t) - y| \leqslant g(t).$$

Since $\lim_{\substack{t\to x\\ t\in D}} 0 = \lim_{\substack{t\to x\\ t\in D}} g(t) = 0$, we conclude by Theorem (4.7) that

$$\lim_{\substack{t\to x\\ t\in D}} |f(t) - y| = 0.$$

which in turn, is equivalent to the conclusion of our corollary. ☐

3.4. **Limits and Algebra.**

Theorem 3.25. *Let $D \subseteq \mathbb{R}$ and let $x \in \mathrm{cl}(D)$.*

Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$. If f and g both have a limit at x along D, and if $\lambda \in \mathbb{R}$, then $\lambda f + g$ has a limit at x along D, and:

$$\lim_{\substack{t\to x\\ t\in D}} (\lambda f(t) + g(t)) = \lambda \lim_{\substack{t\to x\\ t\in D}} f(t) + \lim_{\substack{t\to x\\ t\in D}} g(t).$$

Proof. Write $l_f = \lim_{\substack{t\to x\\ t\in D}} f(t)$ and $l_g = \lim_{\substack{t\to x\\ t\in D}} g(t)$.

Let $\varepsilon > 0$.

By definition, there exists $\delta_f > 0$ such that if $t \in D$ and $\mathrm{dist}(t, x) < \delta_f$, we have $|f(x) - l_f| < \frac{\varepsilon}{2(|\lambda|+1)}$.

By definition, there exists $\delta_g > 0$ such that if $t \in D$ and $\mathrm{dist}(t, x) < \delta_g$, we have $|g(x) - l_g| < \frac{\varepsilon}{2}$.

Let $\delta = \min\{\delta_f, \delta_g\}$ and note that $\delta > 0$. Moreover, if $t \in D$ and $\mathrm{dist}(t, x) < \delta$, then:

$$\begin{aligned}
|(\lambda f(t) + g(t)) - (\lambda l_f + l_g)| &\leqslant |\lambda(f(t) - \lambda_f)| + |g(t) - l_g| \\
&= |\lambda||f(t) - \lambda_f| + |g(t) - l_g| \\
&\leqslant |\lambda|\frac{\varepsilon}{2(|\lambda| + 1)} + \frac{\varepsilon}{2} \\
&\leqslant \varepsilon,
\end{aligned}$$

which completes our proof. ☐

Theorem 3.26. *Let $D \subseteq \mathbb{R}$. Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$, and let $x \in \mathrm{cl}(D)$. If f and g both have a limit at x along D, then fg has a limit at x along D, and:*

$$\lim_{\substack{t\to x\\ t\in D}} f(t)g(t) = \lim_{\substack{t\to x\\ t\in D}} f(t) \cdot \lim_{\substack{t\to x\\ t\in D}} g(t).$$

Proof. Write $l_f = \lim_{\substack{t \to x \\ t \in D}} f(t)$ and $l_g = \lim_{\substack{t \to x \\ t \in D}} g(t)$.

Let $\varepsilon > 0$.

Since f has a limit at x along D, there exists $M > 0$ and a $\delta_0 > 0$ such that if $t \in D$ and $\operatorname{dist}(t, x) < \delta_0$ then $|f(t)| \leqslant M$, using Theorem (3.21).

Moreover, there exists $\delta_1 > 0$ such that if $t \in D$ and $\operatorname{dist}(t, x) < \delta_1$ then $|f(t) - l_f| < \frac{\varepsilon}{2(|l_g|+1)}$.

Since g has a limit at x along D, there exists $\delta_3 > 0$ such that if $t \in D$ and $\operatorname{dist}(t, x) < \delta_3$ then $|g(t) - l_g| < \frac{\varepsilon}{2M}$, again using Theorem (3.21).

Let $\delta = \min\{\delta_0, \delta_1, \delta_2\}$ and note that $\delta > 0$. Moreover, if $t \in D$ and $\operatorname{dist}(t, x) < \delta$ then:

$$\begin{aligned} |f(t)g(t) - l_f l_g| &= |f(t)g(t) - f(t)l_g + f(t)l_g - l_f l_g| \\ &\leqslant |f(t)||g(t) - l_g| + |l_g||f(t) - l_f| \\ &\leqslant M\frac{\varepsilon}{2M} + |l_g|\frac{\varepsilon}{2(|l_g| + 1)} \\ &\leqslant \varepsilon, \end{aligned}$$

which concludes our proof. □

Putting together Theorems (3.25) and (3.26), we then obtain a new class of continuous functions: polynomial functions over $\mathbb{R}$. We first record the following basic rules about continuous functions.

Corollary 3.27. *Let $D \subseteq \mathbb{R}$. If $f : D \subseteq F$ and $g : D \to F$ are functions which are continuous at $x \in D$, then $\lambda f + g$ is continuous at x for all $\lambda \in \mathbb{R}$.*

Proof. By assumption, $\lim_{t\to x} f(t) = f(x)$ and $\lim_{t\to x} g(t) = g(x)$. Therefore, by Theorem (3.25), for all $\lambda \in \mathbb{R}$, we have:

$$\lim_{t\to x}(\lambda f(t) + g(t)) = \lambda f(x) + g(x),$$

and thus $f + \lambda g$ is continuous at $x \in D$. □

We are now ready to prove that polynomials are continuous. In fact, polynomials can be proven to be locally Lipschitz, but the following proof is much more natural.

Corollary 3.28. *If P is a polynomial function over $\mathbb{R}$ then P is continuous on $\mathbb{R}$.*

Proof. Of course $X : x \in \mathbb{R} \mapsto x$ is continuous. If for some $n \in \mathbb{N} \setminus \{0\}$, the function $X^n : x \in \mathbb{R} \mapsto x^n$ is continuous, then $X \cdot X^n : x \in \mathbb{R} \mapsto x \cdot x^n = x^{n+1}$ is continuous by Theorem (3.26). Thus by induction, X^n is continuous for all $n \in \mathbb{N} \setminus \{0\}$.

If, for some $n \in \mathbb{N} \setminus \{0\}$, every function $P = \sum_{j=1}^{n} a_j X^j$ is continuous on $\mathbb{R}$, and if $Q : \mathbb{R} \to \mathbb{R}$ is given by $Q = \sum_{j=0}^{n+1} a_j X^j$, then $Q = P + a_{n+1}X^{n+1}$, which is continuous on $\mathbb{R}$ by Theorem (3.25). By induction, we have thus shown that if $P : \mathbb{R} \to \mathbb{R}$ is a polynomial function such that $P(0) = 0$ then P is continuous on $\mathbb{R}$. We conclude, thanks to Theorem (3.25), that all polynomial functions over $\mathbb{R}$ are continuous. □

We next prove that the multiplicative inverse function is also continuous on its domain. Notably, it is sufficient to know that a function has a nonzero limit at a point to conclude that it is nonzero on some neighborhood of that point.

Theorem 3.29. *Let $D \subseteq \mathbb{R}$. If $f : D \to \mathbb{R}$, if $x \in \mathrm{cl}\,(D)$, and if f has a* nonzero *limit at x along D, then:*

(1) there exists an open neighborhood U of x such that $f(y) \neq 0$ for all $y \in U$,
(2) $\lim_{\substack{t\to x\\ t\in D\cap U}} \frac{1}{f(t)} = \frac{1}{\lim_{\substack{t\to x\\ t\in D}} f(t)}$.

Proof. Let $l = \lim_{\substack{t\to x\\ t\in D}} f(t)$.

Let $\alpha = \frac{|l|}{2} > 0$. There exists $\delta_0 > 0$ such that if $t \in D$ such that $\mathrm{dist}(t,x) < \delta_0$ then $|f(t) - l| < \alpha$, and thus in particular $|f(t)| > \alpha$. Thus setting $U = E(x,\delta_0)$, we have proven Assertion (1).

Let $\varepsilon > 0$. There exists $\delta_1 > 0$ such that if $t \in D$ and $\mathrm{dist}(t,x) < \delta_1$ then $|f(t) - l| < \varepsilon |l| \alpha$ (note that $|l|\alpha > 0$). If $t \in D$ and $\mathrm{dist}(x,t) < \min\{\delta_0,\delta_1\} > 0$, then:

$$\begin{aligned}\left|\frac{1}{f(t)} - \frac{1}{l}\right| &= \left|\frac{l - f(t)}{l f(t)}\right| \\ &\leqslant |l - f(t)| \frac{1}{l\alpha} \leqslant \varepsilon.\end{aligned}$$

Thus Assertion (2) is proven as well. □

Corollary 3.30. *Let $D \subseteq \mathbb{R}$. If $f : D \to \mathbb{R}$, $g : D \to \mathbb{R}$, if $x \in \mathrm{cl}\,(D)$, if f has a limit at x along D, and if g has a* nonzero *limit at x along D, then:*

$$\lim_{\substack{t\to x\\ t\in D}} \frac{f(t)}{g(t)} = \frac{\lim_{\substack{t\to x\\ t\in D}} f(t)}{\lim_{\substack{t\to x\\ t\in D}} g(t)}.$$

Proof. By Theorem (3.29), $\lim_{\substack{t\to x\\ t\in D\cap U}} \frac{1}{g(t)} = \frac{1}{\lim_{\substack{t\to x\\ t\in D}} g(t)}$. By Theorem (3.26), we then conclude our theorem. □

Corollary 3.31. *Let $D \subseteq \mathbb{R}$. Let $f : D \to \mathbb{R}$ and $g : D \to \mathbb{R}$. If $x \in D$, if f is continuous at x, if g is continuous at x, and $0 \notin g(E)$, then $\frac{f}{g}$ is continuous at x. In particular, if f and g are both continuous on D, then $\frac{f}{g}$ is continuous on D*

Proof. This is immediate. □

We can thus prove that rational functions are also continuous. Again, rational functions are locally Lipschitz on their domain, but the following proof is again much easier.

Corollary 3.32. *If R is a rational function, then R is continuous on its domain.*

Proof. This follows from Corollary (3.28) and Corollary (3.30). □

3.5. **Composition of Limits.**

Theorem 3.33. *Let $D \subseteq \mathbb{R}$ and $H \subseteq \mathbb{R}$. Let $a \in \mathrm{cl}\,(D)$.*

If $f : D \to \mathbb{R}$ and $g : H \to \mathbb{R}$ are two functions such that:

(1) f converges to l at a along D,
(2) $l \in \mathrm{cl}\,(H)$,
(3) g converges to y at l along H,
(4) $f(D) \subseteq H$,

then $g \circ f$ converges to y at a along D.

Proof. Let $\varepsilon > 0$. There exists $\delta_g > 0$ such that if $t \in H$ and $|t - l| < \delta_g$ then $|g(t) - y| < \varepsilon$. Now there exists $\delta_f > 0$ such that if $x \in D$ and $|x - a| < \delta_f$ then $|f(x) - l| < \delta_g$.

Thus, if $x \in D$ and $|x - a| < \delta_f$, then $f(x) \in H$ and $|f(x) - l| < \delta_g$ and thus $|g(f(x)) - y| < \varepsilon$. □

4. Sequences

Sequences are functions from $\mathbb{N}$ which may be used to characterize all the topological concepts of $\mathbb{R}$, thanks to the notion of convergence of sequences. In essence, convergent sequences provide approximations of their limits. Two core concepts which make sequences helpful are: the ability to prove a sequence has a limit without knowledge of what the limit might be, and the notion of a subsequence.

4.1. Convergence of Sequences.

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of real numbers. If we set $f\left(\frac{1}{n+1}\right) = x_n$ for each $n \in \mathbb{N}$, then we define a function $f : S \to \mathbb{R}$ with $S = \left\{\frac{1}{n+1} : n \in \mathbb{N}\right\}$. Since $\mathbb{R}$ is Archimedean, $0 = \inf S$; thus $0 \in \mathrm{cl}\,(S)$. We thus have a ready-made notion of limit at 0 (along S) for f: the function f converges to l at 0 when for all $\varepsilon > 0$, there exists $\delta > 0$ such that if $t \in S$ and $|t| < \delta$, then $|f(t) - l| < \varepsilon$. Now, $t \in S$ if and only if $t = \frac{1}{n+1}$ for some $n \in \mathbb{N}$. If $N = \lceil \frac{1}{\delta} \rceil$, then we note that $n \geqslant N$ implies that $\frac{1}{n+1} < \delta$, and thus $|x_n - l| = \left|f\left(\frac{1}{n+1}\right) - l\right| < \varepsilon$. It is easily checked that this last statement is in fact equivalent to $\lim_0 f = l$.

Informally, convergence of f at 0 is understood as the convergence of $(x_n)_{n\in\mathbb{N}}$ at ∞. Thus the notion of convergence for sequences is nothing more than a very special case of our general study of limits of functions. All the theory developed there applies to sequences directly.

We restate the definition of convergence of sequence for reference.

Definition 4.1. A sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ converges to $l \in \mathbb{R}$ when

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \in \mathbb{N} \quad n \geqslant N \implies |x_n - l| < \varepsilon.$$

Our theory of limits for functions immediately applies to prove that limits are unique, if they exist, and satisfy various properties related to order and algebra. We will include proofs once more as this is an important topic.

Theorem 4.2. *If $(x_n)_{n\in\mathbb{N}}$ is a sequence converging to both l and l', then $l = l'$.*

Proof. Let $\varepsilon > 0$. Since $(x_n)_{n\in\mathbb{N}}$ converges to l, there exists $N_1 \in \mathbb{N}$ such that, for all $n \in \mathbb{N}$, if $n \geqslant N_1$, then $|x_n - l| < \frac{\varepsilon}{2}$. Since $(x_n)_{n\in\mathbb{N}}$ converges to l', there exists $N_2 \in \mathbb{N}$ such that, for all $n \in \mathbb{N}$, if $n \geqslant N_2$, then $|x_n - l'| < \frac{\varepsilon}{2}$. Thus, for $n = \max\{N_1, N_2\}$, we conclude:

$$|l - l'| \leqslant |l - x_n| + |x_n - l'| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

As $\varepsilon > 0$ is arbitrary, we conclude that $l = l'$. □

Notation 4.3. If $(x_n)_{n\in\mathbb{N}}$ converges to l, then we write $\lim_{n\to\infty} x_n = l$.

4.2. **Limit and Order.**

Theorem 4.4. *If a sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ converges, then it is bounded.*

Proof. Let $l = \lim_{n\to\infty} x_n$. There exists $N \in \mathbb{N}$ such that for all $n \geqslant N$, we have $|x_n - l| < 1$ so $|x_n| < |l| + 1$. Let $M = \max\{|l| + 1, |x_n| : n \in \mathbb{N}, n \leqslant N\}$. Then for all $n \in \mathbb{N}$ we have $|x_n| \leqslant M$. □

Theorem 4.5. *Let $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ be two convergent sequences in $\mathbb{R}$. If there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $x_n \leqslant y_n$, then $\lim_{n\to\infty} x_n \leqslant \lim_{n\to\infty} y_n$.*

Proof. Let $l_x = \lim_{n\to\infty} x_n$ and $l_y = \lim_{n\to\infty} y_n$. If $l_x > l_y$. There exists $N_x \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ with $n \geqslant N_x$, we have $l_x - \frac{\varepsilon}{2} < x_n < l_x + \frac{\varepsilon}{2}$.

There exists $N_y \in \mathbb{N}$ such that for all $n \geqslant N_y$ we have $l_y - \frac{\varepsilon}{2} < y_n < l_y + \frac{\varepsilon}{2}$.

For $n \geqslant \max\{N_x, N_y\}$, we have $l_x - \frac{\varepsilon}{2} \leqslant x_n \leqslant y_n \leqslant l_y + \frac{\varepsilon}{2}$. Therefore, $l_x \leqslant l_y + \varepsilon$ for all $\varepsilon > 0$. Hence $l_x \leqslant l_y$, as claimed. □

Theorem 4.6. *If $(x_n)_{n\in\mathbb{N}}$ converges to $l > 0$ then there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $x_n > \frac{l}{2}$.*

Proof. Let $\varepsilon = \frac{1}{2}l$. There exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $x_n > l - \varepsilon = \frac{l}{2}$. □

Theorem 4.7 (Squeeze Theorem)**.** *Let $(x_n)_{n\in\mathbb{N}}$, $(y_n)_{n\in\mathbb{N}}$ and $(z_n)_{n\in\mathbb{N}}$ be three sequences in $\mathbb{R}$ such that:*

1. *there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $x_n \leqslant y_n \leqslant z_n$,*
2. *the sequences $(x_n)_{n\in\mathbb{N}}$ and $(z_n)_{n\in\mathbb{N}}$ converge and $\lim_{n\to\infty} x_n = \lim_{n\to\infty} z_n$.*

Then $(y_n)_{n\in\mathbb{N}}$ converges and $\lim_{n\to\infty} y_n = \lim_{n\to\infty} x_n$.

Proof. Let $l = \lim_{n\to\infty} x_n = \lim_{n\to\infty} z_n$. Let $\varepsilon > 0$. There exists $N_x \in \mathbb{N}$ such that for all $n \geqslant N_x$ we have $|x_n - l| < \varepsilon$ so $l - \varepsilon < x_n$. There exists $N_z \in \mathbb{N}$ such that for all $n \geqslant N_z$ we have $|z_n - l| < \varepsilon$ so $z_n < l + \varepsilon$. Let $n \geqslant \max\{N, N_x, N_z\}$. Then:

$$l - \varepsilon < x_n \leqslant y_n \leqslant z_n < l + \varepsilon.$$

This concludes our proof. □

Corollary 4.8. *If $(x_n)_{n\in\mathbb{N}}$ converges to $l \in \mathbb{R}$ then $(|x_n|)_{n\in\mathbb{N}}$ converges to $|l|$.*

Proof. For all $n \in \mathbb{N}$ we have $0 \leqslant ||x_n| - |l|| \leqslant |x_n - l|$. Conclude by applying Theorem (4.7). □

4.2.1. *Limits and Algebra.*

Theorem 4.9. *If $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are two convergent sequences and $t \in \mathbb{R}$ then $(x_n + ty_n)_{n\in\mathbb{N}}$ converges to $\lim_{n\to\infty} x_n + t\lim_{n\to\infty} y_n$.*

Proof. Let $l_x = \lim_{n\to\infty} x_n$ and $l_y = \lim_{n\to\infty} y_n$. Let $\varepsilon > 0$. There exists $N_x \in \mathbb{N}$ such that for all $n \geqslant N_x$ we have $|x_n - l_x| < \frac{\varepsilon}{2}$. There exists $N_y \in \mathbb{N}$ such that for all $n \geqslant N_y$ we have $|y_n - l_y| \leqslant \frac{\varepsilon}{2(|t|+1)}$.

Let $n \geqslant \max\{N_x, N_y\}$. Then:

$$|(x_n + ty_n) - (l_x + tl_y)| \leqslant |x_n - l_x| + |t||y_n - l_y| < \varepsilon.$$

This concludes our proof. □

Theorem 4.10. *If $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are two convergent sequences then $(x_n y_n)_{n\in\mathbb{N}}$ converges to $\lim_{n\to\infty} x_n \cdot \lim_{n\to\infty} y_n$.*

Proof. Let $l_x = \lim_{n\to\infty} x_n$ and $l_y = \lim_{n\to\infty} y_n$. Since $(y_n)_{n\in\mathbb{N}}$ converges, it is bounded: there exists $M \in \mathbb{R}$ such that for all $n \in \mathbb{N}$ we have $|y_n| \leqslant M$. For all $n \in \mathbb{N}$ we have:

$$\begin{aligned} 0 \leqslant |x_n y_n - l_x l_y| &= |x_n y_n - l_x y_n + l_x y_n - l_x l_y| \\ &\leqslant |x_n - l_x||y_n| + |y_n - l_y||l_x| \\ &\leqslant |x_n - l_x|M + |y_n - l_y||l_x|. \end{aligned}$$

By assumption, $(|x_n - l_x|)_{n\in\mathbb{N}}$ and $(|y_n - l_y|)_{n\in\mathbb{N}}$ converge to 0. By Theorem (4.9), $(|x_n - l_x|M + |y_n - l_y||l_x|)_{n\in\mathbb{N}}$ converges to 0 as well. By Theorem (4.7), we conclude $(|x_n y_n - l_x l_y|)_{n\in\mathbb{N}}$ converges to 0. Hence $(x_n y_n)_{n\in\mathbb{N}}$ converges to $l_x l_y$. □

Theorem 4.11. *If $(x_n)_{n\in\mathbb{N}}$ is a sequence in $\mathbb{R}$ converging to some $l \in \mathbb{R} \setminus \{0\}$, then there exists $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}$ we have $x_n \neq 0$, and moreover $\left(\frac{1}{x_n}\right)_{n\geqslant N}$ converges to $\frac{1}{l}$.*

Proof. If $l < 0$, then replace $(x_n)_{n\in\mathbb{N}}$ by $(-x_n)_{n\in\mathbb{N}}$, so that we may assume $l > 0$. By Theorem (4.6), there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $x_n \geqslant \frac{l}{2} > 0$. Now let $\varepsilon > 0$. There exists $P \in \mathbb{N}$ such that for all $n \geqslant P$ we have $|x_n - l| < \frac{2\varepsilon}{l^2}$. Thus, for all $n \geqslant \max\{P, N\}$ we have:

$$\begin{aligned} \left|\frac{1}{x_n} - \frac{1}{l}\right| &= \frac{|l - x_n|}{l|x_n|} \\ &\leqslant \frac{|x_n - l|}{\frac{1}{2}l^2} \\ &< \varepsilon. \end{aligned}$$

This completes our proof. □

4.3. **Sequences and Topology.**

Theorem 4.12. *If $A \subseteq \mathbb{R}$, then*

$$\mathrm{cl}\,(A) = \left\{ l \in \mathbb{R} : \exists (x_n)_{n\in\mathbb{N}} \textit{ sequence in } A \quad l = \lim_{n\to\infty} x_n \right\}.$$

Proof. Write

$$S = \left\{ l \in \mathbb{R} : \exists (x_n)_{n\in\mathbb{N}} \text{ sequence in } A \quad l = \lim_{n\to\infty} x_n \right\}.$$

Let $l \in \mathrm{cl}\,(A)$. Let $n \in \mathbb{N}$. Since $\mathrm{dist}(l, A) = 0$, there exists $x_n \in A$ such that $|x_n - l| < \frac{1}{n+1}$. By Theorem (4.7), we have $\lim_{n\to\infty} x_n = l$. Hence $l \in S$, i.e. $\mathrm{cl}\,(A) \subseteq S$.

Let $l \in S$. By definition of S, there exists $(x_n)_{n\in\mathbb{N}}$ in A such that $l = \lim_{n\to\infty} x_n$. Let now $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}$, if $n \geqslant N$ then $|x_N - l| < \varepsilon$. Hence, $\mathrm{dist}(l, A) \leqslant |x_N - l| < \varepsilon$ (since $x_N \in A$). Thus $\mathrm{dist}(l, A) = 0$, i.e. $l \in \mathrm{cl}\,(A)$. Hence $S \subseteq \mathrm{cl}\,(A)$. □

Corollary 4.13. *A subset F in $\mathbb{R}$ is closed if, and only if, the limit of every convergent sequence of elements of F is also in F.*

Proof. Immediate. □

Theorem 4.14. *Let $D \subseteq \mathbb{R}$. Let $f : D \to \mathbb{R}$, $a \in \operatorname{cl}(D)$ and $l \in \mathbb{R}$. The function f converges to l at a along D if, and only if, for all sequences $(x_n)_{n\in\mathbb{N}}$ of elements of D converging to a, the sequence $(f(x_n))_{n\in\mathbb{N}}$ converges to l.*

Proof. Assume that $\lim_{\substack{x\to a\\ x\in D}} f(x) = l$. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in D converging to a. We could employ Theorem (3.33) directly to conclude that $\lim_{n\to\infty} f(x_n) = l$. We however provide an explicit proof here, for didactic purposes. Let $\varepsilon > 0$. Since $\lim_{\substack{x\to a\\ x\in D}} f(x) = l$, there exists $\delta > 0$ such that, for all $t \in D$, if $|t - a| < \delta$, then $|f(t) - l| < \varepsilon$. Since $\lim_{n\to\infty} x_n = a$, there exists $N \in \mathbb{N}$ such that, for all $n \geqslant N$, we have $|x_n - a| < \delta$. Therefore, for all $n \geqslant N$, we conclude $|f(x_n) - l| < \varepsilon$. Thus, $\lim_{n\to\infty} f(x_n) = l$.

Assume now that there exists $\varepsilon > 0$ such that, for all $\delta > 0$, there exists $t \in D$ such that $|t - a| < \delta$ and $|f(t) - l| \geqslant \varepsilon$. For each $n \in \mathbb{N}$, let $t_n \in D$ such that $|t_n - a| < \frac{1}{n+1}$ and $|f(t_n) - l| \geqslant \varepsilon$. By Theorem (4.7), we conclude that $\lim_{n\to\infty} t_n = a$. By construction, $(f(t_n))_{n\in\mathbb{N}}$ does not converge to l. Thus, our result follows by contraposition. □

4.4. **Monotone Sequences in $\mathbb{R}$.**

Definition 4.15. A sequence $(x_n)_{n\in\mathbb{N}}$ of real numbers is *increasing* when

$$\forall n \in \mathbb{N} \quad x_n \leqslant x_{n+1}.$$

A sequence $(x_n)_{n\in\mathbb{N}}$ of real numbers is *decreasing* when

$$\forall n \in \mathbb{N} \quad x_n \geqslant x_{n+1}.$$

A sequence $(x_n)_{n\in\mathbb{N}}$ is *monotone* when it is increasing or decreasing.

Remark 4.16. A sequence is strictly increasing, strictly decreasing and strictly monotone, when we replace the order by the strict order in Definition (4.15), though this will not be a very important concept for us.

The Dedekind completeness of the field $\mathbb{R}$ of real numbers imply the following very important theorem.

Theorem 4.17 (Monotone Convergence Theorem)**.** *If $(x_n)_{n\in\mathbb{N}}$ is a monotone sequence of real numbers, then $(x_n)_{n\in\mathbb{N}}$ converges if and only if $(x_n)_{n\in\mathbb{N}}$ is bounded.*

Proof. Assume first that $(x_n)_{n\in\mathbb{N}}$ is an increasing, bounded sequence. The set $\{x_n : n \in \mathbb{N}\}$ is not empty and bounded by assumption, thus $L = \sup\{x_n : n \in \mathbb{N}\}$ exists.

Let $\varepsilon > 0$. By characterization of suprema, there exists $N \in \mathbb{N}$ such that $L - \varepsilon < x_M \leqslant L$. Since $(x_n)_{n\in\mathbb{N}}$ is increasing, if $n \geqslant N$, we have $L - \varepsilon < x_N \leqslant x_n$. Moreover $x_n \leqslant L$. Thus for all $n \geqslant N$, we conclude $L - \varepsilon < x_n \leqslant L$, i.e. $|x_n - L| < \varepsilon$. So $(x_n)_{n\in\mathbb{N}}$ converges to L.

If $(x_n)_{n\in\mathbb{N}}$ is a decreasing, bounded sequence, then $(-x_n)_{n\in\mathbb{N}}$ is an increasing, bounded sequence, and thus it converges. Therefore, $(x_n)_{n\in\mathbb{N}}$ converges.

If $(x_n)_{n\in\mathbb{N}}$ is any convergent sequence — monotone or not — it is bounded. Thus our theorem is proven. □

We record an interesting application of the monotone convergence theorem, which may be seen either as a special case of compactness, or as a special case of completeness.

Theorem 4.18. *If $(I_n)_{n\in\mathbb{N}}$ is a sequence of nested closed intervals, i.e. $\forall n \in \mathbb{N} \quad I_{n+1} \subseteq I_n$, then $\bigcap_{n\in\mathbb{N}} I_n \neq \emptyset$.*

Proof. For each $n \in \mathbb{N}$, we write $I_n = [a_n, b_n]$. Since $I_{n+1} \subseteq I_n$, we conclude that $a_n \leqslant a_{n+1}$ and $b_{n+1} \leqslant b_n$. In particular, $(a_n)_{n\in\mathbb{N}}$ is an increasing, bounded above (by b_0) sequence, and thus it converges to some $x \in \mathbb{R}$. Of course, $x \geqslant a_n$ for all $n \in \mathbb{N}$. Similarly, the sequence $(b_n)_{n\in\mathbb{N}}$ is increasing and bounded, so it converges to some $y \in \mathbb{R}$; moreover $y \leqslant b_n$ for all $n \in \mathbb{N}$. Since $a_n \leqslant b_n$ for all $n \in \mathbb{N}$, we also have $x \leqslant y$. Thus, for all $n \in \mathbb{N}$, we have $a_n \leqslant x \leqslant y \leqslant b_n$ i.e. $[x, y] \subseteq \bigcap_{n\in\mathbb{N}}$. □

4.5. **Subsequences.**

Definition 4.19. A subsequence of a sequence $(x_n)_{n\in\mathbb{N}}$ in a set E is a sequence of the form $(x_{\phi(n)})_{n\in\mathbb{N}}$ for some strictly increasing function $\phi : \mathbb{N} \to \mathbb{N}$.

It is helpful to record the following result.

Lemma 4.20. *Let $\phi : \mathbb{N} \to \mathbb{N}$ be strictly increasing. Then for all $n \in \mathbb{N}$ we have $\phi(n) \geqslant n$.*

Proof. By definition, $\phi(0) \geqslant 0$. Assume that for some $n \in \mathbb{N}$ we have $\phi(n) \geqslant n$. Then by assumption, $\phi(n+1) > \phi(n) \geqslant n$ so $\phi(n+1) \geqslant n+1$. The lemma holds by the theorem of induction. □

Subsequences of a convergent sequence converge as well, and to the same limit. This tool may be used to show that a sequence such as $((-1)^n)_{n\in\mathbb{N}}$ does not have a limit.

The set of all limits of subsequences of a fixed sequence can actually be nicely described using the closure of tails, and is always a closed subset.

Theorem 4.21. *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence. The closed set:*

$$\bigcap_{n\in\mathbb{N}} \mathrm{cl}\left(\{x_k : k \geqslant n\}\right)$$

is the set of all limits of subsequences of $(x_n)_{n\in\mathbb{N}}$.

Proof. Let $l \in \bigcap_{n\in\mathbb{N}} \mathrm{cl}\left(\{x_k : k \geqslant n\}\right)$. Since $l \in \mathrm{cl}\left(\{x_k : k \geqslant 0\}\right)$, there exists $\varphi(0) \in \mathbb{N}$ such that $|x_{\varphi(0)} - l| \leqslant 1$. Assume now that we have constructed, for some $n \in \mathbb{N}$, natural numbers $\varphi(0) < \varphi(1) < \ldots < \varphi(n)$ such that $|x_{\varphi(k)} - l| < \frac{1}{k+1}$ for all $k \in \{0, \ldots, n\}$. Since $l \in \mathrm{cl}\left(\{x_k : k \neq \varphi(n)+1\}\right)$, there exists $\varphi(n+1) > \varphi(n)$ such that $|x_{\varphi(n+1)} - l| < \frac{1}{n+2}$. Thus by induction, we have constructed a subsequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ such that $|x_{\varphi(n)} - l| < \frac{1}{n+1}$ for all $n \in \mathbb{N}$. By the Squeeze Theorem, we conclude that $(x_{\varphi(n)})_{n\in\mathbb{N}}$ converges to l.

Let now $l \in \mathbb{R}$ be the limit of some subsequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$. Thus l is the limit of $(x_{\varphi(n)})_{n\geqslant N}$ for all $N \in \mathbb{N}$, and thus $l \in \mathrm{cl}\left(\{x_k : k \geqslant \varphi(n)\}\right)$ for all $n \in \mathbb{N}$; hence $l \in \bigcap_{n\in\mathbb{N}} \mathrm{cl}\left(\{x_k : k \geqslant \varphi(n)\}\right)$. Since $\varphi(n) \geqslant n$ for all $n \in \mathbb{N}$, we have $\{x_k : k \geqslant \varphi(n)\} \subseteq \{x_k : k \geqslant n\}$ and thus $\mathrm{cl}\left(\{x_k : k \geqslant \varphi(n)\}\right) \subseteq \mathrm{cl}\left(\{x_k : k \geqslant n\}\right)$. Therefore, $l \in \bigcap_{n\in\mathbb{N}} \mathrm{cl}\left(\{x_k : k \geqslant n\}\right)$, as desired. □

Theorem 4.22. *If a sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ converges to $l \in \mathbb{R}$, then all subsequences of $(x_n)_{n\in\mathbb{N}}$ converge to l as well.*

Proof. Assume that $(x_n)_{n\in\mathbb{N}}$ converges to l and let $\phi : \mathbb{N} \to \mathbb{N}$ be a strictly increasing function. Let $\varepsilon > 0$. By Definition (4.1), there exists $N \in \mathbb{N}$ such that for all $n \geqslant N$ we have $|x_n - l| < \varepsilon$. Therefore, if $n \geqslant N$ then $|x_{\phi(n)}, l| < \varepsilon$, since $\phi(n) \geqslant n \geqslant N$. Hence the subsequence $(x_{\phi(n)})_{n\in\mathbb{N}}$ converges to l as well. □

Theorem 4.23 (Monotone Subsequence Theorem)**.** *If $(x_n)_{n\in\mathbb{N}}$ is a sequence in $\mathbb{R}$, then there exists a monotone subsequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$.*

Proof. Let
$$\mathscr{P} = \{n \in \mathbb{N} : \forall k \in \mathbb{N} \quad k \geqslant n \implies x_n \geqslant x_k\}.$$

If $\mathscr{P}$ is infinite, then let $\varphi(0) = \min \mathscr{P}$. Assume that for some $n \in \mathbb{N}$ we have constructed $\varphi(0) < \ldots < \varphi(n) \in \mathscr{P}$. The set $\mathscr{P} \setminus \{j : j \leqslant \varphi(n)\}$ is not empty since $\mathscr{P}$ is infinite. Let $\varphi(n+1)$ be its smallest element. Now $\varphi : \mathbb{N} \to \mathbb{N}$ is strictly increasing, and by construction, if $n \leqslant m$ then $x_{\varphi(m)} \leqslant x_{\varphi(n)}$. Thus $(x_{\varphi(n)})_{n\in\mathbb{N}}$ is decreasing.

If $\mathscr{P}$ is finite, then $\mathscr{P}$ is bounded by M. Let $\varphi(0) = M + 1$. Now assume we have constructed $\varphi(0) < \ldots < \varphi(n)$ for some $n \in \mathbb{N}$ with $x_{\varphi(j)} \leqslant x_{\varphi(j+1)}$ for all $0 \leqslant j \leqslant n-1$. Since $\varphi(n) \notin \mathscr{P}$ there exists $\varphi(n+1) > \varphi(n)$ such that $x_{\varphi(n+1)} \geqslant x_{\varphi(n)}$. The sequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$ thus constructed by induction is increasing.

Our theorem is thus proven. □

4.6. **Sequentially Compact Subsets of $\mathbb{R}$.** Closed subsets of $\mathbb{R}$ share with finite subsets of $\mathbb{R}$ the property that a point almost in the set is in fact in the set. When working with sequences, in particular, finite sets have another desirable property: any sequence in a finite set must contain a constant, hence convergent, subsequence. Closed sets do not possess this property (consider $\mathbb{N}$), but this property is important enough to warrant the following definition.

Definition 4.24. A subset $K \subseteq \mathbb{R}$ of $\mathbb{R}$ is (sequentially) compact when every sequence in K has a convergent subsequence whose limit is in K.

Theorem 4.25. *A subset K of $\mathbb{R}$ is sequentially compact if, and only if, it is closed and bounded.*

Proof. Assume that K is closed and bounded. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in K. By Theorem (4.23), there exists a monotone subsequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$. Since K is closed, $(x_{\varphi(n)})_{n\in\mathbb{N}}$ is bounded as well. Thus by Theorem (4.17), the bounded monotone sequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$ converges to some l. Since K is closed, we conclude that $l \in K$. Therefore, K is compact.

Assume that K is not closed or not bounded. If K is not bounded, then for all $n \in \mathbb{N}$, there exists $x_n \in K$ such that $|x_n| \geqslant n$. Thus no subsequence of $(x_n)_{n\in\mathbb{N}}$ is bounded, and thus no subsequence of $(x_n)_{n\in\mathbb{N}}$ converges. So K is not sequentially compact.

If K is not closed, then there exists a convergent sequence $(x_n)_{n\in\mathbb{N}}$ of elements of K converging to $l \in \mathbb{R} \setminus K$. All the subsequences of $(x_n)_{n\in\mathbb{N}}$ converge to l. Thus K is not sequentially compact.

By contraposition, if K is sequentially compact, then K is closed and bounded. □

Corollary 4.26. *If $K \subseteq \mathbb{R}$ is sequentially compact and $F \subseteq \mathbb{R}$ is closed, then $K \cap F$ is sequentially compact.*

Proof. The set $K \cap F \subseteq K$ is included in K, so $K \cap F$ is bounded. Moreover, the intersection of two closed sets is closed. As a closed, bounded subset of $\mathbb{R}$, $K \cap F$ is sequentially compact in $\mathbb{R}$. $\square$

Theorem 4.27. *If $K \subseteq \mathbb{R}$ is sequentially compact, if $f : K \to \mathbb{R}$ is continuous, then $f(K)$ is sequentially compact.*

Proof. Let $(y_n)_{n\in\mathbb{N}}$ be a sequence in $f(K)$. By definition of $f(K)$, there exists $x_n \in K$ such that $f(x_n) = y_n$, for each $n \in \mathbb{N}$. Since K is sequentially compact, and $(x_n)_{n\in\mathbb{N}}$ is a sequence in the sequentially compact K, there exists a convergent subsequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$ with limit $l \in K$. Since f is continuous on K, we conclude $l = \lim_{n\to\infty} f(x_{\varphi(n)})$, i.e. $(y_{\varphi(n)})_{n\in\mathbb{N}} = (f(x_{\varphi(n)}))_{n\in\mathbb{N}}$ converges. Thus $f(K)$ is sequentially compact. $\square$

Corollary 4.28. *If $f : K \to \mathbb{R}$ is continuous over a sequentially compact subset $K \subseteq \mathbb{R}$ of $\mathbb{R}$, then there exists $m, M \in K$ such that*

$$\forall x \in K \quad f(m) \leqslant f(x) \leqslant f(M).$$

Proof. Since $f(K)$ is sequentially compact, it is closed and bounded. Hence $\sup K \in K$, so there exists $M \in K$ such that $f(M) = \sup K$, and $\inf K \in K$, so there exists $m \in K$ such that $f(m) = \inf K$. $\square$

We can now use sequential compactness to prove that $\mathbb{R}$ is uncountable.

Theorem 4.29. *If $P \subseteq \mathbb{R}$ is a nonempty perfect subset of $\mathbb{R}$, then P is uncountable. In particular, $\mathbb{R}$ is uncountable.*

Proof. Since P is perfect, P is not finite. Assume $P = \{x_n : n \in \mathbb{N}\}$.

Let $a_0, b_0 \in \mathbb{R}$ such that $a_0 < x_0 < b_0$ (for instance $a_0 = x_0 - 1$ and $b_0 = x_0 + 1$). Let $I_0 = [a_0, b_0]$. Set $\varphi(0) = 0$. Since P is perfect, the set $\{n \in \mathbb{N}, n > 0 : x_n \in P \cap (I_0 \setminus \{x_0\})\}$ is infinite. Let $\varphi(1)$ be the minimum of this set. Setting $\delta_1 = \min\{|a_0 - x_{\varphi(1)}|, |b_0 - x_{\varphi(1)}|, \frac{|x_{\varphi(1)} - x_0|}{2}\}$, we define $a_1 = x_{\varphi(1)} - \delta_1$ and $b_1 = x_{\varphi(1)} + \delta_1$. We thus observe: $x_0 \notin I_1$, $x_{\varphi(1)} \in I_1$ and $I_1 \subseteq I_0$.

We proceed by induction to construct a sequence $(I_n)_{n\in\mathbb{N}}$ of nested closed intervals and a strictly increasing function $\varphi : \mathbb{N} \to \mathbb{N}$ such that for all $n \in \mathbb{N}$, we have $x_n \notin I_{n+1}$, yet $x_{\varphi(n)} \in I_n \cap P$.

Since $I_0 \cap P$ is compact, the sequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$ has a convergent subsequence $(x_{\varphi\circ\psi(n)})_{n\in\mathbb{N}}$ with limit in P, hence with limit x_N for some $N \in \mathbb{N}$.

By construction, $(x_{\varphi\circ\psi(n)})_{n\geqslant N+1}$ lies in the closed set $P \cap I_{N+1}$, which does not contain x_N. This is a contradiction. Hence P is not countable.

We conclude noting that $\mathbb{R}$ is perfect. $\square$

Sequentially Compact sets can be infinite — for instance, $[0, 1]$ is sequentially compact and uncountably infinite since perfect — and have rather complex topology — consider the Cantor set — but they share two properties of finite sets: any point almost in a sequentially compact set is in fact in the sequentially compact set, and any sequence in a sequentially compact set has a convergent subsequence. Thus sequentially compactness is an analysis-motivated extension of the notion of a finite set. We will return to this important fact when trying to gain a deeper topological insight into sequentially compactness.

4.7. **Uniform Continuity.** Compactness is a helpful tool to obtain some uniformity results in analysis. A very important example of this phenomenon is the following strengthening of the notion of continuity.

Definition 4.30. A function $f : D \to \mathbb{R}$ is uniformly continuous over the subset $D \subseteq \mathbb{R}$ of $\mathbb{R}$ when

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x, y \in D \quad |x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

It is immediate that a uniformly continuous function is always continuous.

Example 4.31. If $f : D \to \mathbb{R}$ is k-Lipschitz, then f is uniformly continuous over D: for all $\varepsilon > 0$, if $x, y \in D$ with $|x - y| < \frac{\varepsilon}{k}$ then $|f(x) - f(y)| \leqslant k|x - y| < \varepsilon$.

We note that the composition of uniformly continuous functions are uniformly continuous.

Theorem 4.32. *If $f : D \to \mathbb{R}$ and $g : E \to \mathbb{R}$ are both uniformly continuous, and if $f(D) \subseteq E$, then $g \circ f : D \to \mathbb{R}$ is uniformly continuous.*

Proof. Let $\varepsilon > 0$. Since g is uniformly continuous, there exists $\delta_g > 0$ such that if $x, y \in E$ with $|x - y| < \delta_g$, then $|g(x) - g(y)| < \varepsilon$. Since f is uniformly continuous, there exists $\delta_f > 0$ such that, if $x, y \in D$ with $|x - y| < \delta_f$ then $|f(x) - f(y)| < \delta_g$. Thus if $x, y \in D$ with $|x - y| < \delta_f$ then $|g(f(x)) - g(f(y))| < \varepsilon$, as needed. □

A core observation is the following theorem.

Theorem 4.33. *If $f : K \to \mathbb{R}$ is a continuous function over a sequentially compact subset $K \subseteq \mathbb{R}$ of $\mathbb{R}$, then f is uniformly continuous over K.*

Proof. Assume that there exists $\varepsilon > 0$ such that, for all $\delta > 0$, there exists $x, y \in K$ such that $|x - y| < \delta$ and $|f(x) - f(y)| \geqslant \varepsilon$. For each $n \in \mathbb{N}$, there exists $x_n, y_n \in K$ such that $|x_n - y_n| < \frac{1}{n+1}$ and $|f(x_n) - f(y_n)| \geqslant \varepsilon$. Since K is sequentially compact, there exists a convergent subsequence $(x_{\varphi(n)})_{n \in \mathbb{N}}$ of $(x_n)_{n \in \mathbb{N}}$ converging to $l \in K$. Since $\lim_{n \to \infty} x_n - y_n = 0$, we conclude that $\lim_{n \to \infty} y_{\varphi(n)} = l$. Since $|f(x_{\varphi(n)}) - f(y_{\varphi(n)})| \geqslant \varepsilon$, at least one of the sequences $(f(x_{\varphi(n)}))_{n \in \mathbb{N}}$ or $(f(y_{\varphi(n)}))_{n \in \mathbb{N}}$ does not converge to $f(l)$. Thus f is not continuous over K.

Our theorem is proven by contraposition. □

4.8. **Continuous image of intervals.**

Theorem 4.34. *If $f : D \to \mathbb{R}$ is continuous over $D \subseteq \mathbb{R}$, and if $I \subseteq D$ is an interval, then $f(I)$ is an interval.*

Proof. Let $a, b \in I$. Let ξ between $f(a)$ and $f(b)$. Without loss of generality, assume $a \leqslant b$ and $f(a) \leqslant \xi < f(b)$ (if $f(b) \leqslant f(a)$ then work with $-f$).

Set $a_0 = a$ and $b_0 = b$. Assume that, for some $n \in \mathbb{N}$, we have constructed:

$$a_0 \leqslant a_1 \leqslant \ldots \leqslant a_n < b_n \leqslant b_{n-1} \leqslant \ldots \leqslant b_1 \leqslant b_0,$$

with $b_j - a_j = \frac{1}{2^j}(b - a)$ and $f(a_j) \leqslant \xi \leqslant f(b_j)$ for all $j \in \{0, \ldots, n\}$. Note that the assumption holds in particular for $n = 0$.

Let $m_n = \frac{a_n + b_n}{2}$. If $f(m_n) \leqslant \xi$ then set $a_{n+1} = m_n$ and $b_{n+1} = b_n$; otherwise set $b_{n+1} = m_n$ and $a_{n+1} = a_n$. We check that we then have proven our induction hypothesis for $n + 1$.

The sequence $(a_n)_{n\in\mathbb{N}}$ thus constructed is increasing and bounded above, so it converges by Theorem (4.17). Let $c = \lim_{n\to\infty} a_n$. We then note that $f(c) = \lim_{n\to\infty} f(a_n) \leqslant \xi$.

Since $\lim_{n\to\infty}(b_n - a_n) = 0$, we conclude that $\lim_{n\to\infty} b_n = c$, and thus $f(c) = \lim_{n\to\infty} f(b_n) \geqslant \xi$. Thus $f(c) = \xi$, and our theorem is proven. □

5. Cauchy Sequences

Sequences are a powerful tool of analysis, and their power get revealed when one realizes that one can use sequence to prove the existence of certain limits, without knowing the limit. This leads us to the concept of a Cauchy sequence, and the associated notions, such as functions preserving Cauchy sequences. The latter is best understood via a uniform version of continuity. We will connect these ideas with compactness as well, offering a metric view of sequential compactness.

5.1. Cauchy Property.

The Cauchy criterion for sequence allows the discussion of convergence without knowledge of the limit.

Definition 5.1. A sequence $(x_n)_{n\in\mathbb{N}}$ is a *Cauchy sequence* when

$$\forall\varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall p, q \in \mathbb{N} \quad (p \geqslant N \text{ and } q \geqslant N) \implies |x_p - x_q| < \varepsilon.$$

Theorem 5.2. *If $(x_n)_{n\in\mathbb{N}}$ is a convergent sequence, then $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence.*

Proof. Write $l = \lim_{n\to\infty} x_n$. Let $\varepsilon > 0$. Since $(x_n)_{n\in\mathbb{N}}$ converges, there exists $N \in \mathbb{N}$ such that, if $n \geqslant N$, then $|x_n - l| < \frac{\varepsilon}{2}$.

Thus, if $p, q \in \mathbb{N}$ and $p \geqslant N$, $q \geqslant N$, then

$$|x_p - x_q| \leqslant |x_p - l| + |l - x_q| \leqslant \frac{l}{2} + \frac{l}{2} = l.$$

Thus, $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence. □

Theorem 5.3. *If $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, then $(x_n)_{n\in\mathbb{N}}$ is bounded.*

Proof. Since $(x_n)_{n\in\mathbb{N}}$ is Cauchy, there exists $N \in \mathbb{N}$ such that, if $n \geqslant N$, then $|x_n - x_N| \leqslant 1$, i.e. $|x_n| \leqslant |x_N| + 1$. Thus, for all $n \in \mathbb{N}$, we have

$$|x_n| \leqslant \max\{1 + |x_N|, |x_j| : j \in \{1, \ldots, N-1\}\}.$$

This completes our proof. □

Of central importance is the observation that a Cauchy sequence wants to converge, and the only reason it might not is that its limit may not exist in the space where the sequence is defined.

Theorem 5.4. *If $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence with a convergent subsequence, then $(x_n)_{n\in\mathbb{N}}$ converges.*

Proof. We thus assume $(x_{\varphi(n)})_{n\in\mathbb{N}}$ converges to some $l \in \mathbb{R}$. Let $\varepsilon > 0$. Since $(x_n)_{n\in\mathbb{N}}$ is Cauchy, there exists $N \in \mathbb{N}$ such that, for all $p, q \geqslant N$, we have $|x_p - x_q| < \frac{\varepsilon}{2}$.

Since $(x_{\varphi(n)})_{n\in\mathbb{N}}$ converges to l, there exists $M \in \mathbb{N}$ such that, for all $n \geqslant M$, we have $|x_{\varphi(n)} - l| < \frac{\varepsilon}{2}$.

Let now $n \geqslant \max\{N, M\}$. Note that $\varphi(n) \geqslant n \geqslant \max\{M, N\}$. thus

$$|x_n - l| \leqslant |x_n - x_{\varphi(n)}| + |x_{\varphi(n)} - l|$$

$$\leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This concludes our proof. □

5.2. **Completeness.** All Cauchy sequences in $\mathbb{R}$ converge.

Theorem 5.5. *If $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$, then it converges.*

Proof. Since $(x_n)_{n\in\mathbb{N}}$ is Cauchy, it is bounded. By Theorem (4.23), $(x_n)_{n\in\mathbb{N}}$ has a monotone subsequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$. By Theorem (4.17), $(x_{\varphi(n)})_{n\in\mathbb{N}}$ converges. By Theorem (5.4), the sequence $(x_n)_{n\in\mathbb{N}}$ converges. □

We can now ask a more subtle question: if $A \subseteq \mathbb{R}$, under what condition do all Cauchy sequences of A converge in A?

Definition 5.6. A subset A of $\mathbb{R}$ is complete when all Cauchy sequences in A converge in A.

Thus, in particular, $\mathbb{R}$ is complete. More generally,

Theorem 5.7. *A subset A of $\mathbb{R}$ is complete if, and only if A is closed.*

Proof. If $A \subseteq \mathbb{R}$ is complete, then given any sequence $(x_n)_{n\in\mathbb{N}}$ in A converging to some $x \in \mathbb{R}$, we have $x \in A$, otherwise, $(x_n)_{n\in\mathbb{N}}$ (as a convergent sequence in $\mathbb{R}$ which takes values in A) is a Cauchy sequence in A which can not converge, and thus A is not complete. So A contains the limits of all its convergent sequences, and thus A is closed.

If $A \subseteq \mathbb{R}$ is closed, and if $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence of element in A, then $(x_n)_{n\in]N}$ converges as $\mathbb{R}$ is complete, and its limit lies in A as A is closed. □

It is not true in general that completeness and closed-ness are synonymous. Indeed, the notion of completeness makes sense in $\mathbb{Q}$ endowed with its usual metric. We can define a topology on $\mathbb{Q}$ using this metric, but suffices to say that $\mathbb{Q}$ itself would be closed for this topology and yet, it is not complete (there are sequences in $\mathbb{Q}$ approaching $\sqrt{2}$ in $\mathbb{R}$). It is however true that completeness always implies closeness.

5.3. **Total boundedness.** We noted the importance of the notion of sequential compactness, where any sequence has a convergent subsequence. As we introduced the notion of a Cauchy sequence, it is reasonable to wonder what property of a set guarantees the existence of a Cauchy subsequence to any sequence in the set. We are led to the following.

Definition 5.8. A subset F of a subset A of $\mathbb{R}$ is δ-dense when

$$\forall x \in A \quad \exists y \in F \quad |x - y| < \delta.$$

Definition 5.9. A subset $A \subseteq \mathbb{R}$ is *totally bounded* when, for all $\delta > 0$, there exists a finite δ-dense subset $F \subseteq A$ of A.

Theorem 5.10. *If a subset $A \subseteq \mathbb{R}$ is totally bounded, then it is bounded.*

Proof. Since A is totally bounded, it contains a 1-dense finite subset F. Let $M = \max\{|x - y| : x, y \in F\}$. If $a, b \in A$ then there exists $x, y \in F$ such that $|a - x| < 1$ and $|b - y| < 1$. Thus

$$|a - b| \leqslant |a - x| + |x - y| + |y - b| < 2 + M.$$

Thus $\operatorname{diam}(A) \leqslant 2 + M$ and our proof is complete. □

A simple feature of total boundedness is that it is automatically inherited by subsets.

Theorem 5.11. *A subset $A \subseteq \mathbb{R}$ is totally bounded if, and only if, for all $\delta > 0$, there exists a finite subset F of $\mathbb{R}$ such that $A \subseteq \bigcup_{x \in F}(x-\delta, x+\delta)$.*

Proof. The condition is necessary by definition. Let us now assume that, for all $\delta > 0$ there exists a finite subset $F \subseteq \mathbb{R}$ of $\mathbb{R}$ such that $A \subseteq \bigcup_{x\in F}(x - delta, x+\delta)$. Let $\delta > 0$ be chosen, and $F \subseteq \mathbb{R}$ be given so that $A \subseteq \bigcup_{x \in F}\left(x - \frac{\delta}{2}, x + \frac{\delta}{2}\right)$.

As our result is obvious for $A = \emptyset$, we henceforth assume that we fixed $z \in A$. For each $x \in F$, let $y(x)$ be either any element of $A \cap \left(x - \frac{\delta}{2}, x + \frac{\delta}{2}\right)$ — if the latter intersection is not empty — or $y(x) = z$ otherwise.

By construction, y is a surjection from the finite set F onto $G = \{y(x) : x \in F\}$, so the latter set is finite. Moreover, G is a subset of A by construction. Last, if $a \in A$, then there exists $x \in F$ such that $|a - x| < \frac{\delta}{2}$. In particular, $y(x) \in A \cap \left(x - \frac{\delta}{2}, x + \frac{\delta}{2}\right)$ (since $a \in A \cap \left(x - \frac{\delta}{2}, x + \frac{\delta}{2}\right)$). So $|x - y(x)| < \frac{\delta}{2}$ and thus

$$|a - y(x)| \leqslant |a - x| + |x - y(x)| < \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

The set $\{y(x) : x \in F\}$ is thus a finite δ-dense subset of A, and our proof is complete. □

Corollary 5.12. *If $A \subseteq \mathbb{R}$ is totally bounded and if $B \subseteq A$, then B is totally bounded.*

Proof. This is immediate from Theorem (5.11). □

The crucial observation about total boundedness is its relation with Cauchy sequences.

Theorem 5.13. *A subset A of $\mathbb{R}$ is totally bounded if, and only if, every sequence in A has a Cauchy subsequence.*

Proof. Assume that $A \subseteq \mathbb{R}$ is totally bounded.

Let $(x_n)_{n\in\mathbb{N}}$ be a sequence of elements in A. We define the sets:

$$X_{n,p} = \left\{ k > n : |x_k - x_n| < \frac{\operatorname{diam}(A)}{p+1} \right\}.$$

We define $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ by induction. Set $\varphi(0) = 0$ and note that $X_{0,0} = \mathbb{N}$. Assume we have built $\varphi(0) < \ldots < \varphi(n)$ for some $n \in \mathbb{N}$ in such a way that $X_{\varphi(n),n}$ is infinite, and $X_{\varphi(n),n} \subseteq X_{\varphi(n-1),n-1} \subseteq \cdots \subseteq X_{0,0}$.

Since A is totally bounded, there exists a finite $\frac{\operatorname{diam}(A)}{2(n+2)}$-dense subset F_n of A. For each $x \in F_n$, let $A_{n,x} = \{k \in X_{\varphi(n),n} : |x_k - x| < \frac{1}{2(n+2)}|\}$. Since $X_{\varphi(n),n} = \bigcup_{x\in F_n} A_{n,x}$ is infinite by the induction hypothesis, while F_n is finite, there exists $x \in F_n$ such that $A_{n,x}$ is infinite. Let $\varphi(n+1) = \min A_{n,x}$. By construction, if $k \in A_{n,x}$, then $k > \varphi(n)$, and

$$\begin{aligned}|x_k - x_{\varphi(n+1)}| &\leqslant |x_k - x| + |x - x_{\varphi(n+1)}| \\ &\leqslant \frac{1}{2(n+2)} + \frac{1}{2(n+2)} = \frac{1}{n+2}.\end{aligned}$$

Since $A_{n,x}$ is infinite and $A_{n,x} \subseteq X_{n+1} = \{k > \varphi(n+1) : |x_k - x_{\varphi(n)}| < \frac{1}{n+2}\}$, we conclude that we have met out induction hypothesis at $n+1$. The subsequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$ thus constructed is indeed Cauchy. Let $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that $\frac{1}{N+1} < \frac{\varepsilon}{2}$ (since $\mathbb{R}$ is Archimedean). By construction, for all $p, q \geqslant N$, we have

$$\begin{aligned}|x_{\varphi(p)} - x_{\varphi(q)}| &< |x_{\varphi(p)} - x_{\varphi(N)}| + |x_{\varphi(N)} - x_{\varphi(q)}| \\ &< \frac{1}{N} + \frac{1}{N} < \varepsilon.\end{aligned}$$

We thus have shown that every sequence in a totally bounded set has a Cauchy subsequence.

Let us now assume that A is not totally bounded. Let $x_0 \in A$. Since A is not totally bounded, there exists $x_1 \in A \setminus (x_0 - 1, x_0 + 1)$. Suppose now that we have constructed $x_0, x_1, \ldots, x_n \in A$ such that $|x_j - x_k| \geqslant 1$ for all $0 \leqslant j \neq k \leqslant n$. Since A is not totally bounded, there exists $x_{n+1} \in A \setminus \bigcup_{j=0}^{n}(x_j - 1, x_j + 1)$. Thus, by induction, there exists a sequence $(x_n)_{n\in\mathbb{N}}$ such that, for all $j, k \in \mathbb{N}$, if $j \neq k$ then $|x_j - x_k| \geqslant 1$. Therefore, $(x_n)_{n\in\mathbb{N}}$ has no Cauchy subsequence. $\square$

We are thus led to an important result. The correct way to understand sequential compactness (in the general metric space situation) is given in the following corollary.

Corollary 5.14. *A subset $D \subseteq \mathbb{R}$ is sequentially compact if, and only if, D is complete and totally bounded.*

Proof. If D is sequentially compact, then every sequence of elements of D has a convergence, hence Cauchy, subsequence. So D is totally bounded. As D is sequentially compact, it is closed, and hence complete, since $\mathbb{R}$ is complete.

If D is complete and totally bounded, then every sequence in D has a Cauchy subsequence, which must converge in D since D is complete. So D is sequentially compact. $\square$

We invite the reader to note that $[0,1] \cap \mathbb{Q}$ is a totally bounded subset of $\mathbb{Q}$, yet it contains sequences with no convergent subsequences. The characterization of sequentially compactness in metric space as complete and totally bounded subsets is very important; it is sometimes obscured by the easier observation that sequentially compact subsets of $\mathbb{R}$ are simply closed (which is equivalent to complete since $\mathbb{R}$ is complete) and bounded (which, in the very special case of finite dimensional vector spaces, is equivalent to totally bounded). We can see from the theory so far that the following result is now easy, though it is instructive to give it a direct proof as well.

Theorem 5.15. *A subset of $\mathbb{R}$ is totally bounded if, and only if, it is bounded.*

First method: using the monotone subsequence theorem. A totally bounded subset of $\mathbb{R}$ is always bounded. If $A \subseteq \mathbb{R}$ is bounded, then any sequence in A contains a bounded monotone subsequence by Theorem (4.23), which is Cauchy since bounded by Theorem (4.17). (Note: we did not assume that A is closed, so a bounded monotone sequence in A may have a limit outside of A). $\square$

Second method: using sequentially compactness. A totally bounded subset of $\mathbb{R}$ is always bounded. If A is bounded, then so is its closure by Theorem (2.14), and thus $\mathrm{cl}(A)$ is closed and bounded, hence sequentially compact in $\mathbb{R}$; thus every sequence in A has a Cauchy subsequence (which converges in $\mathrm{cl}(A)$). □

Third method: direct approach. A totally bounded subset of $\mathbb{R}$ is always bounded. Let now A be a bounded subset of $\mathbb{R}$. Thus there exists $a < b$ such that $A \subseteq [a, b]$. Since subsets of totally bounded sets are totally bounded by Theorem (5.11), it is sufficient to show that $[a, b]$ is totally bounded.

Let $\delta > 0$. Then

$$[a,b] \subseteq \bigcup_{n=0}^{N} ((a-1)+n\delta, a+(n+1)\delta) \text{ where } N = \min\{k \in \mathbb{N} : a + k\delta > b\}.$$

□

Remark 5.16. The first two proofs are in fact exactly the same argument, since we used Theorem (4.23) and Theorem (4.17) in our initial treatment of sequentially compactness in $\mathbb{R}$.

5.4. **Cauchy Continuity.** Continuous functions are characterized as mapping convergent sequences to convergent sequences, and sequentially compact sets to sequentially compact sets. We now investigate functions with similar roles regarding Cauchy sequences and totally bounded sets.

Definition 5.17. A function $f : D \to \mathbb{R}$ is *Cauchy continuous* over D when, for all Cauchy sequences $(x_n)_{n\in\mathbb{N}}$ in D, the sequence $(f(x_n))_{n\in\mathbb{N}}$ is Cauchy.

It is easy to check that the composition of two Cauchy continuous functions is again Cauchy continuous. It is not true that a continuous function must be Cauchy continuous (for instance, consider $f : x \in (0, 1] \mapsto \frac{1}{x}$). However, every Cauchy continuous function is indeed continuous.

Theorem 5.18. *If $f : D \to \mathbb{R}$ is Cauchy continuous over D, then f is continuous over D.*

Proof. Let $x \in D$. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in D converging to x. For each $n \in \mathbb{N}$, we set $z_{2n} = x_n$ and $z_{2n+1} = x$. The sequence $(z_n)_{n\in\mathbb{N}}$ converges to x and thus, it is Cauchy. Since f is Cauchy continuous, $(f(z_n))_{n\in\mathbb{N}}$ is a Cauchy sequence. It contains the subsequence $(f(z_{2n+1}))_{n\in\mathbb{N}} = (f(x))_{n\in\mathbb{N}}$ which converges to $f(x)$. By Theorem (5.4), we conclude that $(f(z_n))_{n\in\mathbb{N}}$ converges to $f(x)$ as well. As a subsequence of a convergent sequence with limit $f(x)$, the sequence $(f(x_n))_{n\in\mathbb{N}}$ converges to $f(x)$ as well.

Thus f is continuous at all $x \in D$. □

The fundamental example of Cauchy continuous functions is given by the uniformly continuous functions — thus include, in particular, continuous functions over compact.

Theorem 5.19. *If $f : D \to \mathbb{R}$ is uniformly continuous over D, then f is Cauchy continuous over D.*

Proof. Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in D. Let $\varepsilon > 0$. Since f is uniformly continuous over D, there exists $\delta > 0$ such that, if $x, y \in D$ with $|x - y| < \delta$, then $|f(x) - f(y)| < \varepsilon$. Since $(x_n)_{n\in\mathbb{N}}$ is Cauchy, there exists $N \in \mathbb{N}$ such that, for all $p, q \in \mathbb{N}$, we have $|x_p - x_q| < \delta$. Thus $|f(x_p) - f(x_q)| < \varepsilon$, i.e. $(f(x_n))_{n\in\mathbb{N}}$ is Cauchy, as claimed. □

Uniform continuity is in fact very closely related to Cauchy continuity, as seen in the following theorem.

Theorem 5.20. *A function $f : D \to \mathbb{R}$ is Cauchy continuous, if and only if the restriction of f to any totally bounded subset of D is uniformly continuous.*

Proof. Assume that the restriction of f to any totally bounded subset of D is uniformly continuous. Let $(x_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in D. Since $\{x_n : n \in \mathbb{N}\}$ is totally bounded, the restriction of f to the range of $(x_n)_{n\in\mathbb{N}}$ is uniformly continuous, and thus $(f(x_n))_{n\in\mathbb{N}}$ is a Cauchy sequence.

Assume now that there exists a totally bounded subset A of D to which f does not restrict to a uniformly continuous function. Assume that the restriction of f to A is not uniformly continuous over A. Thus, there exists $\varepsilon > 0$ such that for all $n \in \mathbb{N}$, there exists $x_n, y_n \in A$ such that $|x_n - y_n| < \frac{1}{2^n}$ yet $|f(x_n) - f(y_n)| \geqslant \varepsilon$. Since A is totally bounded, there exists Cauchy subsequences $(x_{\varphi(n)})_{n\in\mathbb{N}}$ of $(x_n)_{n\in\mathbb{N}}$, and then a Cauchy subsequence $(y_{\varphi\circ\psi(n)})_{n\in\mathbb{N}}$. Setting $\theta = \varphi \circ \psi$, we thus have two Cauchy sequences $(x_{\theta(n)})_{n\in\mathbb{N}}$ and $(y_{\theta(n)})_{n\in\mathbb{N}}$, and up to extracting further subsequences, we assume that $x_{\theta(n+1)} - x_{\theta(n)}| < \frac{1}{2^n}$ and $|y_{\theta(n+1)} - y_{\theta(n)}| < \frac{1}{2^n}$. Therefore (noting $\theta(n) \geqslant n$ for all $n \in \mathbb{N}$), we conclude that the sequence $(z_n)_{n\in\mathbb{N}}$ defined by

$$\forall n \in \mathbb{N} \quad z_{2n} = x_n \text{ and } z_{2n+1} = y_n$$

is a Cauchy sequence, since $|z_{n+1} - z_n| < \frac{1}{2^n}$ by construction. Yet $(f(z_n))_{n\in\mathbb{N}}$ is not a Cauchy sequence since $|f(z_{2n+1}) - f(z_{2n})| = |f(x_{\theta(n)}) - f(y_{\theta(n)})| \geqslant \varepsilon$ for all $n \in \mathbb{N}$. So f is not Cauchy continuous. Our proof is complete by contraposition. □

We now turn to the following result about preservation of total boundedness.

Theorem 5.21. *If $f : A \to \mathbb{R}$ is Cauchy continuous, and if A is totally bounded, then $f(A)$ is totally bounded.*

Proof using Cauchy sequences. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in $f(A)$. For each $n \in \mathbb{N}$, let $y_n \in A$ such that $f(y_n) = x_n$. Since A is totally bounded, there exists a Cauchy subsequence $(y_{\varphi(n)})_{n\in\mathbb{N}}$ of $(y_n)_{n\in\mathbb{N}}$. Since f is Cauchy continuous, the sequence $(x_{\varphi(n)})_{n\in\mathbb{N}} = (f(y_{\varphi(n)}))_{n\in\mathbb{N}}$ is Cauchy. So every sequence of $f(A)$ has a Cauchy subsequence. By Theorem (5.13), the set $f(A)$ is totally bounded. □

Proof using uniform continuity. The restriction of f to A is uniformly continuous, since f is Cauchy continuous and A is totally bounded. Let $\varepsilon > 0$. Since f is uniformly continuous, there exists $\delta > 0$ such that, for all $x, y \in A$, if $|x - y| < \delta$, then $|f(x) - f(y)| < \varepsilon$. Since A is totally bounded, there exists a finite δ-dense finite subset F of A. If $x \in f(A)$, then $x = f(y)$ for some $y \in A$. There exists $z \in F$ such that $|y - z| < \delta$, and thus $|x - f(z)| = |f(y) - f(z)| < \varepsilon$. So $\{f(z) : z \in F\}$ is a finite ε-dense subset of $f(A)$. Therefore, $f(A)$ is indeed totally bounded. □

6. Topology of $\mathbb{R}$: Compactness

Sequentially compact subsets of $\mathbb{R}$ are the complete, totally bounded subsets of $\mathbb{R}$ — equivalently, in this case, the closed bounded subsets of $\mathbb{R}$, and they exhibit many desirable properties: the continuous images of compact are compact, points almost in a compact set are in fact in it, any sequence in a compact has a convergent subsequence, any Cauchy sequence in a compact set must converge in that set, continuity over a compact set is equivalent to Cauchy continuity which is also equivalent in this case to uniform continuity. All these properties are far reaching generalizations of the properties of finite sets to analysis. There are two informal reasons why this is important. Compactness is a tool to obtain existence results from mere topological considerations (convergent sequences and their limit, maxima and minima of functions, and so on), and it provides means to strengthen properties toward uniformity (as in uniform continuity from continuity).

This notion is very important and deserves deeper exploration. In particular, the proper definition of compactness, when leaving the realm of metric spaces, is challenging, but it provides instructive properties of compact sets yet to be explored in these notes. This is the matter of this section.

6.1. **Compact Sets.**

We propose a definition of compactness inspired by our work on subsequences. We say in Theorem (4.21) that the set of all limits of convergent subsequences of a sequence $(x_n)_{n\in\mathbb{N}}$ is given as $\bigcap_{n\in\mathbb{N}} \mathrm{cl}\left(\{x_k : k \geqslant n\}\right)$, i.e. an infinite intersection of closed subsets of $\mathbb{R}$. Thus, a set K if sequentially compact if, and only if, for any sequence $(x_n)_{n\in\mathbb{N}}$ in K, the intersection $\bigcap_{n\in\mathbb{N}} \mathrm{cl}\left(\{x_k : k \geqslant n\}\right)$ is never empty. Now, we note that for all $N \in \mathbb{N}$, we have $\bigcap_{n=0}^{N} \mathrm{cl}\left(\{x_k : k \geqslant n\}\right) \neq \emptyset$. So, we start with the idea that compactness can be understood via the *finite intersection property*:

Definition 6.1. A subset $A \subseteq \mathbb{R}$ is *compact* if it has the finite intersection property: for any collection $\mathscr{F}$ of closed subsets of $\mathbb{R}$ such that $A \cap \bigcap \mathscr{F} = \emptyset$, there exists a *finite* subset $\mathscr{G} \subseteq \mathscr{F}$ of $\mathscr{F}$ such that $A \cap \bigcap \mathscr{G} = \emptyset$.

The finite intersection property is easily checked to be equivalent to the following:

Theorem 6.2. *A subset $K \subseteq \mathbb{R}$ of $\mathbb{R}$ is compact if, and only if, for any collection $\mathscr{F}$ of closed subsets of $\mathbb{R}$, if $K \cap \bigcap \mathscr{G} \neq \emptyset$ whenever $\mathscr{G} \subseteq \mathscr{F}$ is a* finite *subset of $\mathscr{F}$, then $K \cap \bigcap \mathscr{F} \neq \emptyset$.*

Proof. This follows immediately by contraposition. □

We thus see that the finite intersection property is an existence property. We immediately record:

Theorem 6.3. *If $K \subseteq \mathbb{R}$ is compact, then K is sequentially compact.*

Proof. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a compact K. Let $X_n = \mathrm{cl}\left(\{x_k : k \geqslant n\}\right)$. Note that $X_p \subseteq X_n$ for all $p, n \in \mathbb{N}$ with $p \geqslant n$. Thus $\bigcap_{j\in F} X_j = X_{\max F} \neq \emptyset$ for all finite subset F of $\mathbb{N}$. Since K is compact, $\bigcap_{n\in\mathbb{N}} X_n \neq \emptyset$. By Theorem (4.21), the sequence $(x_n)_{n\in\mathbb{N}}$ has at least one convergent subsequence. Thus K is sequentially compact. □

Corollary 6.4. *A compact subset of $\mathbb{R}$ is closed and bounded.*

Proof. Since compact sets are sequentially compact, this follows from Theorem (4.25). □

Of great interest to us is the converse of this result. To obtain it, we first provide some useful characterizations of compactness.

Theorem 6.5. *A subset $K \subseteq \mathbb{R}$ of $\mathbb{R}$ is compact if, and only if, for any collection $\mathscr{U}$ of open subsets of $\mathbb{R}$ such that $K \subseteq \bigcup \mathscr{U}$, there exists a* finite *subset $\mathscr{V} \subseteq \mathscr{U}$ of $\mathscr{U}$ such that $K \subseteq \bigcup \mathscr{V}$.*

Proof. This equivalence follows by taking the complements in the definition of compactness. □

We thus see explicitly how compactness generalizes finiteness in our topological setting.

It is interesting to prove that compact sets are closed and bounded directly from their definition. If K is not bounded, then there can be no finite open subcover from $K \subseteq (n, n+1)$, so K cannot be compact. A very general, albeit technical, proof compactness implies closedness is as follows. For each $x \notin K$ and $y \in K$, let $U_{x,y}$ and $V_{x,y}$ be two *disjoint* open subsets of $\mathbb{R}$ such that $y \in U_{x,y}$ and $x \in V_{x,y}$. Then $K \subseteq \bigcup_{y \in K} U_{x,y}$ by construction; as K is compact, there exists a finite subset $F_x \subseteq K$ such that $K \subseteq \bigcup_{y \in F_x} U_{x,y}$. It is easy to check that $\bigcap_{y \in F_x} V_{x,y} \cap K \subseteq \bigcap_{y \in F_x} V_{x,y} \cap \bigcup_{y \in F_x} U_{x,y} = \emptyset$, so $\bigcap_{y \in F_x} V_{x,y} \subseteq \mathbb{R} \setminus K$. Now, as a *finite* intersection of open subsets of $\mathbb{R}$, the set $\bigcap_{y \in F_x} V_{x,y}$ is open and contain x. We conclude from this that $\mathbb{R} \setminus K$ is open, and thus K is closed.

We also have already enough knowledge of compact sets and continuous functions to (re)prove the following result.

Theorem 6.6. *If $f : K \to \mathbb{R}$ is a continuous function, and if $K \subseteq \mathbb{R}$ is a compact subset of $\mathbb{R}$, then $f(K)$ is compact.*

Proof. Let $\mathscr{U}$ be a collection of open subsets of $\mathbb{R}$ such that $f(K) \subseteq \bigcap \mathscr{U}$. For each $U \in \mathscr{U}$, let G_U be an open subset of $\mathbb{R}$ such that $f^{-1}(U) = G_U \cap K$. Thus

$$K \subseteq f^{-1}(f(K)) \subseteq \bigcup \{f^{-1}(U) : U \in \mathscr{U}\} \subseteq \bigcup \{G_U : U \in \mathscr{U}\}.$$

Since K is compact, there exists a finite subset $\mathscr{V} \subseteq \mathscr{U}$ such that $K \subseteq \bigcup \{G_U : U \in \mathcal{V}\}$. Therefore, $f(K) \subseteq \bigcup \mathcal{V}$ (note: $f(G_U \cap K) \subseteq U$). Therefore, $f(K)$ is compact. □

6.2. **Equivalence of compactness and sequential compactness in $\mathbb{R}$.** We now proceed to show that sequential compactness implies compactness in our setting.

Theorem 6.7 (Lebesgue's covering number)**.** *Let $K \subseteq \mathbb{R}$ be a sequentially compact subset of $\mathbb{R}$. If $\mathscr{U}$ is an open cover of K, then there exists $\delta > 0$ such that, for all $x \in K$, there exists $U \in \mathscr{U}$ such that $(x - \delta, x + \delta) \subseteq U$.*

Proof. Let $K \subseteq \mathscr{U}$ be an open cover of a sequentially compact set K. Assume that for all $\delta > 0$, there exists $x \in K$ such that $(x - \delta, x + \delta)$ is not a subset of any $U \in \mathscr{U}$.

For each $n \in \mathbb{N}$, there exists $x_n \in K$ such that $\left(x_n - \frac{1}{n+1}, x_n + \frac{1}{n+1}\right)$ is not contained in any $U \in \mathscr{U}$.

Since K is sequentially compact, the sequence $(x_n)_{n\in\mathbb{N}}$ has a convergent subsequence $(x_{\varphi(n)})_{n\in\mathbb{N}}$ with limit in K. Let $l \in K$ be its limit. Since $K \subseteq \bigcup \mathscr{U}$, there exists $U \in \mathscr{U}$ such that $l \in U$. Since U is open, there exists $\delta > 0$ such that $(l - \delta, l + \delta) \subseteq U$. Since $(x_{\varphi(n)})_{n\in\mathbb{N}}$ converges to l, there exists $N \in \mathbb{N}$ such that, for all $n \geqslant N$, we have $|x_{\varphi(n)} - l| < \frac{\delta}{2}$. Since $\lim_{n\to\infty} \frac{1}{\varphi(n)} = 0$, there exists $N' \in \mathbb{N}$ such that, if $n \geqslant N'$, then $\frac{1}{\varphi(n)} < \frac{\delta}{2}$. Let $n = \max\{N, N'\}$. If $t \in (x_{\varphi(n)} - \frac{1}{\varphi(n)+1}, x_{\varphi(n)} + \frac{1}{\varphi(n)+1}) \subseteq (x - \frac{\delta}{2}, x + \frac{\delta}{2})$, then:

$$|t - l| \leqslant |t - x_{\varphi(n)}| + |x_{\varphi(n)} - l| < \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

Thus we have reached a contradiction. Our theorem is proven. ☐

Theorem 6.8. *A subset K of $\mathbb{R}$ is compact if, and only if, it is sequentially compact.*

Proof. Theorem (4.21) proves that if K is compact, then it is sequentially compact. Let us now assume that K is sequentially compact. Let $\mathscr{U}$ be an open cover of K. By Theorem (6.7), there exists $\delta > 0$ such that for all $x \in K$, there exists $U_x \in \mathscr{U}$ such that $(x - \delta, x + \delta) \subseteq U$. By Theorem (5.14), since K is sequentially compact, it is totally bounded. So, there exists a finite δ-dense subset F of K. It then follows that $K \subseteq \bigcup_{x\in F}(x - \delta, x + \delta) \subseteq \bigcup_{x\in F} U_x$. Thus $\{U_x : x \in F\}$ is our open subcover of K from $\mathscr{U}$. So K is compact. ☐

7. The Spaces of Bounded Functions

7.1. Norms.

Definition 7.1. A *norm* $\|\cdot\|_E$ on a vector space E (over $\mathbb{R}$ or $\mathbb{C}$) is a function from E to $[0, \infty)$ such that

- **triangle inequality:** $\forall x, y \in E \quad \|x + y\|_E \leqslant \|x\|_E + \|y\|_E$,
- **homogeneity:** $\forall x \in E \quad \forall t \in \mathbb{R} \quad \|tx\|_E = |t|\,\|x\|_E$,
- **coincidence property:** $\forall x \in E \quad \|x\|_E = 0 \iff x = 0$.

A vector space E endowed with a norm $\|\cdot\|_E$ is called a *normed vector space.*

A norm induces a distance on E by setting the distance between any two $x, y \in E$ to be $\|x - y\|_E$.

A simple but important example of a norm is the absolute value on $\mathbb{R}$.

Now, our presentation above about the geometry of $\mathbb{R}$ can be applied to obtain a lot of basic results about normed vector spaces "as is", replacing absolute values with norms. We will however try for a more self-contained presentation of the essential results for our purpose. To this end, we start from the convergence of sequences with values in a normed vector space since we can easily bootstrap a lot of results straight from our knowledge about $\mathbb{R}$.

Definition 7.2. A sequence $(x_n)_{n\in\mathbb{N}}$ in a normed vector space E converges to $l \in E$ when the sequence $(\|x_n - l\|_E)_{n\in\mathbb{N}}$ converges to 0 (as a sequence of real numbers).

Therefore, $(x_n)_{n\in\mathbb{N}}$ converges to l if, and only if

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geqslant N \quad \|x_n - l\|_E < \varepsilon.$$

Indeed, norms are always nonnegative. In fact, we will very often use the following reasoning below: if we can find some sequence $(t_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ with limit 0 such that $\|x_n - l\|_E \leqslant t_n$, then using the squeeze theorem, we can conclude that $(x_n)_{n\in\mathbb{N}}$ converges to l.

Theorem 7.3. *If $(x_n)_{n\in\mathbb{N}}$ is a convergent sequence in a normed vector space E converging to both l and l' in E, then $l = l'$.*

Proof. We observe, for all $n \in \mathbb{N}$:

$$0 \leqslant \underbrace{\|l - l'\|_E}_{=l-x_n+x_n-l'} \leqslant \underbrace{\|l - x_n\|_E + \|x_n - l\|_E}_{\text{triangle inequality}} \xrightarrow{n\to\infty} 0 + 0 = 0.$$

Thus $l = l'$. $\square$

Notation 7.4. If $(x_n)_{n\in\mathbb{N}}$ is a convergent sequence in a normed vector space E, then the unique vector to which it converges is called the *limit* of $(x_n)_{n\in\mathbb{N}}$ and is denoted by $\lim_{n\to\infty} x_n$.

We now establish the continuity of the basic ingredients of a normed vector space: the norm, the addition, and the scalar multiplication. We define continuity using sequences here, to take the shortest route.

Definition 7.5. Let E and F be normed vector spaces. A function $f : D \subseteq E \to F$ from $D \subseteq E$ is *continuous at* $x \in E$ when for any sequence $(z_n)_{n\in E}$ in E converging to x, the sequence $(f(z_n))_{n\in\mathbb{N}}$ converges to $f(x)$ in F.

Definition 7.6. Let E and F be normed vector spaces. A function $f : D \subseteq E \to F$ from $D \subseteq E$ is *continuous on* D when it is continuous at every $x \in D$.

We begin with the continuity of the norm.

Theorem 7.7 (reverse triangle inequality)**.** *For all $x, y \in E$, we have*

$$|\|x\|_E - \|y\|_E| \leqslant \|x - y\|_E .$$

Proof. Since $\|x\|_E \leqslant \|x - y\|_E + \|y\|_E$, we conclude $\|x\|_E - \|y\|_E \leqslant \|x - y\|_E$. Similarly,

$$\|y\|_E - \|x\|_E \leqslant \|y - x\|_E = \|(-1)(x - y)\|_E = |1| \|x - y\|_E = \|x - y\|_E .$$

Thus our theorem follows. $\square$

Corollary 7.8. *If $(x_n)_{n\in\mathbb{N}}$ converges to z in a normed vector space E, then $\lim_{n\to\infty} \|x_n\|_E = \|z\|_E$.*

Proof. By the reverse triangle inequality,

$$0 \leqslant |\|x_n\|_E - \|z\|_E| \leqslant \|x_n - z\|_E \xrightarrow{n\to\infty} 0$$

hence our result follows from the squeeze theorem. $\square$

We now turn to the addition. It is useful to note that if E and F are two normed vector space, and if we define

$$\forall x \in E \quad \forall y \in F \quad \|(x,y)\|_{E\times F} \coloneqq \max\{\|x\|_E, \|y\|_F\}$$

then $\|\cdot\|_{E\times F}$ is indeed a norm on the product vector space $E\times F$, with the property that $(x_n, y_n)_{n\in\mathbb{N}}$ converges to (z,w) in $E\times F$ if, and only if, $(x_n)_{n\in\mathbb{N}}$ converges to z in E and $(y_n)_{n\in\mathbb{N}}$ converges to w in F. With this in mind, we prove the following.

Theorem 7.9. *If $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are two convergent sequences in a normed vector space E, then $(x_n+y_n)_{n\in\mathbb{N}}$ converges as well, and*

$$\lim_{n\to\infty}(x_n+y_n) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n.$$

Proof. Let $l_x \coloneqq \lim_{n\to\infty} x_n$ and $l_y \coloneqq \lim_{n\to\infty} y_n$. Then

$$\begin{aligned} 0 \leqslant \|(x_n+y_n)-(l_x+l_y)\|_E &\leqslant \|x_n-l_x\|_E + \|y_n-l_y\|_E \\ &\xrightarrow{n\to\infty} 0+0=0. \end{aligned}$$

Thus our theorem is proven. □

Theorem 7.10. *If $(t_n)_{n\in\mathbb{N}}$ is a convergent sequence in $\mathbb{R}$ and $(x_n)_{n\in\mathbb{N}}$ is a convergent sequence in a normed vector space E, then $(t_nx_n)_{n\in\mathbb{N}}$ converges in E as well, and*

$$\lim_{n\to\infty} t_nx_n = \lim_{n\to\infty} t \lim_{n\to\infty} x_n.$$

Proof. We write $l \coloneqq \lim_{n\to\infty} t_n$ and $z = \lim_{n\to\infty} x_n$. We then compute:

$$\begin{aligned} 0 \leqslant \|t_nx_n - lz\|_E &\leqslant \|t_nx_n - t_nz\|_E + \|t_nz - lz\|_E \\ &\leqslant |t_n|\,\|x_n - z\|_E + |t_n - l|\,\|z\|_E \\ &\xrightarrow{n\to\infty} |l|0 + 0\,\|z\|_E = 0. \end{aligned}$$

Thus, our result is proven. □

We characterize continuity just as we did for functions of a real variable.

Theorem 7.11. *Let E and F be normed vector spaces. A function $f : D\subseteq E\to F$ from $D\subseteq E$ is* continuous *at $x\in E$ if, and only if,*

$$\forall\varepsilon>0 \quad \exists\delta>0 \quad \forall z\in E \quad \|z-x\|_E<\delta \implies \|f(z)-f(x)\|_E<\varepsilon. \tag{7.1}$$

Proof. Assume Equation (7.1). If $(x_n)_{n\in\mathbb{N}}$ be a sequence in E converging to $x\in E$. Let $\varepsilon>0$. There exists $\delta>0$ such that if $t,s\in E$ with $\|t-x\|_E<\delta$, then $\|f(t)-f(x)\|_E<\varepsilon$. There exists $N\in\mathbb{N}$ such that if $n\geqslant N$, then $\|x_n-x\|_E<\delta$. Thus for all $n\geqslant N$, then $\|f(x_n)-f(x)\|_E<\varepsilon$. Thus f is continuous at x.

Assume Equation (7.1) does not hold. Thus, there exists $\varepsilon>0$ such that, for any $\delta>0$, there exists $x_\delta\in E$ such that $\|x-x_\delta\|_<\delta$ yet $\|f(x)-f(x_\delta)\|_F\geqslant\varepsilon$. Thus, for all $n\in\mathbb{N}$, there exists $x_n\in E$ with $\|x_n-x\|_E<\frac{1}{n+1}$ and $\|f(x_n)-f(x)\|_E\geqslant\varepsilon$. By construction, $(x_n)_{n\in\mathbb{N}}$ converges to x, yet $(f(x_n))_{n\in\mathbb{N}}$ does not converge to $f(x)$, so f is not continuous at x. □

A crucial result about continuity concerns linear functions between normed vector spaces.

Theorem 7.12. *Let $\varphi : E\to F$ be a linear function between two normed vector spaces E and F. The following assertions are equivalent.*

(1) φ is continuous,
(2) φ is continuous at 0,
(3) φ is bounded on the open unit sphere $\{x \in E : \|x\|_E < 1\}$ of E,
(4) φ is bounded on the unit sphere $\{x \in E : \|x\|_E = 1\}$ of E,
(5) φ is bounded on the closed unit sphere $\{x \in E : \|x\|_E \leqslant 1\}$ of E,
(6) there exists $K > 0$ such that, for all $x \in E$,

$$\|\varphi(x)\|_F \leqslant K \|x\|_E ,$$

(7) φ is Lipschitz.

Proof. Of course, (1) implies (2).

Assume (2). Therefore, there exists $\delta > 0$ such that if $x \in E$ with $\|x\|_E < \delta$ then $\|\varphi(x)\|_F < 1$. Let now $\|x\|_E < 1$. If $x = 0$ then $\|\varphi(x)\|_F = 0$. If $x \neq 0$, then we note that if $y \coloneqq \frac{\delta}{2\|x\|_E} x$, then

$$\|y\|_E = \frac{\delta}{2\|x\|_E} \|x\|_E = \frac{\delta}{2} < \delta$$

and thus $\|\varphi(y)\|_F < 1$. Hence $\|\varphi(x)\|_E \leqslant \frac{2}{\delta}$ by linearity of φ.
Therefore, φ is bounded by $\frac{2}{\delta}$ over the open unit ball of E. So (3) holds.

We continue with Assumption (2). Let $x \in E$ with $\|x\|_E = 1$. Then $\left\|\frac{\delta}{2}x\right\|_E = \frac{\delta}{2} < \delta$, so $\left\|\varphi(\frac{\delta}{2}x)\right\|_F \leqslant 1$ so $\|\varphi(x)\|_F < \frac{2}{\delta}$. So (4) holds.

Of course, (3) and (4) imply (5), so (2) implies (3), (4) and (5). Moreover, (5) implies (3) and (4).

We now assume (5). Thus there exists $K > 0$ such that for all $x \in E$, if $\|x\|_E \leqslant 1$, then $\|\varphi(x)\|_F \leqslant K$. Let $x \in E \setminus \{0\}$. Then

$$\begin{aligned} \|\varphi(x)\|_F &= \|x\|_E \left\| \varphi\left(\frac{1}{\|x\|_E} x\right) \right\|_F \\ &\leqslant \|x\|_E K, \end{aligned}$$

as claimed. Thus (6) holds.

Assume (6). Let $x, y \in E$. By linearity,

$$\|\varphi(x) - \varphi(y)\|_F = \|\varphi(x-y)\|_F \leqslant K \|x - y\|_E .$$

Thus φ is Lipschitz. So (7) holds.

Assume (7). Then (1) holds. So our theorem is proven. $\square$

Corollary 7.13. *Let E and F be two normed vector spaces. If $\mathscr{L}(E,F)$ be the set of all continuous linear functions from E to F, and if we set for any $\varphi \in \mathscr{L}(E,F)$:*

$$|||\varphi|||_E^F \coloneqq \sup\{\|\varphi(x)\|_F : \|x\|_E = 1\},$$

then $\left(\mathscr{L}(E,F), |||\cdot|||_E^F\right)$ is a normed vector space.

Proof. Let $\varphi, \psi \in \mathscr{L}(E,F)$ and $t \in \mathbb{R}$. Of course, $t\varphi + \psi$ is linear. Moreover, if $x \in E$ with $\|x\|_E = 1$, then

$$\|t\varphi(x) + \psi(x)\|_F \leqslant |t| \|\varphi(x)\|_F + \|\psi(x)\|_F \leqslant |t| |||\varphi|||_E^F + |||\psi(x)|||_E^F < \infty$$

so $t\varphi+\psi$ is bounded on the unit sphere of E, so it is continuous as a linear function, i.e. $t\varphi + \psi \in \mathscr{L}(E,F)$. Of course $0 \in \mathscr{L}(E,F)$. So $\mathscr{L}(E,F)$ is a subspace of the space of all functions from E to F. Moreover, $|||t\varphi + \psi|||_E^F \leqslant |t| |||\varphi|||_E^F + |||\psi|||_E^F$.

Last, $|||\varphi|||_E^F = 0$ for some $\varphi \in \mathscr{L}(E,F)$. Thus $\varphi(x) = 0$ for all $x \in E$ with $\|x\|_E = 1$. If $x \in E \setminus \{0\}$ then

$$\varphi(x) = \|x\|_E \, \varphi\left(\frac{1}{\|x\|_E} x\right) = \|x\|_E \, 0 = 0.$$

So $\varphi = 0$. This completes our proof. ☐

Normed vector spaces are metric spaces, and we can use these structures to discuss Cauchy sequences and uniform continuity.

Definition 7.14. A sequence $(x_n)_{n\in\mathbb{N}}$ is *Cauchy* in a normed vector space E when

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall p,q \in \mathbb{N} \quad p \geqslant N \text{ and } q \geqslant N \implies \|x_p - x_q\|_E < \varepsilon.$$

Theorem 7.15. *If $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in a normed vector space E, then $\sup_{n\in\mathbb{N}} \|x_n\|_E < \infty$.*

Proof. There exists $N \in \mathbb{N}$ such that if $p, q \geqslant N$, then $\|x_p - x_q\|_E < 1$, so for all $p \geqslant N$, we have $\|x_p\|_E < \|x_N\|_E + 1$. Thus $\sup_{n\in\mathbb{N}} \|x_n\|_E \leqslant \max\{\max_{j\in\{0,\ldots,N\}} \|x_j\|_E, 1 + \|x_N\|_E$. ☐

Theorem 7.16. *If $(x_n)_{n\in\mathbb{N}}$ is a convergent sequence in a normed vector space E, then $(x_n)_{n\in\mathbb{N}}$ is Cauchy.*

Proof. Let $l = \lim_{n\to\infty} x_n$. Let $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that for all $n \in \mathbb{N}$, if $n \geqslant N$, then $\|x_n - l\|_E < \frac{\varepsilon}{2}$.

Thus, for all $p, q \geqslant N$, we have

$$\|x_p - x_q\|_E \leqslant \|x_p - l\|_E + \|l - x_q\|_E < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This completes our proof. ☐

Not all normed vector spaces are complete. We introduce the following definition.

Definition 7.17. A normed vector space E is a *Banach space* when every Cauchy sequence in E converges.

We now proceed with the basic topology of normed vector spaces (taking in essence the opposite road to our work on $\mathbb{R}$).

Definition 7.18. Let E be a normed vector space. The *closure* $A \subseteq E$ of a subset A of E is

$$\mathrm{cl}\,(A) \coloneqq \left\{\lim_{n\to\infty} x_n : (x_n)_{n\in\mathbb{N}} \text{ is a convergent sequence in } E\right\}.$$

Theorem 7.19. *If E is a normed vector space, and if $A \subseteq E$, then*

$$\mathrm{cl}\,(A) \coloneqq \left\{x \in E : \inf_{t\in A} \|x - t\|_E = 0\right\}.$$

Proof. If $x \in \mathrm{cl}(A)$ then there exists $(x_n)_{n\in\mathbb{N}}$ in A such that $\lim_{n\to\infty} \|x_n - x\|_E = 0$; hence for all $\varepsilon > 0$, there exists $n \in \mathbb{N}$ such that $\|x_n - x\|_E < \varepsilon$; since $x_n \in A$, we conclude $\inf_{t\in A} \|x - t\|_E = 0$.

Conversely, assume $\inf_{t\in A} \|x - t\|_E = 0$. Let $n \in \mathbb{N}$. There exists $t_n \in A$ such that $\|t - x\|_E < \frac{1}{n+1}$. Thus by definition, $x = \lim_{n\to\infty} t_n$. This proves our equality. □

Theorem 7.20. *Let E be a normed vector spaces. The following assertions hold.*

1. $\mathrm{cl}(\emptyset) = \emptyset$,
2. $\forall A \subseteq E \quad A \subseteq \mathrm{cl}(A)$,
3. $\forall A \subseteq E \quad \mathrm{cl}(\mathrm{cl}(A)) = \mathrm{cl}(A)$,
4. $\forall A, B \subseteq E \quad \mathrm{cl}(A \cup B) = \mathrm{cl}(A) \cup \mathrm{cl}(B)$,
5. $\forall A, B \subseteq E \quad A \subseteq B \implies \mathrm{cl}(A) \subseteq \mathrm{cl}(B)$.

Proof. The proof follows the same path as in the real case. □

Thus the closure operator is used to define a topology on E.

Definition 7.21. Let E be a normed vector space. A subset A of E is closed when $A = \mathrm{cl}(A)$. A subset B of E is closed if its complement $E \setminus A$ is closed.

We can now apply all the topological proofs on $\mathbb{R}$ to conclude the following assertions.

Theorem 7.22. *Let E be a normed vector space. The following assertions hold.*

1. *$\emptyset$ and E are both closed and open,*
2. *if $U, V \subseteq E$ are open, then $U \cap V$ is open,*
3. *if $(U_j)_{j\in J}$ is a family of open sets then $\bigcup_{j\in J} U_j$ is open,*
4. *if F, G are closed, then $F \cup G$ is closed,*
5. *if $(F_j)_{j\in J}$ is a family of closed sets then $\bigcap_{j\in J} F_j$ is closed,*
6. *the closure $\mathrm{cl}(A)$ of a subset $A \subseteq E$ is the smallest closed subset of E containing A,*
7. *the closure $\mathrm{cl}(A)$ is given by*
$$\mathrm{cl}(A) = \{x \in E : \exists U \text{ open} \quad x \in U, U \cap A \neq \emptyset\},$$
8. *$f : E \to F$ is continuous over E if, and only if, $f^{-1}(G)$ is closed for all closed subset G of F,*
9. *$f : E \to F$ is continuous over E if, and only if, $f^{-1}(U)$ is open for all open subset U of F.*

We now establish a general tool to extend continuous linear functions from a dense subspace.

Theorem 7.23. *Let E be a normed vector space and let F be a Banach space. Let $G \subseteq E$ be a subspace of E such that $\mathrm{cl}(G) = E$. If $\varphi : G \to F$ is a continuous linear function, then there exists a unique continuous linear function $\psi : E \to F$ whose restriction to G is ψ; moreover $|||\psi|||_E^F = |||\varphi|||_G^F$.*

Proof. Since φ is continuous, there exists $K > 0$ such that, for all $x \in E$, we have $\|\varphi(x)\|_F \leqslant K \|x\|_E$.

Let $x \in E$. Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in E which converge to x. Thus $(x_n)_{n\in\mathbb{N}}$ is Cauchy in E. Since, for all $p, q \in \mathbb{N}$, we have $\|\varphi(x_p) - \varphi(x_q)\|_F \leqslant \|x_p - x_q\|_E$, we conclude that $(\varphi(x_n))_{n\in\mathbb{N}}$ is a Cauchy sequence in F. Since F is complete, $(\varphi(x_n))_{n\in\mathbb{N}}$ converges to some $y \in F$. For now, our number depends on the choice of the sequence $(x_n)_{n\in\mathbb{N}}$.

Assume for this paragraph that $x = 0$. Then observe that, for any sequence $(x_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}}$, we have

$$|\varphi(x_n)| \leqslant K \|x_n\|_E \xrightarrow{n\to\infty} K \|0\|_E = 0.$$

So, in that case, no matter what our choice of $(x_n)_{n\in\mathbb{N}}$ is, we always have $\lim_{n\to\infty} \varphi(x_n) = 0$.

Let us now return to a general $x \in E$. If $(x_n)_{n\in\mathbb{N}}$ and $(x'_n)_{n\in\mathbb{N}}$ both converge uniformly to x in E, then $\lim_{n\to\infty} \varphi(x_n - x'_n) = 0$, since $(x_n - x'_n)_{n\in\mathbb{N}}$ converges to 0 in E. On the other hand, by linearity, $\varphi(x_n - x'_n) = \varphi(x_n) - \varphi(h_n)$, and we have shown that $(\varphi(g_n))_{n\in\mathbb{N}}$ and $(\varphi(h_n))_{n\in\mathbb{N}}$ converge. We thus conclude:

$$\lim_{n\to\infty} \varphi(h_n) = \lim_{n\to\infty} \varphi(g_n).$$

Thus our limit does not depend on the choice of the sequence $(x_n)_{n\in\mathbb{N}}$ in E converging to x. We denote this limit by $\psi(x)$.

We immediately observe that ψ is linear: if $x, y \in E$ and $t \in \mathbb{R}$, then $x = \lim_{n\to\infty} x_n$ and $y = \lim_{n\to\infty} y_n$ for two sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ in G. Then $(tx_n + y_n)_{n\in\mathbb{N}}$ converges to $tx + y$, and

$$\begin{aligned}\psi(tx+y) &= \lim_{n\to\infty} \varphi(tx_n + y_n) \\ &= \lim_{n\to\infty} t\varphi(x_n) + \varphi(y_n) \\ &= t\psi(x) + \psi(y).\end{aligned}$$

Moreover, $\|\psi(x)\|_F = \lim_{n\to\infty} \|\varphi(x_n)\|_F \leqslant \lim_{n\to\infty} |||\varphi|||_G^F \|x_n\|_E \leqslant |||\varphi|||_G^E \|x\|_E$. So ψ is continuous over E with norm $|||\varphi|||_G^F$.

Last, assume θ is a continuous linear function over E whose restriction to G is φ. Then $\theta - \psi$ is zero on G, i.e. $G \subseteq \ker(\theta - \psi)$. On the other hand, $\ker(\theta - \psi) = (\theta - \psi)^{-1}(\{0\})$ and $\{0\}$ is closed in F, so $\ker(\theta - \psi)$ is closed. Therefore, $E = \mathrm{cl}\,(G) \subseteq \mathrm{cl}\,(\ker(\theta - \psi)) = \ker(\theta - \psi)$ and thus $\theta = \psi$. □

Remark 7.24. This result is just a very special case of the unique extension property of uniformly continuous functions from a dense subspace to the entire space.

7.2. **Uniform Norms.**

Definition 7.25. If $f : D \to \mathbb{R}$ is a function over a set D, we set

$$\|f\|_D \coloneqq \sup\{|f(x)| : x \in D\},$$

allowing for the value ∞. We call $\|\cdot\|_D$ the *uniform norm* over D (note: it can be infinite so technically, it is not a norm).

Definition 7.26. A function $f : D \to \mathbb{R}$ over a set D is bounded when $\|f\|_D < \infty$.

Theorem 7.27. *The set of all bounded functions over a set D, denoted by $\mathscr{B}(D)$, is a normed vector space with the uniform norm $\|\cdot\|_D$.*

Proof. Let $f, g \in \mathscr{B}(D)$ and $t \in D$. If $x \in D$ then $|tf(x) + g(x)| \leqslant |t| \|f\|_D + \|g\|_D$. Thus $\|tf + g\|_D \leqslant |t| \|f\|_D + \|g\|_D$. Since $0 \in \mathscr{B}(D)$ (with $\|0\|_D = 0$), we conclude that $\mathscr{B}(D)$ is a subspace of the space of all functions form E to F. Since $\|f\|_D = 0$

implies $|f(x)| = 0$ for all $x \in D$, we conclude that $\|f\|_D = 0 \implies f = 0$, so $\|\cdot\|_D$ is a norm as well. □

Definition 7.28. A sequence $(f_n)_{n\in\mathbb{N}}$ of functions defined over a set D converges *uniformly on D* when

$$\lim_{n\to\infty} \|f_n - f\|_D = 0.$$

Theorem 7.29. *If $(f_n : D \to \mathbb{R})_{n\in\mathbb{N}}$ converges to f uniformly on D, then $(f_n(x))_{n\in\mathbb{N}}$ converges to $f(x)$ for all $x \in D$.*

Proof. Simply note that $|f(x)-f_n(x)| \leqslant \|f - f_n\|_D$ and apply the squeeze theorem. □

When working with *bounded* functions over D, we can understand the previous theorem as follows.

Theorem 7.30. *Let $D \subseteq \mathbb{R}$ and let $x \in D$. The* evaluation map *$f \in \mathscr{B}(D) \mapsto f(x) \in \mathbb{R}$ is a continuous linear map.*

Proof. We simply observe that $|f(x)| \leqslant \|f\|_D$ for all $f \in \mathscr{B}(D)$. □

Now, a deeper result is the completeness of the space of bounded functions.

Theorem 7.31. *The space $\mathscr{B}(D)$ is complete for any nonempty $D \subseteq \mathbb{R}$.*

Proof. Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence for the uniform norm. Let $x \in D$. Since the evaluation map at x is Lipschitz, the sequence $(f_n(x))_{n\in\mathbb{N}}$ is Cauchy. Since $\mathbb{R}$ is complete, the sequence $(f(x_n))_{n\in\mathbb{N}}$ is convergent; let $f(x) = \lim_{n\to\infty} f_n(x)$. Since for all $x \in D$, we have $|f_n(x)| \leqslant \|f_n\|_D$, and $(\|f_n\|_D)_{n\in\mathbb{N}}$ is convergent, so it is bounded. Hence, $f : D \to \mathbb{R}$ is bounded.

It remains to prove that $(f_n)_{n\in\mathbb{N}}$ converges uniformly to f over D. Let $\varepsilon > 0$. Since $(f_n)_{n\in\mathbb{N}}$ is Cauchy for the uniform norm, there exists $N \in \mathbb{N}$ such that, for all $p, q \geqslant N$, we have $\|f_p - f_q\|_D < \frac{\varepsilon}{2}$. Let $n \geqslant N$.

Let now $x \in D$. Since $(f_n(x))_{n\in\mathbb{N}}$ converges to $f(x)$, there exists $M_x \in \mathbb{N}$ such that, if $m \geqslant M_x$, then $|f_n(x) - f(x)| < \frac{\varepsilon}{2}$. Therefore, if $p \geqslant \max\{N, M_x\}$, then:

$$\begin{aligned}|f(x) - f_n(x)| &\leqslant |f(x) - f_p(x)| + |f_p(x) - f_n(x)| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

Thus $\|f_n - f\|_D < \varepsilon$. This concludes our proof. □

7.3. **The spaces of continuous functions.** Since linear combinations of continuous functions are continuous, the following concept is well-defined.

Definition 7.32. The subspace of bounded continuous functions over a subset $D \subseteq \mathbb{R}$ is denoted by $C_b(D)$.

We now prove that $C_b(D)$ is closed in $\mathscr{B}(D)$. In particular, it is complete.

Theorem 7.33. *Let $D \subseteq \mathbb{R}$ and $a \in \mathrm{cl}\,(D)$. If $(f_n)_{n\in\mathbb{N}}$ is sequence of functions over D such that, for all $n \in \mathbb{N}$, the function f_n has a limit l_n at a along D, then:*

1. *$(l_n)_{n\in\mathbb{N}}$ converges to some $l \in \mathbb{R}$,*
2. *the function $f : D \to \mathbb{R}$ has limit l at a along D.*

Proof. Let $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that, if $n \geqslant N$, then $\|f_n - f\|_D < \frac{\varepsilon}{3}$.

Moreover, for each $n \in \mathbb{N}$, there exists $\delta_n > 0$ such that if $x \in D$ with $|x-a| < \delta_n$ then $|f_n(x) - l_n| < \frac{\varepsilon}{3}$.

Let $p, q \geqslant N$. Let $x \in D$ such that $|x - a| < \min\{\delta_p, \delta_q\}$, which exists since $a \in \mathrm{cl}(D)$. Then

$$\begin{aligned} |l_p - l_q| &\leqslant |l_p - f_p(x)| + |f_p(x) - f_q(x)| + |f_q(x) - l_q| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Thus, $(l_n)_{n\in\mathbb{N}}$ is Cauchy in $\mathbb{R}$. Since $\mathbb{R}$ is complete, $(l_n)_{n\in\mathbb{N}}$ converges to some $l \in \mathbb{R}$.

Let now $N' \in \mathbb{N}$ such that, for all $n \geqslant M$, we have $|l_n - l| < \frac{\varepsilon}{3}$. Let $M = \max\{N, N'\}$.

For all $x \in D$, if $|x - a| < \delta_M$, then:

$$\begin{aligned} |f(x) - l| &\leqslant |f(x) - f_M(x)| + |f_M(x) - l_M| + |l_M - l| \\ &\leqslant \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Therefore, f converges to l at a along D. □

Corollary 7.34. *Let $D \subseteq \mathbb{R}$ and $x \in D$. If $(f_n)_{n\in\mathbb{N}}$ is a sequence of functions over D which converges uniformly to some f, and if f_n is continuous at x for all $n \in \mathbb{N}$, then f is continuous at x.*

Proof. Since $f_n(x) = \lim_{\substack{t\to x\\ t\in D}} f_n(t)$ for each $n \in \mathbb{N}$, and since $\lim_{n\to\infty} f_n(x) = f(x)$, our result follows from Theorem (7.33). □

Theorem 7.35. *The space $C_b(D)$ is a Banach space for the uniform norm.*

Proof. Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $C_b(D)$ for the uniform norm. By Theorem (7.33), the sequence $(f_n)_{n\in\mathbb{N}}$ converges uniformly to a bounded function $f : D \to \mathbb{R}$ over D. By Corollary (7.34), the function f is continuous over D. □

We conclude this section by exhibiting a dense subspace of functions in $C_b(K)$ for any compact set $K \subseteq \mathbb{R}$. The idea here is to introduce some class of functions on which, for instance, defining the integral as a continuous linear function is relatively easy, and then extend our integral to the closure of that space of nice functions. This is one reason, among many, to find dense subspaces of continuous functions. We thus have the following.

Theorem 7.36. *If $f : K \to \mathbb{R}$ is a continuous function over a compact subset $K \subseteq \mathbb{R}$ of $\mathbb{R}$, then for all $\varepsilon > 0$, there exists a continuous piecewise affine function $g : K \to \mathbb{R}$ such that $\|f - g\|_K < \varepsilon$.*

Proof. Let $\varepsilon > 0$. Since $f : K \to \mathbb{R}$ is continuous over the compact set K, it is uniformly continuous. Thus, there exists $\delta > 0$ such that, for all $x, y \in K$, if $|x - y| < \delta$, then $|f(x) - f(y)| < \varepsilon$. Let $N \coloneqq \max\{n \in \mathbb{N} : a + n\delta \leqslant b\}$, and for all $j \in \{0, \ldots, N\}$, let $\sigma_j \coloneqq a + j\delta$ and $\sigma_{N+1} \coloneqq b$.

For all $x \in [a, b]$, there exists $j \in \{0, \ldots, N\}$ such that $\sigma_0 \leqslant x < \sigma_{j+1}$, and we set

$$g(x) \coloneqq \frac{x - \sigma_j}{\sigma_{j+1} - \sigma_j} f(\sigma_j) + \frac{\sigma_{j+1} - x}{\sigma_{j+1} - \sigma} f(\sigma_{j+1}).$$

By construction, g is continuous and piecewise affine. Moreover, for all $x \in K$, we estimate:

$$\begin{aligned}|g(x) - f(x)| &= \left| \frac{x - \sigma_j}{\sigma_{j+1} - \sigma_j} f(\sigma_j) + \frac{\sigma_{j+1} - x}{\sigma_{j+1} - \sigma_j} f(\sigma_{j+1}) - f(x) \right| \\ &\leqslant \frac{1}{\sigma_{j+1} - \sigma_j} \left(|x - \sigma_j| |f(\sigma_j) - f(x)| + |x - \sigma_{j+1}| |f(\sigma_{j+1}) - f(x)| \right) \\ &\leqslant \frac{1}{\sigma_{j+1} - \sigma_j} ((x - \sigma_j) + (\sigma_{j+1} - x))\varepsilon \\ &= \varepsilon.\end{aligned}$$

So $\|f - g\|_K < \varepsilon$, as claimed. □

8. The regulated integral

8.1. Subdivisions.

Definition 8.1. A *subdivision* $(\sigma_j)_{j=0}^n$ of $[a, b]$, where $n \in \mathbb{N} \setminus \{0\}$, is a family of numbers in $[a, b]$ indexed by $\{0, \ldots, n\}$, such that $a = \sigma_0 < \sigma_1 < \cdots < \sigma_n = b$. The *size* $\#\sigma$ of σ is n.

Theorem 8.2. *If $F \subseteq [a, b]$, there is a unique subdivision σ of $[a, b]$ such that $\{\sigma_j : j \in \{0, \ldots, \#\sigma\} = F \cup \{a, b\}$.*

Proof. This is immediate via a simple induction since the order on $\mathbb{R}$ is linear. □

Definition 8.3. If σ and η are two subdivisions of $[a, b]$, then we declare $\sigma \leqslant \eta$ when $\{\eta_0, \ldots, \eta_n\} \subseteq \{\sigma_0, \ldots, \sigma_n\}$.

Theorem 8.4. *The relation $\leqslant$ is an order on the set of subdivisions of $[a, b]$.*

Proof. This is immediate (note that $\sigma \leqslant \eta$ and $\eta \leqslant \sigma$ implies these two subdivisions have the same range, and there is only one subdivision with a given range). □

Definition 8.5. If σ and η are two subdivisions of $[a, b]$, then we denote by $\sigma \wedge \eta$ the unique subdivision of $[a, b]$ with range $\{\sigma_0, \ldots, \sigma_{\#\sigma}\} \cup \{\eta_0, \ldots, \eta_{\#\eta}\}$.

Theorem 8.6. *If σ and η are two subdivisions of $[a, b]$, then $\sigma \wedge \eta$ is the infimum of $\{\sigma, \eta\}$.*

Proof. By construction, $\sigma \wedge \eta \leqslant \sigma$ and $\sigma \wedge \eta \leqslant \eta$. Moreover, let ζ be a subdivision of $[a, b]$ such that $\zeta \leqslant \sigma$ and $\zeta \leqslant \eta$. Then $\zeta \leqslant \sigma \wedge \eta$ by construction. □

8.2. Our Scheme.

We can now use our work to define our first integral on the space of continuous functions over a compact interval, as a special continuous linear functional. To this end, we first identify a dense subspace of $C_b([a, b])$ where it is easy to construct the integral as a continuous linear function. We propose here the continuous piecewise affine functions as our “nice functions”.

We then define the integral on the space of “nice functions”. We will say a subdivision σ of $[a, b]$ is *adapted* to a piecewise affine continuous function $f : [a, b] \to \mathbb{R}$ when f restricts to an affine function on each of $[\sigma_{j-1}, \sigma_j]$ for $j \in \{1, \ldots, \#\sigma\}$.

Lemma 8.7. *If $f : [a, b] \to \mathbb{R}$ is a continuous piecewise affine function, and if σ and η are two subdivisions of $[a, b]$ adapted to f, then*

$$\sum_{j=1}^{\#\sigma} \frac{\sigma_j - \sigma_{j-1}}{2} \left(f(\sigma_j) + f(\sigma_{j-1}) \right) = \sum_{j=1}^{\#\eta} \frac{eta_j - \eta_{j-1}}{2} \left(f(\eta_j) + f(\eta_{j-1}) \right).$$

Proof. First, assume $\sigma \leqslant \eta$. Then

$$\begin{aligned}\sum_{j=1}^{\#\sigma} \frac{(\sigma_j - \sigma_{j-1})}{2}\left(f(\sigma_j) + f(\sigma_{j-1})\right) &= \sum_{j=1}^{\#\eta} \sum_{\{k:\eta_{j-1} < \sigma_k \leqslant \eta_{j+1}\}} \frac{\sigma_k - \sigma_{k-1}}{2}\left(f(\sigma_k) + f(\sigma_{k-1})\right) \\ &= \sum_{j=1}^{\#\eta} \frac{\eta_j - \eta_{j-1}}{2}\left(f(\eta_j) + f(\eta_{j-1})\right).\end{aligned}$$

Now, assume η and σ are adapted to f. Then $\zeta \coloneqq \eta \wedge \sigma$ is also adapted to f, and our result follows from the transitivity of equality. □

Definition 8.8. If $f : [a, b] \to \mathbb{R}$ is a continuous piecewise affine function with adapted subdivision $\sigma_0 < \sigma_1 < \cdots < \sigma_n$, then we define

$$\int_a^b f \coloneqq \sum_{j=1}^{n} \frac{(\sigma_j - \sigma_{j-1})}{2}\left(f(\sigma_j) + f(\sigma_{j-1})\right).$$

We then check that our integral is indeed linear and continuous.

Theorem 8.9. *If f, g are continuous piecewise affine over $[a, b]$, and if $t \in \mathbb{R}$, then $\int_a^b (tf + g) = t \int_a^b f + \int_a^b g$.*

Proof. Let σ be a subdivision of $[a, b]$ adapted to f and η be a subdivision of $[a, b]$ adapted to g. Let $\zeta = \sigma \wedge \eta$. Then ζ is adapted to both f and g. A direct computation then concludes our proof. □

Theorem 8.10. *If $f \in C_b([a, b])$ is piecewise affine, then $\left|\int_a^b f\right| \leqslant \int_a^b |f| \leqslant (b - a)\|f\|_{[a,b]}$.*

Proof. We observe first that $|f|$ is a continuous piecewise affine function over $[a, b]$. We simply compute:

$$\begin{aligned}\left|\int_a^b f\right| &= \left|\sum_{j=1}^{n} \frac{\sigma_j - \sigma_{j-1}}{2}(f(\sigma_j) + f(\sigma_{j-1})\right| \\ &\leqslant \sum_{j=1}^{n} \frac{\sigma_j - \sigma_{j-1}}{2}(|f(\sigma_j)| + |f(\sigma_{j-1})|) \\ &= \int_a^b |f| \\ &\leqslant \sum_{j=1}^{n} \frac{\sigma_j - \sigma_{j-1}}{2}(\|f\|_{[a,b]} + \|f\|_{[a,b]}) \\ &= (b-a)\|f\|_{[a,b]}.\end{aligned}$$

This concludes our proof. □

Thus, we can extend uniquely $\int_a^b$, using Theorem (7.23), as a continuous linear function of norm $b - a$ to the closure of the space of continuous piecewise affine functions over $[a, b]$, which is the entire space of continuous functions over $[a, b]$.

Thus defined, our integral on $C[a, b]$ is linear and continuous. It is easy from the definition to also check that $\int_a^b f = \int_a^c f + \int_c^b f$ for any $f \in C[a, b]$ and $c \in [a, b]$.

It is easy to check that, for any $g \in C_b[a,b]$, the function $f \in C_b[a,b] \mapsto \int_a^b fg$ is again a continuous linear function. Thus, we have discovered many continuous linear functionals of the space $C_b[a,b]$: integration with some "density" function and evaluation maps (and all their linear combinations). While this is not an exhaustive list, it is indeed progress toward understanding the continuous linear functionals of $C[a,b]$ — a full picture is indeed all about integration, but in the sense of Lebesgue.

Indeed, one interesting feature of integration is that it can be extended to much more general spaces of functions. We now repeat our scheme but we go out of the space of continuous functions; we remain within the general Riemannian framework, however.

8.3. Regulated functions.

Definition 8.11. A function $f : [a,b] \to \mathbb{R}$ is a *step function* when there exists some subdivision $(\sigma_j)_{j=0}^n$ of $[a,b]$ such that the restriction of f to (σ_j, σ_{j+1}) is constant for all $j \in \{0, \ldots, n\}$.

Theorem 8.12. *The space $\mathscr{E}([a,b])$ of step functions on $[a,b]$ is a vector space.*

Proof. Straightforward. □

Definition 8.13. A function $f : [a,b] \to \mathbb{R}$ is *regulated* over $[a,b]$ when for all $\varepsilon > 0$, there exists a step function $g : [a,b] \to \mathbb{R}$ such that $\|f - g\|_{[a,b]} < \varepsilon$.

Theorem 8.14. *The space $\mathscr{R}([a,b])$ of all regular functions over $[a,b]$ is a vector space.*

Theorem 8.15. *A function $f : [a,b] \to \mathbb{R}$ is regulated over $[a,b]$ if, and only if, the function f has a left limit at every $t \in (a,b]$ and a right limit at every $t \in [a,b)$.*

Proof. First, assume f is regulated. Since every step function has a left limit at every point in $(a,b]$, and f is the uniform limit of step functions, f has a left at every point of $(a,b]$ as well, by Theorem (7.33). The same reasoning applies for the right limits.

Second, assume that f has left and right limits at every point in $[a,b]$. Let $K = \left\{c \in [a,b] : \exists g \text{ step function } \|f-g\|_{[a,c]} < \varepsilon\right\}$. Note that $a \in K$ and K is bounded above by b, so K has a supremum d.

We first prove that $d \in K$. Since f has a left limit at d, there exists $\delta > 0$ such that for all $t \in [a,d]$, if $d - t < \delta$, then $|\lim_{d^-} f - f(t)| < \varepsilon$. Since $d = \sup K$, there exists $t \in K$ with $d - t < \delta$. Thus, there exists a step function $g : [a,t] \to \mathbb{R}$ such that $\|f - g\|_{[a,t]} < \varepsilon$. Now, we define the following function:

$$h : x \in [a,d] \longmapsto \begin{cases} g(x) & \text{if } x \in [a,t], \\ \lim_{d^-} f & \text{if } x \in (t,d), \\ f(d) & \text{if } x = d. \end{cases}$$

The function h is a step function by construction, and $\|f - h\|_{[a,d]} < \varepsilon$ by our choice of δ. Therefore, $d \in K$.

We now prove that $d = b$. Assume that $d < b$. Since f then has a right limit at d, we conclude that there exists $\delta > 0$ such that if $t \in [d,b]$ and $t - b < \delta$ then

$|f(t) - \lim_{d^+} f| < \varepsilon$. Since $d \in K$, there exists a step function $g : [a, d] \to \mathbb{R}$ such that $\|f - g\|_{[a,d]} < \varepsilon$. Write $\varpi = \frac{\delta}{2} > 0$. We now set:

$$h : t \in [a, d + \delta] \longmapsto \begin{cases} g(t) \text{ if } t \in [a, d], \\ \lim_{d^+} f \text{ if } t \in (d, d + \varpi]. \end{cases}$$

Then h is a step function over $[a, d + \varpi]$ with $\varpi > 0$ such that $\|f - h\|_{[a,d+\varpi]} < \varepsilon$, by construction. Thus $d + \varpi \in K$, which is a contradiction since $d + \varpi > d$. So $d = b$.

Therefore, $b \in K$, and therefore, there exists a step function $g : [a, b] \to \mathbb{R}$ such that $\|f - g\|_{[a,b]} < \varepsilon$, as claimed. $\square$

Corollary 8.16. *Every continuous function on $[a, b]$ is regulated on $[a, b]$.*

8.4. **Integral of step functions.**

Lemma 8.17. *If φ is a step function, and both σ and η are subdivisions of $[a, b]$ adapted to φ, then*

$$\sum_{j=1}^{\#\sigma} (\sigma_j - \sigma_{j-1})\varphi(\xi_j) = \sum_{j=1}^{\#\eta} (\eta_j - \eta_{j-1})\varphi(\zeta_j)$$

for any family $(\xi_1, \ldots, \xi_{\#\sigma})$ such that $\sigma_0 < \xi_1 < \sigma_1 < \xi_2 < \cdots < \sigma_{\#\sigma}$ and any family $(\zeta_1, \ldots, \zeta_{\#\eta})$ such that $\eta_0 < \zeta_1 < \eta_1 < \zeta_2 < \cdots < \eta_{\#\eta}$.

Proof. First, assume $\sigma \leqslant \eta$. Then, noting that φ is constant on any interval of the form $[\eta_{j-1}, \eta]$ for $j \in \{1, \ldots, \#\eta - 1\}$,

$$\begin{aligned} \sum_{j=1}^{\#\sigma} (\sigma_j - \sigma_{j-1})\varphi(\xi_j) &= \sum_{j=1}^{\#\eta} \sum_{\{k : \eta_{j-1} < \sigma_k \leqslant \eta_j\}} (\sigma_j - \sigma_{j-1})\varphi(\zeta_j) \\ &= \sum_{j=1}^{\#\eta} (\eta_j - \eta_{j-1})\varphi(\zeta_j). \end{aligned}$$

We obtain our theorem by setting, in general, $\sigma' = \sigma \wedge \eta$. $\square$

Definition 8.18. If $\varphi \in \mathscr{E}[a, b]$ then we define

$$\int_a^b \varphi \coloneqq \sum_{j=1}^{\#\sigma} (\sigma_j - \sigma_{j-1})\varphi(\xi_j)$$

for some subdivision σ of $[a, b]$ adapted to φ, and for some numbers $\xi_1, \ldots, \xi_{\#\sigma}$ such that

$$\sigma_0 < \xi_1 < \sigma_1 < \xi_2 < \sigma_2 < \cdots < \sigma_{\#\sigma}.$$

We check that our integral is indeed linear and continuous.

Theorem 8.19. *For all $\varphi, \psi \in \mathscr{E}[a, b]$ and for all $t \in \mathbb{R}$, we have:*

$$\int_a^b (t\varphi + \psi) = t \int_a^b \varphi + \int_a^b \psi.$$

Proof. This follows from a direct computation. $\square$

Theorem 8.20. *For all $f \in \mathscr{E}[a,b]$, we have:*

$$\left| \int_a^b f \right| \leqslant \int_a^b |f| \leqslant (b-a) \|f\|_{[a,b]} .$$

Proof. We simply note that if $f \in \mathscr{E}[a,b]$ then $|f| \in \mathscr{E}[a,b]$, and moreover:

$$\begin{aligned} \left| \int_a^b \varphi \right| &\leqslant \sum_{j=1}^{\#\sigma} (\sigma_j - \sigma_{j-1}) \left| \varphi(\xi_j) \right| \\ &= int_a^b |f| \\ &\leqslant \sum_{j=1}^{\#\sigma} (\sigma_j - \sigma_{j-1}) \|f\|_{[a,b]} = (b-a) \|f\|_{[a,b]} . \end{aligned}$$

This completes our proof. ∎

Thus the integral of step function is a continuous linear function over the space $\mathscr{E}[a,b]$. It has two additional and helpful properties, which are both very easy to check.

Theorem 8.21. *If $\varphi \in \mathscr{E}[a,b]$ is nonnegative, then $\int_a^b \varphi \geqslant 0$.*

Theorem 8.22 (Chasles' relation)**.** *If $\varphi \in \mathscr{E}[a,b]$ and $c \in [a,b]$ then $\int_a^b \varphi = \int_a^c \varphi + \int_c^b \varphi$.*

Note that $\int_a^c \varphi = \int_a^b \varphi \cdot 1_{[a,c]}$, so Chasles' relation is a consequence of linearity.

8.5. **Integral of regulated functions.**

Definition 8.23. The regulated integral $\int_a^b$ over the space $\mathscr{R}[a,b]$ of regulated functions over $[a,b]$ is the unique, continuous linear extension of the integral on step functions.

From all our efforts, we derive:

Theorem 8.24. *The regulated integral is a continuous linear functional over $\mathscr{R}[a,b]$ of norm $b-a$ and such that*

$$\forall f \in \mathscr{R}[a,b] \quad \left| \int_a^b f \right| \leqslant \int_a^b |f| \leqslant (b-a) \|f\|_{[a,b]} ,$$

and for all $c \in [a,b]$, and for all $f \in [a,b]$,

$$\int_a^b f = \int_a^c f + \int_c^b f .$$

We note that $\int_a^b 1 = b - a = \left|\left|\left| \int_a^b \right|\right|\right|_{\mathscr{R}[a,b]}^{\mathbb{R}}$, which implies the *positivity* of the integral as well.

Theorem 8.25. *If $f \in \mathscr{R}[a,b]$ and $f \geqslant 0$, then $\int_a^b f \geqslant 0$.*

Proof. First, observe that if $f \geqslant 0$ then $\left\| \|f\|_{[a,b]} - f \right\|_{[a,b]} \leqslant \|f\|_{[a,b]}$. We then have:

$$\|f\|_{[a,b]} \left\| \left| \int_a^b \right| \right\| - \int_a^b f = \int_a^b (\|f\|_{[a,b]} - f) \leqslant \left\| \left| \int_a^b \right| \right\| \left\| \|f\|_{[a,b]} - f \right\|_{[a,b]} \leqslant \left\| \left| \int_a^b \right| \right\| \|f\|_{[a,b]} .$$

Therefore, $\int_a^b f \geqslant 0$. □

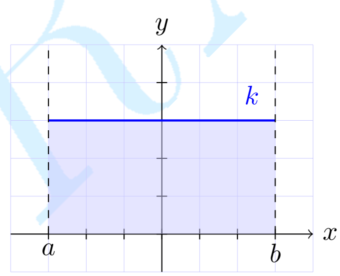
We now have two integrals defined on the space of continuous functions over a compact interval. Choosing another set of "nice" functions could lead to many other constructions. We need to reconcile all these integrals — and learn how to actually compute them. This is the matter of the following section.

9. The integral of continuous functions over segment: characterization

In this section, we prove that there is only one possible positive linear functional over the space of continuous functions over $[a, b]$, which maps the constant 1 to $b - a$, and which has the Chasles property. In fact, we will use only three basic properties which we have established above: Chasles' relation, positivity, and the normalization.

Hypothesis 9.1. We will assume given a real-valued function which, to any segment $[a, b]$ of $\mathbb{R}$ and any continuous function $f : [a, b] \to \mathbb{R}$, associates a real number $\int_a^b f$ (also denoted by $\int_a^b f(x)\,dx$ with x a dummy variable), which satisfies the following three Conditions (9.2), (9.3), and (9.4).

Condition 9.2. *For all $k \in \mathbb{R}$, if $a \leqslant b \in \mathbb{R}$, then $\int_a^b k = k(b - a)$.*



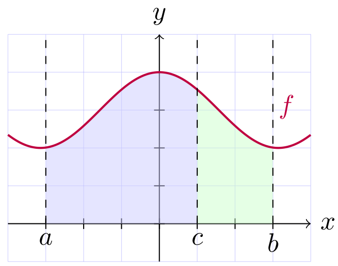
Condition (9.2): $\int_a^b k = k(b - a)$.

Condition 9.3 (Normalization)**.** *For all $a \leqslant b \in \mathbb{R}$, for all continuous functions $f : [a, b] \to \mathbb{R}$, the following assertion holds for all $c \in [a, b]$:*

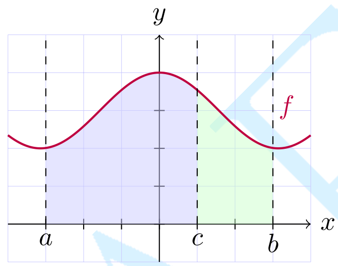
$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Condition 9.4 (Positivity)**.** *For all $a \leqslant b \in \mathbb{R}$, for all continuous functions $f : [a, b] \to \mathbb{R}$ and $g : [a, b] \to \mathbb{R}$, if*

$$\forall x \in [a, b] \quad f(x) \leqslant g(x).$$



Condition (9.3): *If* $a \leqslant c \leqslant b$, *then* $\int_a^b f = \int_a^c f + \int_c^b f$.



Chasles: *If* $a \leqslant c \leqslant b$, *then* $\int_a^b f = \int_a^c f + \int_c^b f$.

We record a few very simple properties our integral posses, immediately following our conditions.

Proposition 9.5. *Let* $a \in \mathbb{R}$ *and* f *be a function defined at* a. *Then:*

$$\int_a^a f = 0.$$

Proof. When restricted to $\{a\} = [a, a]$, the function f is constant (hence, continuous) with value $f(a)$, so by Condition (9.2), we have:

$$\int_a^a f = f(a)(a - a) = 0,$$

as claimed. □

Our conditions immediately imply very crucial bounds on the integral of any continuous function over a segment.

Proposition 9.6. *Let* $a \leqslant b \in \mathbb{R}$ *and let* f *be a real-valued continuous function on* $[a, b]$. *Let* $m, M \in \mathbb{R}$ *such that for all* $x \in [a, b]$ *we have* $m \leqslant f(x) \leqslant M$. *Then:*

$$m(b - a) \leqslant \int_a^b f \leqslant M(b - a).$$

Remark 9.7. Since f is a continuous function on the compact set $[a, b]$, the function f is bounded (and reaches its b4ounds) by the extreme value theorem, so there always exists m, M such that $m \leqslant f \leqslant M$ on $[a, b]$.

Proof. By Condition (9.4), since $f \leqslant M$ on $[a,b]$, we conclude:

$$\int_a^b f \leqslant \int_a^b M.$$

Now by Condition (9.2), we compute that $\int_a^b M = M(b-a)$, hence $\int_a^b f \leqslant M(b-a)$, as claimed.

The proof of the other inequality is similar. □

We can extend the definition of the integral $\int$ in a manner which preserves the Chasles relation.

Definition 9.8. For all $a \leqslant b \in \mathbb{R}$, and for all continuous functions f on $[a,b]$, we set

$$\int_b^a f = -\int_a^b f.$$

Proposition 9.9. *Let $a,b,c \in \mathbb{R}$ and let f be a continuous function on some interval containing a,b,c. Then:*

$$\int_a^b f + \int_b^c f + \int_c^a f = 0.$$

Proof. Assume $a \leqslant b \leqslant c$. By Condition (9.3), we then have $\int_a^c f = \int_a^b f + \int_b^c f$, so

$$\int_a^b f + \int_b^c f + \int_c^a f = \int_a^c f + \int_c^a f = 0.$$

Using commutativity and associativity of the addition for real numbers, this proof also includes the cases $b \leqslant c \leqslant a$ and $c \leqslant a \leqslant b$.

Assume now $b \leqslant a \leqslant c$. Then:

$$\begin{aligned}\int_a^b f + \int_b^c f + \int_c^a f &= -\int_b^a f + \int_b^c f - \int_a^c f \\ &= \int_b^c -(\int_b^a f + \int_a^c f) \\ &= \int_b^c - \int_b^c f = 0,\end{aligned}$$

as expected. Again, this proof also deals with the cases $a \leqslant c \leqslant b$ and $c \leqslant b \leqslant a$. This concludes our proof. □

A first observation is that we can define Lipschitz functions by integrating continuous functions.

Proposition 9.10. *If $f : [a,b] \rightarrow \mathbb{R}$ is a continuous function, and if $F : x \in [a,b] \rightarrow \int_a^x f$, then the function F is $\|f\|_\infty$-Lipschitz on $[a,b]$ — thus, in particular, it is continuous on $[a,b]$.*

Proof. Since f is continuous over the compact interval $[a,b]$, it is bounded; we recall that $\|f\|_\infty = \sup\{|f(x)| : x \in [a,b]\}$. Therefore, for all $x \in [a,b]$, we have $-\|f\|_\infty \leqslant f(x) \leqslant \|f\|_\infty$. By Proposition (9.6), we thus conclude that, for all $x,y \in [a,b]$, if $x \leqslant y$, then

$$-\|f\|_\infty (y-x) \leqslant F(x) - F(y) = \int_x^y f \leqslant \|f\|_\infty (y-x),$$

so $|F(x)-F(y)| \leqslant \|f\|_\infty |x-y|$. If $x, y \in [a,b]$ and $y \leqslant x$, then since $|F(x)-F(y)| = |F(y)-F(x)|$, we also conclude that $|F(x)-F(y)| \leqslant \|f\|_\infty |x-y|$, as needed. Our proof is complete. $\square$

The fundamental idea for this section is to relate the integral and the derivation. We thus introduce a central definition.

Definition 9.11. Let $f : [a,b] \to \mathbb{R}$ be a function. An *antiderivative* $F : [a,b] \to \mathbb{R}$ of f on $[a,b]$ is a continuous function over $[a,b]$, differentiable over (a,b), and such that

$$\forall x \in (a,b) \quad F'(x) = f(x).$$

The reason for the specific properties listed in Definition (9.11) is explained by the following observation: antiderivatives of a given function over a fixed interval differ by a constant.

Proposition 9.12. *If $f : [a,b] \to \mathbb{R}$ is a function, if $F : [a,b] \to \mathbb{R}$ is an antiderivative of f on $[a,b]$, then any function $G : [a,b] \to \mathbb{R}$ is an antiderivative of f on $[a,b]$ if, and only if, there exists $c \in \mathbb{R}$ such that $F = G + c$, i.e.*

$$\forall x \in [a,b] \quad F(x) = G(x) + c.$$

We can now relate our integral on continuous functions and antiderivatives.

Theorem 9.13. *Let $a \leqslant b \in \mathbb{R}$. If $f : [a,b] \to \mathbb{R}$ be a continuous function on $[a,b]$, and if we define:*

$$F : x \in [a,b] \mapsto \int_a^x f,$$

then F is an antiderivative of f on $[a,b]$.

Proof. Proposition (9.10) immediately implies that F is Lipschitz, hence continuous, on $[a,b]$.

Let $x \in (a,b)$ and let $h > 0$ such that $x + h \in (a,b)$. We then compute:

$$\begin{aligned} \frac{1}{h}\left(F(x+h) - F(x)\right) &= \frac{1}{h}\left(\int_a^{x+h} f - \int_a^x f\right) \\ &= \frac{1}{h}\left(\int_x^{x+h} f\right) \text{ by Proposition (9.9).} \end{aligned} \tag{9.1}$$

Now, as f is continuous on $[a,b]$, it is continuous on $[x, x+h]$, so by the Extreme Value Theorem, there exists $m_h, M_h \in [x, x+h]$ such that for all $t \in [x, x+h]$, we have:

$$f(m_h) \leqslant f(t) \leqslant f(M_h).$$

Thus, by Proposition (9.6), we have:

$$h f(m_h) \leqslant \int_x^{x+h} f \leqslant h f(M_h). \tag{9.2}$$

Together with Expression (9.1) we obtain:

$$f(m_h) \leqslant \frac{1}{h}\left(F(x+h) - F(x)\right) \leqslant f(M_h).$$

Now, since $x \leqslant m_h \leqslant x + h$ and $\lim_{\substack{h\to 0\\h>0}} x = \lim_{\substack{h\to 0\\h>0}} x + h = x$, we conclude by the Squeeze Theorem that $\lim_{\substack{h\to 0\\h>0}} m_h = x$. Since f is continuous at x, we then conclude that $\lim_{\substack{h\to 0\\h>0}} f(m_h) = f(x)$.

Similarly, we prove that $\lim_{\substack{h\to 0\\h>0}} f(M_h) = f(x)$. Hence, Expression (9.2) and the Squeeze theorem imply together that:

$$\lim_{\substack{h\to 0\\h>0}} \frac{1}{h}\left(F(x+h) - F(x)\right) = f(x). \tag{9.3}$$

The same method applies to show that $\lim_{\substack{h\to 0\\h<0}} \frac{1}{h}\left(F(x+h) - F(x)\right) = f(x)$, so we only sketch this part.

Let $h < 0$ be given so that $x + h \in (a, b)$. Then:

$$\begin{aligned}\frac{1}{h}\left(F(x+h) - F(x)\right) &= \frac{1}{h}\int_x^{x+h} f\\ &= \frac{-1}{h}\int_{x+h}^{x} f.\end{aligned}$$

By the extreme value theorem, there exist $m_h, M_h \in [x + h, x]$ such that $f(m_h) \leqslant f(t) \leqslant f(M_h)$ for all $t \in [x + h, x]$. Thus by Condition (9.4), and noting that $x + h < x$, we have:

$$-hf(m_h) \leqslant \int_{x+h}^{x} f \leqslant -hf(M_h).$$

Thus:

$$f(m_h) \leqslant \frac{1}{h}\left(F(x+h) - F(x)\right) \leqslant f(M_h)$$

so we conclude,as before, by the Squeeze theorem, that

$$\lim_{\substack{x\to 0\\h<0}} \frac{1}{h}\left(F(x+h) - F(x)\right) = f(x). \tag{9.4}$$

Therefore, by Expressions (9.1) and (9.2), the function F is differentiable at x, and $F'(x) = f(x)$, as stated. Thus, F is continuous on $[a, b]$, differentiable on (a, b), and $F' = f$ on (a, b), so our theorem is proven. □

Remark 9.14. Theorem (9.13) proves that, if an operation satisfying Conditions (9.2), (9.3) and (9.4) exists, then any continuous function $f : [a, b] \to \mathbb{R}$ must have at least one, and thus infinitely many, antiderivatives.

We thus can compute integrals using antiderivatives, as follows.

Corollary 9.15. *Let $a \leqslant b \in \mathbb{R}$ and f be a continuous function on $[a, b]$. Let F be any antiderivative of f on $[a, b]$. Then:*

$$\int_a^b f = F(b) - F(a).$$

Proof. By Theorem (9.13), $F : x \in [a, b] \mapsto \int_a^x f$ is an antiderivative of f on $[a, b]$. Hence, by Proposition (9.12), there exists $c \in \mathbb{R}$ such that $F + c = F$. We thus get

$$F(b) - F(a) = F(b) + c - (F(a) + c) = F(b) - F(a) = F(b) = \int_a^b f,$$

as desired. □

We also record the following simple observation.

Corollary 9.16. *Let $a \leqslant b \in \mathbb{R}$ and let f be a function with a continuous derivative on $[a,b]$. Then:*

$$\int_a^b f' = f(b) - f(a).$$

Proof. The function f is by construction an antiderivative of f' on $[a,b]$. We then apply Corollary (9.15). □

Since antiderivatives are defined independently of our notion of integral, we also record that there is only one operation satisfying Conditions (9.2), (9.3) and (9.4).

Corollary 9.17. *There exists a unique real-valued operation $\int$ defined on the set of pairs $([a,b],f)$ with $[a,b]$ a closed interval of $\mathbb{R}$ and f a continuous function on $[a,b]$, and which satisfy Conditions (9.3,9.4, 9.2).*

Proof. The regulated integral possesses all the needed properties, and we have shown it is the only possible such operation. □

We can now proceed to prove some useful properties of $\int$ under our assumptions. We shall see that indeed, linearity follows from our hypothesis. Since integration is now related to the problem of antiderivation, we can use our knowledge of derivation to deduce mirror properties for the integral. The first property is linearity, and it is a very important property of integration, as we shall see later on.

Theorem 9.18 (linearity)**.** *Let $a \leqslant b \in \mathbb{R}$. If $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ are two continuous functions on $[a,b]$, and if $\mu, \lambda \in \mathbb{R}$, then:*

$$\int_a^b (\lambda f + \mu g) = \lambda \int_a^b f + \mu \int_a^b g.$$

Proof. Let $F : x \in [a,b] \to \int_a^x f$ and $G : x \in [a,b] \to \int_a^x g$. By Theorem (9.13), F and G are, respectively, antiderivatives of f and g over $[a,b]$.

The function $H = \lambda F + \mu G$ is continuous on $[a,b]$, differentiable on (a,b) and $H' = \lambda F' + \mu G' = \lambda f + \mu g$. Hence H is an antiderivative of $\lambda f + \mu g$ on $[a,b]$. Hence, by Corollary (9.15):

$$\begin{aligned}
\int_a^b (\lambda f + \mu g) &= H(b) - H(a) \\
&= \lambda F(b) + \mu G(b) - (\lambda F(a) + \mu G(a)) \\
&= \lambda(F(b) - F(a)) + \mu(G(b) - G(a)) \\
&= \lambda \int_a^b f + \mu \int_a^b g.
\end{aligned}$$

This concludes our proof. □

Of course, $\int_a^b 0 = 0$ — this was in fact assumed in Condition (9.2), but it also follows from the linearity of the integral. The converse of this result is not true, but we can obtain an interesting and important result nonetheless.

Theorem 9.19. *If $f : [a,b] \to \mathbb{R}$ is continuous, if $f(x) \geqslant 0$ for all $x \in [a,b]$, and if $\int_a^b f = 0$, then $f = 0$.*

Proof. Let $F : x \in [a,b] \mapsto \int_a^x f$. By Theorem (9.13), the function F is an antiderivative of f on $[a,b]$, so $F' = f$ on (a,b). By assumption, $f \geqslant 0$, so F is increasing on $[a,b]$, by the Mean Value Theorem: hence, for all $x \in [a,b]$, we conclude that $F(a) \leqslant F(x) \leqslant F(b)$. Now, $F(b) = \int_a^b f = 0 = F(a)$, and thus, for all $x \in [a,b]$, we conclude that $F(x) = 0$. Thus $f(x) = F'(x) = 0$ for all $x \in (a,b)$. As f is continuous on $[a,b]$, we conclude that $f = 0$ on $[a,b]$, as claimed. □

Thus, linearity of derivation leads to linearity of the integral. The matter is more complicated for products and compositions. We start by handling the mirror image of the Leibniz rule for derivations of products.

Theorem 9.20 (Integration by parts)**.** *Let $a \leqslant b \in \mathbb{R}$. If $f : [a,b] \to \mathbb{R}$ and $g : [a,b] \to \mathbb{R}$ are two functions with continuous derivatives on $[a,b]$, then:*

$$\int_a^b fg' = f(b)g(b) - f(a)g(a) - \int_a^b f'g.$$

Proof. We have $(fg)' = f'g + fg'$ so $fg' = (fg)' - f'g$ and thus we have:

$$\begin{aligned}\int_a^b fg' &= \int_a^b [(fg)' - f'g] \\ &= \int_a^b (fg)' - \int_a^b f'g \text{ by Theorem (9.18)} \\ &= f(b)g(b) - f(a)g(a) - \int_a^b f'g \text{ by Theorem (9.16).}\end{aligned}$$

This concludes our proof. □

The chain rule becomes the substitution rule:

Theorem 9.21 (Substitution)**.** *Let $a \leqslant b$. If u is a function with a continuous derivative on $[a,b]$ and if f is a continuous function on $[u(a),u(b)] \cup [u(b),u(a)]$, then:*

$$\int_a^b ((f \circ u) \cdot u') = \int_{u(a)}^{u(b)} f.$$

Proof. Let $F : x \in [a,b] \to \int_a^x f$. Let $g = F \circ u$ on $[a,b]$. By assumption, g is differentiable on $[a,b]$ and:

$$g'(x) = u'(x)F'(u(x)) = u'(x)f(u(x))$$

for all $x \in (a,b)$. Thus:

$$\begin{aligned}\int_a^b f \circ u \cdot u' &= \int_a^b g' \\ &= g(b) - g(a) \\ &= F(u(b)) - F(u(a)) \\ &= \int_{u(a)}^{u(b)} f.\end{aligned}$$

Our proof is now complete. □

We conclude this section with two observations. First, we prove a useful inequality:

Theorem 9.22 (Median inequality)**.** *Let $a \leqslant b \in \mathbb{R}$. If f is a continuous function on $[a,b]$, then:*

$$\left|\int_a^b f\right| \leqslant \int_a^b |f|.$$

Proof. By definition of the absolute value, $|f|$ is continuous on $[a,b]$ and for all $x \in [a,b]$ we have:

$$-|f(x)| \leqslant f(x) \leqslant |f(x)|.$$

By Condition (9.4) and by Theorem (9.18), we have:

$$-\int_a^b |f| = \int_a^b (-|f|) \leqslant \int_a^b f \leqslant \int_a^b |f|,$$

proving our inequality. □

Another observation motivates the idea that integrals allow to compute averages, as found in probability theory.

Theorem 9.23. *Let $a \leqslant b \in \mathbb{R}$. If $f : [a,b] \to \mathbb{R}$ is a continuous function on $[a,b]$, then there exists $c \in (a,b)$ such that:*

$$\frac{1}{b-a}\int_a^b f = f(c).$$

Proof. The function $F : x \in [a,b] \to \int_a^x f$ is continuous on $[a,b]$ and differentiable on (a,b), by Theorem (9.13). So, by the mean value theorem, there exists $c \in (a,b)$ such that:

$$F(b) - F(a) = (b-a)F'(c) = f(c)(b-a).$$

This proves our theorem. □

The quantity $\frac{1}{b-a}\int_a^b f$ is called the *average of f*. In particular, the strong law of large number shows that $\frac{1}{b-a}\int_a^b f$ is the limit of the empirical average for any infinite sample of points taking uniformly at random in $[a,b]$.

Email address: `frederic@math.du.edu`
URL: `http://www.math.du.edu/~frederic`

Department of Mathematics, University of Denver, Denver CO 80208