

Advanced-Calc Notes

Jonathan L.F. King
 University of Florida, Gainesville FL 32611-2082, USA
 squash@ufl.edu
 Webpage <http://squash.1gainesville.com/>
 1 February, 2022 (at 22:02)

Spaces. Various spaces will be used/defined in this pamphlet. Abbrevs: *VS*, *vectorspace*. *NVS*, *normed vector-space*. *IPVS*, *inner-product (vector)space*. *TOS*, *totally-ordered space*. *MS*, *metric space*. *CMS*, *complete MS*. *TS*, *topological space*. *HS*, *Hausdorff (topological) space*.

§Outline

| | |
|--|----|
| Elementary MS/TOS theorems | 2 |
| MS-sequence Thm | 2 |
| Monotone-subsequence Thm | 2 |
| Induced-topology Lemma | 3 |
| LUBP theorem | 3 |
| Important announcement | 3 |
| Monotone-sequence Thm | 3 |
| Bounded-sequence Lemma | 3 |
| \mathbb{R} Thm | 3 |
| Intermediate-value Theorem [IVT] | 3 |
| Stage- <i>n</i> construction | 4 |
| Last step | 4 |
| Addition-Cts thm | 4 |
| Multi-Cts thm | 4 |
| Normed VSes and MSes | 4 |
| Metric Spaces | 5 |
| Examples of metrics | 5 |
| Examples of normed-VSes | 5 |
| ℓ^p spaces are Banach spaces | 6 |
| Topological Spaces | 6 |
| Classification of properties | 6 |
| Closure/Interior/Bdry etc. | 6 |
| Isomorphisms | 7 |
| Defin: Relative topology | 7 |
| Locally Countably Generated spaces | 7 |
| Sequence-Closure Lemma | 7 |
| Compactness | 8 |
| Covers | 8 |
| Diameter/compactness Prop'n | 8 |
| Compact-intervals theorem | 8 |
| Metric ideas related to compactness | 9 |
| TB-iff-CauchySubseq Thm | 9 |
| Compactness notions Theorem | 9 |
| Product-space Convergence Lemma | 10 |
| Heine-Borel theorem | 10 |
| Bufferable pairs of sets | 11 |
| Compactness lemma | 11 |
| Continuity | 12 |
| Baby continuity Lemma | 12 |
| Uniform-continuity Theorem | 12 |
| Lip-Diff Lemma | 13 |
| Piecewise-linear functions | 14 |
| P.L-approximation thm | 14 |
| Uniform Convergence | 15 |
| Uniform-convergence theorem | 15 |
| Weird application of unif-convergence | 15 |
| Nested uniform-convergence thm | 16 |
| Miscellaneous continuity/limit results | 16 |
| Composition notation | 16 |
| Forward-inheritance Lemma | 17 |
| Miscellaneous connectedness results | 17 |

| | |
|--|-----------|
| A Differentiability | 18 |
| How discontinuous can a derivative be? | 18 |
| Deriv-cty Lemma | 18 |
| vdW's no-where differentiable fnc | 19 |
| vdW-function thm | 19 |
| Product-rule thm | 20 |
| Total derivative | 20 |
| B Riemann Integration | 21 |
| Integrability-equivalence Lemma | 23 |
| Basic RI Thm | 23 |
| Closure properties of \mathbb{R} | 23 |
| Integration-is-Linear lemma | 23 |
| AbsValue RI Thm | 24 |
| Product-RI Thm | 24 |
| Closure-RI Thm | 25 |
| The Fundamental Theorem of Calculus | 25 |
| Fundamental Theorem of Calculus [FTC] | 25 |
| Measuring the size of sets | 26 |
| Basic measure lemma | 26 |
| A condition for Integrability | 27 |
| Integrability Theorem | 27 |
| Interchange of limit-operations | 28 |
| Integral-Convergence Theorem | 28 |
| DUC Thm | 28 |
| Series and Sequences | 29 |

Appendix: *Menagerie of Strange Functions*

| | |
|--|----|
| Strictly increasing fnc, with Cty and DisCty dense | 30 |
|--|----|

| | |
|--------------------------------------|-----------|
| Index for ADVANCED-CALC NOTES | 31 |
|--------------------------------------|-----------|

Prelim: VSes. To indicate that \mathbf{u} is a vector in a VS \mathbf{W} , I'll normally write " $\mathbf{u} \in \mathbf{W}$ ", both in notes and on the blackboard; but I can't write boldface on the blackboard, so it will be " $u \in W$ ". In notes, I'll use boldface

$$\mathbf{0} \stackrel{\text{or}}{=} \widehat{\mathbf{0}}, \widehat{i}, \widehat{j}, \widehat{k}$$

for the **zero-vector** and for the three coordinate-vectors in \mathbb{R}^3 . On the blackboard, I'll write these as $\widehat{\mathbf{0}}, \widehat{i}, \widehat{j}, \widehat{k}$. In contrast, I'll use an overarrow –see (3a), below– to indicate *sequences*. (And indeed, these seqs will often *be* vectors in \mathbb{R}^∞ .)

Over a field \mathcal{F} , consider \mathcal{F} -VSes \mathbf{V} and \mathbf{E} . A map $L: \mathbf{V} \rightarrow \mathbf{E}$ is \mathcal{F} -**linear** (or just **linear**) if:

$$1: \quad \forall \alpha, \beta \in \mathcal{F} \quad \text{and} \quad \forall \mathbf{v}, \mathbf{w} \in \mathbf{V}, \quad \text{necessarily} \\ L(\alpha \mathbf{v} + \beta \mathbf{w}) = \alpha L(\mathbf{v}) + \beta L(\mathbf{w}).$$

A map $L: \mathbf{V} \rightarrow \mathcal{F}$ is called a **functional** (abbrev.: *fnc'al*). In the typical case, $L()$ is linear (viewing \mathcal{F} as a 1-dim'al VS over \mathcal{F}) and we call $L()$ a **linear functional**.

Prelim: Sets. For arbitrary sets D and C , I'll sometimes use

2: the symbol C^D to denote the set of functions $D \rightarrow C$.

(This is a std notation.) The “exponent” D is the domain of these fncs, and C is their codomain. As an example, the vectorspace \mathbb{R}^3 can be viewed as the set of fncs $\mathbb{R}^{[1..3]}$, or as $\mathbb{R}^{[0..3]}$, if convenient.

BTWay: When D and C are finite sets,

2': The cardinality $|C^D|$ equals $|C|^{|D|}$.

Elementary MS/TOS theorems

In this section, we have a general totally-ordered space $(Y, <)$. We also have a general metric space (Ω, d) .

Notation for sequences. A symbol \vec{x} means the (by default, infinite) ordered tuple

3a: $\vec{x} = (x_1, x_2, x_3, \dots)$;

however, the index-set might be a different “ray” of integers, e.g. \vec{x} might be denoting (x_3, x_4, x_5, \dots) . Since \vec{x} is a fnc, $\text{Dom}(\vec{x})$ denotes its index-set, and $\text{Range}(\vec{x}) = \{x_n\}_{n \in \text{Dom}(\vec{x})}$ is its set of \vec{x} -values. Most of the notation below assumes the index-set is \mathbb{Z}_+ .

For a set S , expression “ $\vec{x} \subset S$ ” means

$$\forall n \in \text{Dom}(\vec{x}): x_n \in S.$$

A “**list** of indices” shall mean posints

3b: $N_1 < N_2 < N_3 < \dots$

A sequence \vec{c} is a **subsequence** of \vec{x} IFF *there exists* a list (3b) st. $\forall k: c_k = x_{N_k}$. Write “ $\vec{c} \subset \vec{x}$ ” to indicate this relation. Each $N \in \text{Dom}(\vec{x})$ yields a subsequence called “the N^{th} **tail** of \vec{x} ”,

$$\text{Tail}_N(\vec{x}) := (x_N, x_{N+1}, x_{N+2}, \dots).$$

Fix MS (Ω, d) . For $\vec{x} \subset \Omega \ni q$, let “ $d\text{-lim}(\vec{x}) = q$ ” or “ $\Omega\text{-lim}(\vec{x}) = q$ ” or just “ $\text{lim}(\vec{x}) = q$ ” mean

3c: For each ball $B := \text{Bal}(q)$, there exists an index $N = N(B)$ for which $\text{Tail}_N(\vec{x}) \subset B$.

Implicit in our notation is “*Limits, when they exist, are unique*”. Were this not the case, then we'd view $\text{lim}(\vec{x})$

as a *set*, and write “ $q \in \text{lim}(\vec{x})$ ” rather than $q = \text{lim}(\vec{x})$. Uniqueness is proved after (20), P.6.

We will interpret a *sequence* \vec{e} as the *set* $\text{Range}(\vec{e})$ in these two common contexts: “ $\text{Diam}(\vec{e})$ ” and “ $\vec{e} \subset S$ ”. For example, a sequence \vec{x} is **d-Cauchy** if:

3d: $\forall \varepsilon > 0, \exists N$ such that $d\text{-Diam}(\text{Tail}_N(\vec{x})) < \varepsilon$. □

4: MS-sequence Thm. *Facts about seqs in MS (Ω, d) :*

A: If \vec{x} is convergent, then \vec{x} is a Cauchy sequence.

B: If \vec{x} is Cauchy, then $\text{Diam}(\vec{x}) < \infty$.

C: Suppose Cauchy-seq \vec{x} has a convergent subseq $\vec{y} \subset \vec{x}$. Then \vec{x} converges, and $\text{lim}(\vec{x}) = \text{lim}(\vec{y})$. ◇

Proof of (C). The first two parts were proved in class. For the third, let $p := \text{lim}(\vec{y})$. Fix $\varepsilon > 0$, then take N large enough that $\text{Diam}(\text{Tail}_N(\vec{x})) < \varepsilon$.

Write \vec{y} as $(x_{K_j})_{j=1}^\infty$. Let J be the first posint large enough that $\boxed{K := K_J \geq N}$ and $d(x_K, p) < 7\varepsilon$.

For each $\ell \in [K .. \infty)$, observe that

$$\begin{aligned} d(x_\ell, p) &\leq d(x_\ell, x_K) + d(x_K, p) \\ &< \varepsilon + 7\varepsilon = 8\varepsilon. \end{aligned}$$

Thus $\text{Tail}_K(\vec{x}) \subset \text{Bal}_{8\varepsilon}(p)$. ◇

5: Monotone-subsequence Thm. *Each seq $\vec{x} \subset Y$ has a monotone subsequence.* (“Sequence” means ∞ -seq.)

Indeed, either \vec{x} has a strictly decreasing subseq, or has an increasing subsequence. (Dually, \vec{x} has a strictly incr-subseq or a decr-subseq.) ◇

Proof. Let $\mathcal{T} \subset \mathbb{Z}_+$ comprise the “tall” indices N for which: $[\forall k \in (N .. \infty): x_N > x_k]$.

If \mathcal{T} is infinite, then $(x_\tau)_{\tau \in \mathcal{T}}$ is a strictly-decreasing subsequence of \vec{x} .

Now suppose \mathcal{T} finite. Let N_1 be the smallest index exceeding all the tall indices (phrased this way, to cover the case where \mathcal{T} is empty). Arguing inductively, suppose we have indices $N_1 < N_2 < \dots < N_{K-1}$ for which

$$x_{N_1} \leq x_{N_2} \leq \dots \leq x_{N_{K-1}}.$$

Since N_{K-1} is not tall, there exists a smallest integer $N_K > N_{K-1}$ for which x_{N_K} dominates $x_{N_{K-1}}$.

Continuing the induction yields $(x_{N_k})_{k=1}^\infty$, an increasing subsequence of \vec{x} . ◇

6: Induced-topology Lemma. Fix a MS Ω and subset X . Then a further subset $U \subset X$ is X -open IFF there exists an Ω -open set \widehat{U} st. $\widehat{U} \cap X = U$. *Proof.* Exercise. \diamond

Least upper-bound property [LUBP]. In TOS $(Y, <)$, consider sets $A, B \subset Y$ and a point $u \in Y$. Let

$$7: \quad \begin{aligned} A \leq u & \text{ mean } [\forall \alpha \in A, \text{ necessarily } \alpha \leq u]; \\ A \leq B & \text{ mean } [\forall \alpha \in A \text{ and } \forall \beta \in B: \alpha \leq \beta]. \end{aligned}$$

An **upper-bound** for a set $A \subset Y$ is an element $u \in Y$ such that $A \leq u$. Use $\text{UB}_Y(A)$ for the set of upper-bnds, and $\text{LB}_Y(A)$ for the lower-bnd-set. (Dispense with the subscript if clear from context.) Our $(Y, <)$ has the **LUBP** if:

Each non-void $A \subset Y$ which is upper-bnded [i.e. $\text{UB}_Y(A) \neq \emptyset$] has a **least upper-bound**. That is, $\text{UB}_Y(A)$ has a minimum element.

Reversing the inequalities yields the **greatest lower-bound property**, abbreviated **GLBP**.

The LUB of a set A (when it *has* a LUB!) is called the **supremum** of the set, and is written $\text{sup}(A)$ or $\text{sup}_Y(A)$. Similarly, the **infimum** is the GLB, written $\text{inf}(A)$.

7b: LUBP theorem. TOS $(Y, <)$ has the LUBP IFF it has the GLBP. \diamond

Proof of [LUBP \Rightarrow GLBP]. Fix a non-void lower-bnded subset $B \subset Y$; so $A := \text{LB}_Y(B)$ is non-empty. My goal is to produce a (hence the) greatest lower-bound for B , using that

$$\begin{aligned} \dagger: \quad & A \stackrel{\text{def}}{=} \text{LB}_Y(B), \quad \text{and} \\ \ddagger: \quad & \text{UB}_Y(A) \supset B. \end{aligned}$$

Since $\text{UB}_Y(A) \supset B \neq \emptyset$, and A is non-void, the LUBP applies, and tells us that $\lambda := \text{sup}_Y(A)$ exists. In particular

$$\dagger': \quad \lambda \geq A.$$

Since λ is the *least* upper-bnd, $\lambda \leq \text{UB}_Y(A) \supset B$ and so $\lambda \leq B$. Restating, λ is a lower-bound of B . (Note: λ might or might not be in B .)

And, by (\dagger) and (\dagger') , this λ dominates each lower-bound of B . So λ is a **greatest** lower-bound of B . \diamond

Important announcement. A TOS $(Y, <)$ satisfying LUBP [equivalently, GLBP] is said to be **order-complete**. We take as an axiom [or derive via Dedekind cuts or Cauchy sequences] that

$$7c: \quad (\mathbb{R}, <) \text{ is order-complete.}$$

This means that the extended reals, $\overline{\mathbb{R}}$, satisfies a slightly stronger property: Each^{†1} subset $A \subset \overline{\mathbb{R}}$ has a $\text{sup}(A)$ and an $\text{inf}(A)$ in $\overline{\mathbb{R}}$. In consequence, $\text{sup}()$ and $\text{inf}()$ are maps from the full $\mathcal{P}(\overline{\mathbb{R}}) \rightarrow \overline{\mathbb{R}}$.

(See (14), P.5, for the definition of $\overline{\mathbb{R}}$, the extended reals.)

8: Monotone-sequence Thm. Each bounded monotone sequence $\vec{x} \subset \mathbb{R}$ is \mathbb{R} -convergent. \diamond

Proof. WLOG, \vec{x} is increasing, and upper-bnded. Thus $X := \text{Range}(\vec{x})$ has a supremum in \mathbb{R} ; call it L . I claim that $\lim(\vec{x}) \stackrel{?}{=} L$.

Fix $\varepsilon > 0$. Now L is the *least* UB of X , so $L - \varepsilon$ can not be an upper-bnd. Hence there exists N with $x_N > L - \varepsilon$. For each $\ell \geq N$, since \vec{x} is increasing, we have that

$$L - \varepsilon < x_N \leq x_\ell \leq L.$$

Thus $\text{Tail}_N(\vec{x}) \subset \text{Bal}_\varepsilon(L)$. \diamond

9: Bounded-sequence Lemma. Each bounded sequence $\vec{x} \subset \mathbb{R}$ has an \mathbb{R} -convergent subsequence. \diamond

Proof. Use (5), then (8). \diamond

10: \mathbb{R} Thm. The set of reals is (metrically) complete. \diamond

Proof. Fix a Cauchy sequence $\vec{x} \subset \mathbb{R}$. Courtesy (4B), $\text{Diam}(\vec{x}) < \infty$. So (9) applies, yielding a convergent subsequence. Now use (4C). \diamond

Bernard Bolzano (1781–1848) proved the following form of the Intermediate-value Theorem.

11: IVT. Suppose $f: [a, b] \rightarrow \mathbb{R}$ is continuous, with $f(a)$ and $f(b)$ non-zero and having different signs. Then there exists a point $c \in (a, b)$ which is a zero of f , i.e. $f(c) = 0$. \diamond

^{†1}E.g. $\text{sup}(\emptyset) = -\infty$ and $\text{inf}(\emptyset) = +\infty$. Indeed, for $A \subset \overline{\mathbb{R}}$: $[A \neq \emptyset] \iff [\text{inf}(A) \leq \text{sup}(A)]$.

Proof. WLOGenerality, $f(a) < 0$ and $f(b) > 0$; otherwise, simply replace f by $-f$ (which preserves continuity) and note that a zero of $-f$ is a zero of f .

Let $L_0 := a$ and $R_0 := b$. For stage $n = 1, 2, \dots$, either up to some integer K , or out to ∞ , I will produce numbers L_n and R_n such that:

- i[n]: $a \leq L_{n-1} \leq L_n < R_n \leq R_{n-1} \leq b$;
- ii[n]: $R_n - L_n = \frac{1}{2}[R_{n-1} - L_{n-1}]$;
- iii[n]: $f(L_n) < 0 < f(R_n)$.

Stage- n construction. Let M be the midpoint of interval $[L_{n-1}, R_{n-1}]$, i.e, $M := \frac{1}{2}[L_{n-1} + R_{n-1}]$.

CASE: If $f(M)$ is zero, then STOP Set $K := n-1$. By (i[K]), note that M is strictly between a and b . So $c := M$ fulfills the conclusion of the theorem.

CASE: Otherwise, $f(M) \neq 0$. If $f(M)$ negative then let $L_n := M$ & $R_n := R_{n-1}$. If $f(M)$ positive then let $L_n := L_{n-1}$ & $R_n := M$. In either case, conditions (i,ii,iii[n]), automatically hold.

Last step. WLOGenerality, we may assume that our construction never STOPped. So we have two sequences, $\vec{L} := (L_n)_{n=0}^\infty$ and $\vec{R} := (R_n)_{n=0}^\infty$.

By (i), \vec{L} is increasing and is bounded above by b . Since a bounded monotone seq must converge, $L_\infty := \lim_{n \rightarrow \infty} L_n$ exists; it is in interval $[a, b]$, courtesy (i).

Thus f is defined –hence continuous– at L_∞ , so $f(L_\infty)$ equals $\lim_n f(L_n)$. And $f(L_\infty) \stackrel{\text{must}}{\leq} 0$ since each $f(L_n) \leq 0$.

Analogously, $f(R_\infty) := \lim_{n \rightarrow \infty} f(R_n)$ exists, and is non-negative. Furthermore

$$\begin{aligned} R_\infty - L_\infty &= \lim_{n \rightarrow \infty} [R_n - L_n], \quad \text{by what thm?}, \\ &= \lim_{n \rightarrow \infty} \left[\frac{1}{2}\right]^n \cdot [b - a], \quad \text{by (iii) and induction,} \\ &= 0. \end{aligned}$$

Thus R_∞ and L_∞ equal a common value, call it c , in interval $[a, b]$. The preceding paragraphs tell us that $f(c) \leq 0$ and $f(c) \geq 0$; so $f(c)$ must be zero. Hence $c \notin [a, b]$. \blacklozenge

12: Addition-Cts thm. *The addition operation $\mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ is continuous. Restated: Suppose $\vec{x}, \vec{y} \subset \mathbb{C}$ with $\lim(\vec{x}) = \alpha$ and $\lim(\vec{y}) = \beta$. With $p_n := x_n + y_n$, then, $\lim(\vec{p}) = \alpha + \beta$.* \blacklozenge

Proof. Fix a posreal ε . Take N large enough that

$$\text{Tail}_N(\vec{x}) \subset \text{Bal}_{\frac{\varepsilon}{2}}(\alpha) \quad \text{and} \quad \text{Tail}_N(\vec{y}) \subset \text{Bal}_{\frac{\varepsilon}{2}}(\beta).$$

Each index k has $p_k - [\alpha + \beta] = [x_k - \alpha] + [y_k - \beta]$. For each $k \geq N$, then,

$$|p_k - [\alpha + \beta]| \leq |x_k - \alpha| + |y_k - \beta| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \quad \blacklozenge$$

Remark. The same thm and proof hold for addition on a normed vectorspace; simply replace $|\cdot|$ by the norm $\|\cdot\|$. \square

13: Mult-Cts thm. *The multiplication operation $\mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ is continuous. Restated: Suppose $\vec{x}, \vec{y} \subset \mathbb{C}$ with $\lim(\vec{x}) = \alpha$ and $\lim(\vec{y}) = \beta$. With $p_n := x_n \cdot y_n$, then, $\lim(\vec{p}) = \alpha \cdot \beta$.* \blacklozenge

Proof. WELOG $|\beta| \leq 7$. Since \vec{x} converges, necessarily the Diam(\vec{x}) is finite; WELOG

$$\dagger: \quad \forall \text{ posints } n: \quad |x_n| \leq 50.$$

$$\begin{aligned} \text{For each posint } n, \text{ adding and subtracting a term gives} \\ x_n y_n - \alpha \beta &= x_n y_n - x_n \beta + x_n \beta - \alpha \beta \\ &= x_n [y_n - \beta] + [x_n - \alpha] \beta. \end{aligned}$$

Taking absolute-values, then upper-bounding, yields

$$\begin{aligned} \ddagger: \quad |x_n y_n - \alpha \beta| &\leq |x_n| \cdot |y_n - \beta| + |x_n - \alpha| \cdot |\beta| \\ &\leq 50 \cdot |y_n - \beta| + |x_n - \alpha| \cdot 7, \end{aligned}$$

by (\dagger) and the first sentence.

Fix a posreal ε . Since $\lim(\vec{y}) = \beta$ and $\lim(\vec{x}) = \alpha$, we can take K large enough that for each n in $[K .. \infty)$:

$$|y_n - \beta| \leq \frac{\varepsilon/2}{50} \quad \text{and} \quad |x_n - \alpha| \leq \frac{\varepsilon/2}{7}.$$

Plugging these estimates in to (\ddagger) gives that

$$|x_n y_n - \alpha \beta| \leq 50 \cdot \frac{\varepsilon/2}{50} + \frac{\varepsilon/2}{7} \cdot 7 \stackrel{\text{note}}{=} \varepsilon,$$

for each $n \geq K$.

As this holds for every ε positive, $\lim(\vec{x} \cdot \vec{y})$ indeed equals $\alpha \beta$. \blacklozenge

Normed VSes and MSes

A **norm** $\|\cdot\|$, on a real or complex vectorspace \mathbf{W} , is a map $\mathbf{W} \rightarrow [0, \infty)$ such that $\forall \mathbf{u}, \mathbf{v} \in \mathbf{W}$:

- N1: $\|\mathbf{u}\| = 0$ IFF $\mathbf{u} = \mathbf{0}$.
- N2: \forall scalars α : $\|\alpha \mathbf{u}\| = |\alpha| \cdot \|\mathbf{u}\|$.
- N3: $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Metric Spaces. On a set X , a *metric* m is a map $X \times X \rightarrow [0, \infty)$ such that $\forall x, y, z \in X$:

MS1: $m(x, y) = 0$ IFF $x = y$.

MS2: $m(x, y) = m(y, x)$.

MS3: $m(x, z) \leq m(x, y) + m(y, z)$.

Evidently, a norm $\|\cdot\|$ defines a metric m , by

$$\forall \mathbf{u}, \mathbf{v} \in \mathbf{W}: m(\mathbf{u}, \mathbf{v}) := \|\mathbf{u} - \mathbf{v}\|.$$

Equivalent metrics. Use $\text{OPN}(m)$ for the *collection* of open sets that metric m determines; so $\text{OPN}(m) \subset \mathcal{P}(X)$. Say that two metrics m and d , on the same space, are **topologically equivalent** (*topo-equiv*) if $\text{OPN}(m) = \text{OPN}(d)$. We write $m \overset{\text{Topo}}{\asymp} d$.

If $m \overset{\text{Topo}}{\asymp} d$ and m and d have exactly the same Cauchy seqs, then they are **Cauchy equivalent**, written $m \overset{\text{Cau}}{\asymp} d$.

Examples of metrics. Let's first look at one-dimensional examples.

E1. Let \mathbb{S} be the unit circle $\{(x, y) \mid x^2 + y^2 = 1^2\}$. It has an **arclength-metric** \mathbf{d}_{Arc} , and a **chordal metric** \mathbf{d}_{Ch} . E.g.,

$$\begin{aligned} \mathbf{d}_{\text{Arc}}\text{-Diam}(\mathbb{S}) &= \pi, \quad \text{and} \\ \mathbf{d}_{\text{Ch}}\text{-Diam}(\mathbb{S}) &= 2. \end{aligned}$$

Evidently, $\mathbf{d}_{\text{Arc}} \overset{\text{Cau}}{\asymp} \mathbf{d}_{\text{Ch}}$. □

E2. I define the **arctan metric**, α , on \mathbb{R} and on

$$14: \quad \overline{\mathbb{R}} \overset{\text{synon}}{=} \dot{\mathbb{R}} := \{-\infty\} \sqcup \mathbb{R} \sqcup \{+\infty\} = [-\infty, +\infty].$$

For points $x, y \in \overline{\mathbb{R}}$, define (using \mathbf{d}_{Arc})

$$\alpha(x, y) := |\arctan(x) - \arctan(y)|.$$

Note that $\arctan(+\infty) = +\frac{\pi}{2}$ and $\arctan(-\infty) = -\frac{\pi}{2}$. And α is topo-equiv to the usual metric on \mathbb{R} , but they are *not* Cauchy-equivalent.

The set (14) is variously called the **extended reals** or the **2-point compactification** of \mathbb{R} . □

E3. The **stereographic metric**, σ , on \mathbb{R} and on

$$15: \quad \dot{\mathbb{R}} := \mathbb{R} \sqcup \{\infty\},$$

comes from a projection, as did the arctan-metric. Recall the circle \mathbb{S} from (E1). Let $\overset{\circ}{\mathbb{S}}$ be the “punctured circle”, where we removed the “north pole” $\mathbf{NP} := (0, 1)$. We have two homeomorphisms, $f: \overset{\circ}{\mathbb{S}} \rightarrow \mathbb{R}$ and its inverse-fnc $g: \mathbb{R} \rightarrow \overset{\circ}{\mathbb{S}}$. They are defined by a diagram. (See blackboard.) A bit of algebra shows that

$$16: \quad \begin{aligned} f((x, y)) &= \frac{x}{1-y}; \\ g(t) &= \frac{1}{t^2+1} \cdot (2t, t^2 - 1). \end{aligned}$$

We extend these maps to $f: \mathbb{S} \rightarrow \dot{\mathbb{R}}$ and $g: \dot{\mathbb{R}} \rightarrow \mathbb{S}$, by

$$16': \quad f(\mathbf{NP}) := \infty \quad \text{and} \quad g(\infty) := \mathbf{NP}.$$

Finally, our stereographic metric is: $\forall p, q \in \dot{\mathbb{R}}$,

$$16'': \quad \sigma(p, q) := \mathbf{d}_{\text{Ch}}(g(p), g(q)).$$

The set (15) is called the **projectively extended reals** or the **1-point compactification** of \mathbb{R} . □

Examples of normed-VSes. For a posint N , let's define a family of norms on N -dimensional space $\mathbb{R} \times \dots \times \mathbb{R}$. It will be convenient to use (2), P.2, and write this VS as \mathbb{R}^J , where J is the index-set $J := [0 .. N)$.

For exponent $p \in [1, \infty)$, define the ℓ^p -norm (“little-Lp norm”) by

$$17A: \quad \begin{aligned} \|\mathbf{u}\|_p &:= \left[\sum_{k \in J} |u_k|^p \right]^{1/p}. \quad \text{Also define} \\ \|\mathbf{u}\|_\infty &:= \sup_{k \in J} |u_k|. \end{aligned}$$

One often uses $\ell^p = \ell^p(J)$ as the name of the VS; here, since J is finite, the VS is \mathbb{R}^J . A bit of argument shows

$$17B: \quad \forall \mathbf{u} \in \mathbb{R}^J: \lim_{p \nearrow \infty} \|\mathbf{u}\|_p = \|\mathbf{u}\|_\infty.$$

Infinite index-sets. Now let $J := \mathbb{N}$, the set of real-valued sequences. *What should our vectorspace $\ell^p(J)$ be?*

Take the case $p := 1$. As an example, the constant-7 sequence $\vec{7}$ has infinite ℓ^1 -“norm”; so we *don't* want $\vec{7}$ in ℓ^1 . So for each $p \in [1, \infty)$ we define, using (17A),

$$17C: \quad \ell^p(J) := \left\{ \mathbf{v} \in \mathbb{R}^J \mid \|\mathbf{v}\|_p \text{ is finite} \right\}.$$

One can check that this set *is* sealed under vector-addition, so it is a vector subspace of \mathbb{R}^J . □

^{∞2}For each $p \in [1, \infty)$, indeed, $\|\vec{7}\|_p = +\infty$. OTOHand, $\|\vec{7}\|_\infty = 7$.

Supremum-norm. On a MS X , let $\mathbf{C}(X)$ or $\mathbf{C}^0(X)$ denote the set –indeed, the *vectorspace*– of continuous functions $X \rightarrow \mathbb{R}$. For $f \in \mathbf{C}(X)$, define

$$\|f\|_{\text{sup}} := \sup_{x \in X} |f(x)|.$$

Since this can take on the value $+\infty$, we drop to the vector-subspace of *bounded* continuous fncls,

$$17D: \quad \mathbf{C}_{\text{Bnd}}(X) := \{f \in \mathbf{C}(X) \mid \|f\|_{\text{sup}} < \infty\}.$$

This pair $(\mathbf{C}_{\text{Bnd}}(X), \|\cdot\|_{\text{sup}})$ is a normed-VS. If X is *compact* then –we’ll later discover– *every* cts fnc is bounded. \square

The following thm is easy, when J is finite, but takes some work when the index-set is infinite. (A *Banach space* [don’t panic] is a complete normed-vectorspace.)

18: ℓ^p spaces are Banach spaces. Fix an indexing-set J . Then for each $p \in [1, \infty]$, the space $\ell^p(J)$ is complete in the metric induced by $\|\cdot\|_p$. \diamond

All the foregoing holds *mutatis mutandis* for \mathbb{R} replaced by \mathbb{C} , the complex numbers. Equation (17B), when stated appropriately, holds even when J is infinite.

Topological Spaces

A TS Ω has a collection $\mathcal{U} \subset \mathcal{P}(\Omega)$ of sets that we call the *open* sets. Family \mathcal{U} is required to satisfy:

TS1: \mathcal{U} owns \emptyset and owns Ω .

TS2: $\forall A, B \in \mathcal{U}$, the intersection $A \cap B \in \mathcal{U}$.

TS3: For each collection $\mathcal{A} \subset \mathcal{U}$: The union $\bigcup(\mathcal{A})$ is in \mathcal{U} . (Note that $\bigcup(\mathcal{A})$ is the set of points $\omega \in \Omega$ for which *there exists* a set $V \in \mathcal{A}$ with $V \ni \omega$.)

Let’s use $\text{OPN}(\Omega)$ for this collection \mathcal{U} , and use $\text{CLD}(\Omega)$ for the family of *closed* subsets. Topologists tend to be biased toward opens sets, and call $\text{OPN}(\Omega)$ “the *topology* of Ω ”. This TS is *metrizable* if *there exists* a metric m on Ω for which $\text{OPN}(m) = \text{OPN}(\Omega)$.

Classification of properties. A concept/property on/of a space Ω is (purely) *topological* if it can be determined solely by knowing $\text{OPN}(\Omega)$. On a MS, a property is *metric* if it can be determined from the metric. E.g $\text{Diam}(\Omega)$ is a metric property, but whether Ω is *connected* is purely a topological property.

Perhaps surprisingly, convergence of a sequence “ $\lim(\vec{x}) = q$ ” is just a topological property. For it can be stated as

19: For each open $U \ni q$, there exists an index $N = N(U)$ for which $\text{Tail}_N(\vec{x}) \subset U$.

The notation suggests that a sequence can have at most one limit, and this is true for TSes with the *Hausdorff separation property* (which trivially holds in MSes):

20: For each pair of distinct points $\alpha, \beta \in \Omega$, there exist disjoint open sets $A \ni \alpha$ and $B \ni \beta$.

For if seq \vec{x} converges to both α and β , then $\exists J, K$ with $\text{Tail}_J(\vec{x}) \subset A$ and $\text{Tail}_K(\vec{x}) \subset B$. Setting $N := \text{Max}(J, K)$ gives the \ast that nv-set $\text{Range}(\text{Tail}_N(\vec{x}))$ lies in both A and B .

A TS with property (20) is called a *Hausdorff space*; agree to use *HS* to abbreviate this.

Closure/Interior/Bdry etc. Fix a TS Ω and a set $S \subset \Omega$. A point $q \in \Omega$ is a “*closure point* of S ” if:

21: $\forall V^{\text{open}} \ni q$, the intersection $V \cap S \neq \emptyset$.

Use $\text{Cl}_\Omega(S)$ for the *set* of Ω -closure-points of S . Easily

21': $\text{Cl}_\Omega(S)$ is Ω -closed, and equals the intersection of all Ω -closed supersets of S ; hence, it is the smallest such.

A point $q \in \Omega$ is an “**interior point** of S ” if:

22: $\exists V^{\text{open}} \ni q$ such that $V \subset S$.

Use $\text{Itr}_\Omega(S)$ for the *set* of Ω -interior-points of S . And

22': $\text{Itr}_\Omega(S)$ is Ω -open, and equals the union of all Ω -open subsets of S ; hence, it is the largest such.

A set S is “a **Ω -neighborhood** of a point q ” if $\text{Itr}_\Omega(S) \ni q$. Equivalently, $\exists U^{\text{open}}$ with $S \supset U \ni q$. Write this as

$$q \in^{\text{nbhd}} S \quad \text{or} \quad S \ni^{\text{nbhd}} q.$$

Replacing q by a set, A , we say that “ S is a **neighborhood** of set A ” if $\text{Itr}(S) \supset A$. Analogously, write this relation as

$$A \subset^{\text{nbhd}} S \quad \text{or} \quad S \supset^{\text{nbhd}} A.$$

The “ **Ω -boundary** of S ”, written $\partial_\Omega(S)$ or $\text{Bdry}_\Omega(S)$, is $\text{Cl}(S) \cap \text{Cl}(\Omega \setminus S)$.

A point $q \in \Omega$ is a “**cluster point** of S ” iff

21'': $\forall V^{\text{open}} \ni q$: Intersection $V \cap S$ is infinite.

Use $\text{Clust}_\Omega(S)$ for the S 's set of cluster³ points.

Switching from sets to sequences, a point q is

“a **limit-point**³ of sequence \vec{x} ”

if \vec{x} has some subsequence which converges to q .

Isomorphisms. A map $\varphi: \Omega \leftrightarrow X$ between two TSEs is a **homeomorphism**⁴ if φ is a bijection st.:

23: For each open set $\Lambda \subset \Omega$, the forward-image $\varphi(\Lambda)$ is X -open. And for each open set $S \subset X$, the inverse-image $\varphi^{-1}(S)$ is Ω -open. (Looking ahead, each of φ and φ^{-1} is continuous.)

A homeomorphism is a “topological isomorphism”.

Between two MSeS (Ω, μ) and (X, d) , an **isometry**⁵ is a bijection $f: \Omega \leftrightarrow X$ which preserves distance: For all $\alpha_1, \alpha_2 \in \Omega$, we have $d(f(\alpha_1), f(\alpha_2)) = \mu(\alpha_1, \alpha_2)$.

³Terms *cluster point*, *accumulation point* and *limit point* are related. Alas, textbooks vary as to which term they assign to which concept.

⁴From Greek ομοιος (homoios) “similar”, and μορφη (morph) “form”, “shape”.

⁵From Greek ισος (isos), “equal”, and μορφη (morph).

Defn: Relative topology. In a TS Ω with subset X , how should we define the X -open subsets? Motivated by the Induced-topology Lemma, (6), we specify that

24: A subset $U \subset X$ is **X -open** IFF there exists an Ω -open set \widehat{U} such that $\widehat{U} \cap X = U$.

The collection of such sets U is indeed a topology on X (fulfilling axioms (TS1,2,3)). It is called the **relative topology** or **induced topology** on X .

25: Lemma. For a subset S of a Hausdorff TS: A point q is a cluster-point of S IFF each $V^{\text{open}} \ni q$ owns a point of S different from q . **Proof.** Exercise. \diamond

Locally Countably Generated spaces. Consider a MS Ω and point $q \in \Omega$. Evidently, by letting $U_n := \text{Bal}_{1/n}(q)$,

There exists \vec{U} , a countable family $U_1 \supset U_2 \supset \dots$
26: of Ω -open sets, each owning q . Moreover for each open $V \ni q$, there is some n with $V \supset U_n$.

Such a \vec{U} is called a “**countable local-base** for q ”. A TS Ω is **LCG** (*locally countably-generated*) if each $q \in \Omega$ has a countable local-base. (The std phrase is “ Ω is **first-countable**”.)

27: Sequence-Closure Lemma. In TS Ω , consider a subset $S \subset \Omega$ and point $q \in \Omega$.

a: If there exists a sequence $\vec{\sigma} \subset S$ with $\lim(\vec{\sigma}) = q$, then $q \in \text{Cl}(S)$.

b: Now suppose that Ω is LCG. If $q \in \text{Cl}(S)$ then $\exists \vec{\sigma} \subset S$ such that $\lim(\vec{\sigma}) = q$, \diamond

Proof. Leaving (a) as an **exercise**, let's show (b).

Fix \vec{U} as in (26). Each U_n intersects S , since $q \in \text{Cl}(S)$, so we may pick a point $\sigma_n \in U_n \cap S$.

Given an open $V \ni q$, there exists N with $U_N \subset V$. For each $k \geq N$, then, $\sigma_k \in U_k \subset U_N \subset V$. I.e, $\text{Tail}_N(\vec{\sigma}) \subset V$. \blacklozenge

\diamond Every TS satisfies (27a). But conclusion (27b) can fail in a non-LCG space. It fails in the cartesian-power space $\{0, 1\}^{\mathbb{R}}$.

Lemma 27 implies, in an LCG space, that a set is closed IFF it is **(sequentially-)inescapable**. \square

Compactness

A TS X is **sequentially compact** (*seq-cpt*) if each X -sequence has a X -convergent subsequence.

28: Lemma. *In a MS Ω (Hausdorff & LCG, suffices), suppose subset X is sequentially compact. Then X is Ω -closed.* \diamond

Proof. Fix an arbitrary Ω -convergent seq $\vec{x} \subset X$. Let $\omega := \lim(\vec{x}) \in \Omega$. Since X is sequentially-cpt, there exists an X -convergent subseq $\vec{y} \subset \vec{x}$; so $z := \lim(\vec{y})$ is in X . But \vec{x} is Ω -convergent, so $\omega \stackrel{\text{must}}{=} z$. Thus $\omega \in X$.

This holds for each sequence $\vec{x} \subset X$, so X is Ω -inescapable. But Ω is a MS, so (27) applies and tells us that X is Ω -closed. \diamond

A TS X is **cluster-point compact** (*cluster-pt cpt*) if each infinite subset $S \subset X$ has a cluster-point in X .

29: Lemma. *For a general TS Ω :*

a: Sequentially compact \implies Cluster-point compact.

b: If Ω is LCG, then Cluster-point compactness implies Sequential-compactness. \diamond

Pf of (a). Consider an ∞ -subset $S \subset X$. For $n = 1, 2, 3, \dots$, pick a point

$$*: \quad b_n \in S \setminus \{b_1, b_2, \dots, b_{n-1}\};$$

this is possible, since S is infinite. Since X is seq-cpt, there is a subseq $\vec{a} \subset \vec{b}$ which is X -convergent; let $q := \lim(\vec{a})$. Now $\vec{a} \subset \vec{b} \subset S$, so q is a closure-point of S . But \vec{a} has distinct terms, since \vec{b} does, courtesy (*). Thus q is, in fact, a *cluster-point* of S . \diamond

Pf of (b). Fix a seq $\vec{b} \subset X$. A constant subseq is certainly convergent, so WLOG no value in \vec{b} occurs ∞ -often. Hence we can let $\vec{c} \subset \vec{b}$ be the subsequence obtained by keeping just the *first occurrence* of each value in \vec{b} . Automatically, \vec{c} has distinct terms, so $\{c_\ell\}_{\ell=1}^\infty$ is infinite, and thus has a cluster-point; pick one such, and call it q .

For q , fix countable local-base \vec{U} as in (26). Set $N_0 := 0$. For $k = 1, 2, \dots$, let N_k be the smallest index $n > N_{k-1}$ st. $c_n \in U_k$. Such an n exists, since U_k owns ∞ many points from $\{c_\ell\}_1^\infty$, and the $\ell \mapsto c_\ell$ map is injective.

Let $e_k := c_{N_k}$. To show that seq \vec{e} converges to q , fix an open set $V \ni q$, then take K (smallest, say) so that $U_K \subset V$. But each index $k \geq K$ has

$$e_k \in U_k \subset U_K \subset V.$$

I.e, $\text{Tail}_K(\vec{e}) \subset V$. \diamond

Covers. For $X \subset \Omega$, an “ Ω -cover of X ” is a collection $\mathcal{C} \subset \mathcal{P}(\Omega)$ for which $X \subset \bigcup(\mathcal{C})$. A subset $\mathcal{S} \subset \mathcal{C}$ is a **subcover** (of X) if $\bigcup(\mathcal{S}) \supset X$. The elements of a cover are sometimes called *patches*.

An **open cover** has each patch open. Inconsistently, a cover \mathcal{C} is a **finite cover** if $|\mathcal{C}| < \infty$.

A TS X is **compact** IFF each X -open-cover \mathcal{C} , of X , has (some folks say, “admits”) a finite subcover. In practice, X is a subset of some TS Ω . Courtesy (24) (and (6), indirectly):

30: *X is compact IFF each Ω -open-cover of X has a finite subcover.*

31: Diameter/compactness Prop'n. *Suppose $\text{Diam}(X^{MS})$ is infinite. Then X is not compact.* \diamond

Pf. Since X non-void ($\text{Diam} > 0$), we can pick a point $z \in X$. Let B_n be the center= z ball of radius- n . Thus $\mathcal{C} := \{B_n\}_{n=1}^\infty$ is an open-cover of X . It has no finite subcover, since such would force $\text{Diam}(X) < \infty$. \diamond

32: Compact-intervals theorem. *For all reals $a \leq b$, the closed interval $J := [a, b]$ is compact.* \diamond

Pf. WLOG, $J = [3, 7]$. Given an arbitrary cover \mathcal{C} of J by \mathbb{R} -open sets, ISTProduce a finite subcover.

So our job is to show that 7 is *good*, where an $x \in J$ is “good” IFF there exists a finite subcollection $\mathcal{F} \subset \mathcal{C}$ covering $[3, x]$. We’ll first show that this number,

$$\dagger: \quad z := \sup\{x \in J \mid x \text{ is good}\},$$

exceeds 3. We’ll then show that z is good, and equals 7.

Some patch $P \in \mathcal{C}$ owns 3, so $\exists \delta > 0$ with

$$P \supset [3 - \delta, 3 + \delta].$$

So singleton $\{P\}$ covers $[3, 3 + \delta]$. WLOG $3 + \delta \leq 7$; thus $z \geq 3 + \delta$. Hence $\boxed{z > 3}$.

Also, some \mathcal{C} -patch Q owns z . So $\exists \varepsilon > 0$ with

$$\ddagger: \quad Q \supset [z - \varepsilon, z + \varepsilon] \supset [z - \varepsilon, z],$$

and we can shrink ε so that $z - \varepsilon \geq 3$. (Here is where we use that z exceeds 3.) Automatically (why?), the number $z - \varepsilon$ is good; let $\mathcal{F} \subset \mathcal{C}$ be a finite family which covers $[3, z - \varepsilon]$. Then $\mathcal{F} \cup \{Q\}$ covers $[3, z]$. I.e., $\boxed{z \text{ is good}}$.

Lastly, FTSCContradiction suppose $z < 7$. Then we could have taken ε so small that $z + \varepsilon \leq 7$. But $\mathcal{F} \cup \{Q\}$ covers interval $[3, z + \varepsilon]$, thus showing that $z + \varepsilon$ is good; And *that* rudely contradicts (\ddagger). \blacklozenge

Metric ideas related to compactness

A MS Ω is **totally-bounded** (abbrev.: **TB**) if: For each $\varepsilon > 0$, there exists a cover of Ω by finitely many ε -balls.

33: TB-iff-CauchySubseq Thm. *MS X is totally-bounded IFF each seq $\vec{a} \subset X$ has a Cauchy subsequence $\vec{c} \subset \vec{a}$.*

Corollary. *A MS is complete and totally-bounded IFF it is sequentially compact.* \blacklozenge

Proof: TotBnded \Rightarrow Every-seq-has-a-CauchySubseq.

For each $K = 1, 2, \dots$ we can, by hypothesis,

$$33a: \quad \text{let } B_1^K, B_2^K, \dots, B_{L_K}^K \text{ be a finite list of radius-}\frac{1}{K} \text{ balls, whose union is } X.$$

Fixing a sequence $\vec{a} \subset X$, our goal is to produce a subsequence which is Cauchy.

Define index sets $I_1 := \mathbb{Z}_+ \supset I_2 \supset I_3 \supset \dots$, as follows. At stage K , with I_{K-1} defined, let B be the first ball in list (33a) that owns ∞ many indices from I_{K-1} . I.e.,

$$33b: \quad I_K := \{i \in I_{K-1} \mid B \ni a_i\} \quad \text{is infinite.}$$

Automatically

$$33c: \quad \text{Diam}(\{a_i \mid i \in I_K\}) \leq \text{Diam}(B) \leq \frac{2}{K}.$$

Let $N_1 := 1$ and let each N_K be the smallest element of I_K that exceeds N_{K-1} ; possible, courtesy (33b).

To see that sequence $(a_{N_K})_{K=1}^\infty$ is Cauchy, fix $\varepsilon > 0$, then a K with $\frac{2}{K} < \varepsilon$. For each pair of indices j, ℓ dominating K , note that $N_j \in I_j \subset I_K$; ditto $N_\ell \in I_K$. By (33c), then,

$$\text{Dist}(a_{N_j}, a_{N_\ell}) < \varepsilon. \quad \blacklozenge$$

Pf: Every-seq-has-a-CauchySubseq \Rightarrow TB. FTSC, suppose X is *not* TB. So **there exists** a “bad” posreal ε st.

33d: *there is no finite cover of X by ε -balls.*

Use B_p to denote the radius- ε ball centered at a point p .

In X , pick points $p_1, p_2, \dots, p_K, \dots$ st. each

$$p_K \text{ is in none of } B_{p_1}, B_{p_2}, \dots, B_{p_{K-1}}.$$

This process never gets stuck, courtesy (33d). Hence $(p_K)_{K=1}^\infty$ is an (infinite) sequence, which certainly has no Cauchy-subseq, since each two entries are at least ε apart. Contradiction. \heartsuit^6 \blacklozenge

Lebesgue number. In a MS Ω , a posreal r is a **Lebesgue number** of an Ω -cover \mathcal{C} if:

For each $q \in \Omega$, there exists a patch $P \in \mathcal{C}$ for which $\text{Bal}_r(q) \subset P$.

For want of a better term, say that Ω is a “**cover-positive space**” if each open-cover has a Lebesgue number.

Note that $\Omega := \mathbb{Z}$ is cover-positive; indeed $r := 1$ is a Lebesgue number for *every* cover! That \mathbb{Z} fails to be compact does not contradict the below Compactness notions Thm *because*... \mathbb{Z} is *not* totally-bounded. \square

\heartsuit The equivalence in t.bel Compactness notions Thm does not hold in a general TS; neither *Compactness* nor *Sequential Compactness* implies the other. The uncountable product $Y := \{0, 1\}^{\mathbb{R}}$ is compact, but not seq-cpt. Conversely, equipping the first uncountable ordinal, ω_1 , with the order-topology, gives a seq-cpt space that is not cpt. \square

34: Compactness notions Theorem. *In (X, d) , a metric space, TFAEquivalent:*

- a: *X is sequentially-compact.*
- b': *X is totally-bounded and (metrically) complete.*
- b: *X is totally-bounded and cover-positive. (Leb. number.)*
- c: *X is compact.*
- d: *X is cluster-point compact.* \blacklozenge

Pf (a) \Rightarrow (b'). Seq-cptness gives totally-boundedness, using (33). To get completeness, fix a Cauchy-seq \vec{a} . By seq-cptness, \vec{a} has a convergent subseq; so (4C) implies that \vec{a} converges. \blacklozenge

\heartsuit^6 Note: In a space where this process $B_{p_1}, B_{p_2}, B_{p_3}, \dots$ never gets stuck, there is no reason for this collection of balls to cover X . Indeed, there are MSes where for each $r > 0$, no countable collection of r -balls can cover the space.

Pf (a) \Leftarrow (b'). Fix a seq \vec{a} . Hypothesis (b') and (33) show that \vec{a} has a Cauchy-subseq. And completeness forces this subseq to converge. Hence X is sequentially-compact. \blacklozenge

Pf (a) \Rightarrow (b). We get totally-boundedness from (33).

FTSOC, suppose there exists an open-cover \mathcal{C} with no Lebesgue number. So, fixing posreals $\varepsilon_n \searrow 0$, there is a point $y_n \in X$ st. ball $\text{Bal}_{\varepsilon_n}(y_n)$ lies in *no* \mathcal{C} -patch.

By seq-cptness, \vec{y} has a convergent subseq. Pick one such, rename it \vec{y} and let $q := \lim(\vec{y})$. Since \mathcal{C} covers X , there exists a patch $P \in \mathcal{C}$ with $P \ni q$.

Since P is open, there exists $\delta > 0$ st. $\text{Bal}_{2\delta}(q) \subset P$. Pick N big enough that $\varepsilon_N < \delta$ and $\text{Dist}(y_N, q) < \delta$. Now

$$\text{Bal}_{\varepsilon_N}(y_N) \subset \text{Bal}_{2\delta}(q) \subset P.$$

Alas, this contradicts the "FTSOC" paragraph. \blacklozenge

Pf (b) \Rightarrow (c). Given an open-cover \mathcal{C} , take a Lebesgue number $r > 0$. Since X is TB, there is a *finite* collection \mathcal{F} of radius- r balls that cover X . But r is a Leb-number for \mathcal{C} , so for each ball $B \in \mathcal{F}$ there is a patch $P \in \mathcal{C}$ that includes B . Pick one such and call it \tilde{B} .

Hence $\mathcal{C} := \{\tilde{B} \mid B \in \mathcal{F}\}$ is a finite family of \mathcal{C} -patches. But does it cover X ? Yes, since $\bigcup(\mathcal{C}) \supset \bigcup(\mathcal{F}) = X$. \blacklozenge

Pf (c) \Rightarrow (d). (This implication holds in *all* Topological Spaces.)

Fix a subset $S \subset X$ with no cluster-pts. To show S finite, note that each point $z \in X$ must have an open nbhd $V_z \ni z$ having finite intersection with S .

Family $\{V_z \mid z \in X\}$ is an open cover of X . So there exists a *finite* set $F \leq X$ st. $\{V_z\}_{z \in F}$ covers X . Thus

$$S = S \cap X = S \cap \left[\bigcup_{z \in F} V_z \right] = \bigcup_{z \in F} [S \cap V_z].$$

Being a finite union of finite sets, then, S must be finite. \blacklozenge

Pf (d) \Rightarrow (a). Follows from (29). \blacklozenge

For us, **Euclidean space** $\mathbb{R}^D = \mathbb{R} \times \dots \times \mathbb{R}$, is finite dimensional and equipped with $\|\cdot\|_2$, the **Euclidean norm**.

35: Product-space Convergence Lemma. In $\Omega := \mathbb{R}^D$, write the n^{th} term in sequence \vec{x} as

$$x_n = (b_n^1, b_n^2, b_n^3, \dots, b_n^D), \quad \text{with each } b_n^k \in \mathbb{R}.$$

Then \vec{x} converges in Ω IFF for each $k = 1, \dots, D$, the seq $n \mapsto b_n^k$ converges in \mathbb{R} . With $\beta^k := \lim_{n \rightarrow \infty} b_n^k$, moreover, $\lim(\vec{x})$ equals $(\beta^1, \dots, \beta^D) \in \Omega$. **Proof. Exercise.** \blacklozenge

36: Heine-Borel theorem. In Euclidean space $\Omega := \mathbb{R}^D$, a subset K is compact IFF K is Ω -closed and bounded. \blacklozenge

Pf. WELOG, $\Omega = \mathbb{R} \times \mathbb{R}$. Let's show that a closed rectangle

$$S := I \times J, \quad \begin{array}{l} \text{where } I := [a, b] \subset \mathbb{R} \\ \text{and } J := [c, d] \subset \mathbb{R}, \end{array}$$

is sequentially-compact. Consider a seq $\vec{x} \subset S$, with

$$x_n = (\alpha_n, \beta_n) \in I \times J.$$

Courtesy (32), Compact-intervals thm, and (34), our I is seq-cpt. So we can drop to a subseq (and rename) so that, now, $n \mapsto \alpha_n$ converges. Use cptness of J to subsequence again. The new \vec{x} converges, using (35), and this \vec{x} is a subseq of the original.

A closed subset, K , of a compact space is necessarily [Exer.] cpt. Now consider an Ω -closed and bounded set K . Being bnded, there exist closed intervals I and J so that $I \times J \supset K$. Since K is Ω -closed, this K is automatically $I \times J$ -closed; hence K is compact.

The converse. Fix an Ω -compact set K ; necessarily bounded, by (31). Were K not Ω -closed, there there'd be a sequence $\vec{x} \subset K$ which converges to a point $q \in \Omega \setminus K$. So no subseq could K -converge. \blacklozenge

Precompactness. In a topological space Ω , a subset $X \subset \Omega$ is Ω -precompact if $\text{Cl}_\Omega(X)$ is compact.⁷

The Heine-Borel thm is tantamount to saying that the precompact subsets of Euclidean space are precisely the bounded subsets.

Trying to characterize the precompact subsets of a general MS (Ω, \mathbf{d}) , leads naturally to the following nice problem. \square

Exer 1. In MS (Ω, \mathbf{d}) , suppose a subset $X \subset \Omega$ is totally-bounded. Must its closure $Y := \text{Cl}_\Omega(X)$ automatically be totally-bounded too? \square

(Yes, as shown by Andy, Michael R., Lindsay, Taylor, and ...)

⁷Recall that *compactness* is an absolute notion. However, *precompactness* depends on the closure operator, and is a relative notion. As an example, the interval $(0, 1)$ is \mathbb{R} -precompact, but is *not* Ω -precompact for $\Omega := [0, 1)$.

Proof. (All balls here are Ω -balls.) Fix $\varepsilon > 0$. The TBness of X hands us a finite set $F \subset X$ such that

†: The ε -balls $\{\text{Bal}_\varepsilon(c)\}_{c \in F}$ cover X .

ISTProve that the $[2\varepsilon]$ -balls with centers in $F \subset X \overset{\text{note}}{\subset} Y$ indeed cover Y . To this end, fix a point $P \in Y$. Being in the Ω -closure of X , there exists an $x \in X$ with $d(x, P) < \varepsilon$. By (†), there exists a point $c \in F$ with $d(c, x) < \varepsilon$. So

$$d(c, P) \leq d(c, x) + d(x, P) < \varepsilon + \varepsilon,$$

and the cavalry (i.e. Δ nequality) rides up and saves the day. ♦

Bufferable pairs of sets. (The following terminology is *provisional*, and may get changed. But the Mathematics will remain. . .)

In a TS Ω , a disjoint pair of sets E_1 and E_2 is **bufferable** if there exists disjoint open sets $U_j \supset E_j$. Usually just say that “ E_1, E_2 is a *bufferable pair*”.

Suppose that *foo* and *fum* are two properties that a subset of Ω might or might-not have. We’ll say that Ω is “*foo:fum-buffered*” if for each disjoint pair of sets, a *foo* E_1 and a *fum* E_2 , the pair (E_1, E_2) is bufferable. Examples are: Ω might be compact:compact-buffered or compact:closed-buffered.

We’ll call Ω point:compact-buffered if each point p (technically, each *singleton* $E_1 := \{p\}$) can be buffered from each compact set E_2 that does not own p . In this language, “ Ω is Hausdorff” means that Ω is point:point-buffered.

As an abbreviation, let “*fum-buffered*” mean *fum:fum-buffered*. □

37: Compactness lemma. Consider a TS Ω .

a: If Ω is compact, then each Ω -closed subset is compact.

b: Suppose Ω Hausdorff. Then Ω is point:compact-buffered, and each compact subset $Y \subset \Omega$ is Ω -closed. Further, Ω is compact-buffered.

c: Suppose Ω is metrizable. If d is a metric consistent with the topology, then (Ω, d) is complete. ♦

Proof of (a). (Let “open” mean “ Ω -open”.) Take an Ω -closed $Y \subset \Omega$ and Ω -open cover, \mathcal{C} , of Y . Thus $\{\Omega \setminus Y\} \cup \mathcal{C}$ is an Ω -open cover of Ω . So it has a finite subcover (of Ω), which we can write as $\{\Omega \setminus Y\} \cup \mathcal{F}$, where $\mathcal{F} \subset \mathcal{C}$ is finite. And therefore \mathcal{F} covers Y . ♦

Proof of (b). Fix a point $p \in \Omega \setminus Y$. For each point $z \in Y$, Hausdorffness gives disjoint open sets

$$U_z \ni p \quad \text{and} \quad V_z \ni z.$$

Compactness of Y asserts a *finite* set $Z \subset Y$ such that $\{V_z\}_{z \in Z}$ covers Y . It follows that these disjoint sets,

$$\widehat{U} := \bigcap_{z \in Z} U_z \quad \text{and} \quad \widehat{V} := \bigcup_{z \in Z} V_z,$$

are open. Since $\widehat{U} \ni p$ and $\widehat{V} \supset Y$, we have buffered the p, Y pair.

Renaming \widehat{U} to U_p , we have that

$$\bigcup_{p \in \Omega \setminus Y} U_p \overset{\text{note}}{=} \Omega \setminus Y$$

is Ω -open. Thus Y is Ω -closed.

Lastly, fix disjoint compact sets $C, Y \subset \Omega$. For each point $p \in \Omega \setminus Y$, there exist open sets $U_p \ni p$ and $V_p \supset Y$, with $U_p \cap V_p = \emptyset$. Since $\{U_p\}_{p \in C}$ is an open-cover of C , there exists a finite set $F \subset C$ so that $\{U_p\}_{p \in F}$ already covers C . Automatically, these open sets,

$$\widehat{U} := \bigcup_{p \in F} U_p \quad \text{and} \quad \widehat{V} := \bigcap_{p \in F} V_p,$$

are disjoint from each other. Finally, $\widehat{U} \supset C$ and $\widehat{V} \supset Y$. ♦

Pf of (c). Fix a Cauchy sequence \vec{x} . Sequential-compactness says there exists a convergent subseq $\vec{y} \subset \vec{x}$. So (4C) of the MS-sequence Thm tells us that \vec{x} converges. ♦

Continuity

A map $f: (X, \mathbf{d}) \rightarrow (\Omega, \mu)$ is “continuous at $q \in X$ ” if:

$\forall \varepsilon > 0, \exists \delta = \delta(q, \varepsilon) > 0$ such that:

38: $\forall x \in X, \text{ if } \mathbf{d}(x, q) < \delta \text{ then } \mu(f(x), f(q)) < \varepsilon.$

I.e, setting $\omega := f(q)$: $\text{Bal}_\delta(q) \subset f^{-1}(\text{Bal}_\varepsilon(\omega)).$

And “ f is **continuous**” if it is cts at each point q of its domain. Notice that the $\delta = \delta(\varepsilon, q)$ depends on both ε and q . In contrast, f is **uniformly continuous** if we can quantify q after δ :

$\forall \varepsilon > 0, \exists \delta = \delta(\varepsilon) > 0$ st. for each $q \in X$:

39: $\text{Bal}_\delta(q) \subset f^{-1}(\text{Bal}_\varepsilon(f(q))).$

Equivalently: $\forall \varepsilon > 0, \exists \delta > 0$ st. $\forall x, q \in X, \text{ if } \mathbf{d}(x, q) < \delta$ then $\mu(f(x), f(q)) < \varepsilon.$ Exer: Prove this equivalence!

Metricless continuity. Our defn (39) of uniform continuity seems to really *use* a metric. But just “continuity at a point”, (38), can be stated purely in terms of open sets:

38': For each Ω -open $\Lambda \ni f(q)$, its inverse-image $f^{-1}(\Lambda)$ is a neighborhood⁸ of q .

(Again equivalently: Each Ω -nbhd Λ of $f(q)$ has its inverse-image being a nbhd of q .) Indeed, for a map $f: X \rightarrow \Omega$ between general TSES, we take (38') as our *definition* of

“ f is **continuous** at q ”.

We use $\text{Cty}(f)$ for the **continuity set** of f ; those $q \in X$ at which f is continuous. Use

$$\text{DisCty}(f) := X \setminus \text{Cty}(f)$$

for f 's **discontinuity set** See examples (45) and (47). □

In the case where f is continuous *everywhere* we can, in (38'), simplify “neighborhood” to “open set”.

40: Baby continuity Lemma. A map $f: X \rightarrow \Omega$ between topological spaces is continuous IFF $f^{-1}(\Lambda)$ is X -open, for each Ω -open set Λ . Proof. Exercise. ◇

41: Uniform-continuity Theorem. Consider a continuous map $f: (X, \mathbf{d}) \rightarrow (\Omega, \mu)$ between MSes. If X is compact, then f is uniformly continuous. ◇

⁸Even with f continuous at q , discontinuities at *other* points can ruin $f^{-1}(\Lambda)$ being open; whence the weaker requirement that $f^{-1}(\Lambda)$ have q in its interior.

Proof. FTSOC, suppose we have an $\varepsilon > 0$ for which no δ is small enough. I.e, there are seqs $\vec{a}, \vec{b} \subset X$ such that

$$\dagger: \lim_{n \rightarrow \infty} \mathbf{d}(a_n, b_n) = 0. \text{ And } \forall n: \mu(f(a_n), f(b_n)) \geq \varepsilon.$$

Since X is seq-cpt, are indices $N_1 < N_2 < \dots$ so that $\alpha := \lim_{j \rightarrow \infty} a_{N_j}$ exists in X . Rename \vec{a} to this $(a_{N_j})_1^\infty$, by re-indexing \vec{a} and \vec{b} . Now $\lim(\vec{a}) = \alpha$, and (\dagger) still holds.

Use seq-cptness again to drop to a convergent subseq of \vec{b} ; then re-index. So now, $\beta := \lim(\vec{b})$ exists.

Continuity of f at α and β , and (\dagger) , implies (Exer: do this!) that

$$\ddagger: \mu(f(\alpha), f(\beta)) \geq \varepsilon \stackrel{\text{recall}}{>} 0.$$

OTOHand, the Δ inequality and (\dagger) imply (Exer: show this!) that $\mathbf{d}(\alpha, \beta) = 0$. Hence $\alpha = \beta$. But this contradicts (\ddagger) . ◆

2nd proof. Fix $\varepsilon > 0$. Let \mathcal{D} be the set of ε -balls in Ω . So

$$\mathcal{C} := \{f^{-1}(\Lambda) \mid \Lambda \in \mathcal{D}\}$$

is an open-cover of X , courtesy (40). By the Compactness notions theorem, \mathcal{C} has a Lebesgue number $r > 0$.

Consider two points $x, y \in X$ less than r apart. Since $x, y \in \text{Bal}_r(x)$, there exists a \mathcal{C} -patch $P = f^{-1}(\Lambda)$ owning both. Hence $f(x)$ and $f(y)$ lie in a common Ω - ε -ball, Λ . ◆

Defn. Examine map $f: (X, \mathbf{d}) \rightarrow (\Omega, \mu)$ between MSes. The posreal γ is a **Lipschitz bound** for f if:

$$\forall x, y \in X: \text{Distance } \mu(f(x), f(y)) \leq \gamma \cdot \mathbf{d}(x, y) .$$

A fnc f is **Lipschitz continuous** IFF $\exists \mathcal{U} \in [0, \infty)$ so that:

$$42: \forall x, y \in X: \text{Distance } \mu(f(x), f(y)) \leq \mathcal{U} \cdot \mathbf{d}(x, y)$$

Such a \mathcal{U} is called “a Lipschitz bound for f ”. The infimum of such is “the **Lipschitz constant** of f ”, and is written $\boxed{\text{Lip}(f)}$. Easily,

$$\text{Lipschitz continuity} \implies \text{uniform continuity}.$$

The converse does not hold: The function $\mathbb{R} \rightarrow \mathbb{R}$ by $x \mapsto x^{1/3}$ is uniformly –but not Lipschitz– continuous. This also is an example of an invertible uniformly-cts function whose fnc-inverse is *not* uniformly continuous. □

43: Lip-Diff Lemma. On an interval J , suppose $f: J \rightarrow \mathbb{R}$ is differentiable. Then f is Lipschitz continuous IFF

$$\mathcal{U} := \sup_{x \in J} |f'(x)|$$

is finite; and then \mathcal{U} is $\text{Lip}(f)$, the Lipschitz constant of f . \diamond

Proof of (\Leftarrow). Fix $x \leq y$ in J . The Mean-Value Theorem asserts a point $c \in [x, y]$ such that

$$f(x) - f(y) = f'(c) \cdot [x - y].$$

Consequently, $|f(x) - f(y)| \leq \mathcal{U} \cdot |x - y|$. \diamond

Proof of (\Rightarrow). Exercise. \diamond

Definition. A map $h: (X, d) \rightarrow (\Omega, \mu)$ is **biLipschitz** if h is invertible, and both h^{-1} and h are Lipschitz maps.

Two metrics m and d , on the same space X , are **Lip-schitz equivalent** (*Lip-equiv*) if the identity map

$$x \mapsto x \text{ from } (X, d) \rightarrow (X, m)$$

is biLipschitz. We write $m \overset{\text{Lip}}{\asymp} d$. \square

44: Lemma. If $m \overset{\text{Lip}}{\asymp} d$ then $m \overset{\text{Cau}}{\asymp} d$. **Proof.** Exercise. \diamond

45: Indicator functions. Fix a set Ω . Each subset $S \subset \Omega$ yields a fnc $\mathbf{1}_S: \Omega \rightarrow \{0, 1\}$, the **indicator function**

$$\mathbf{1}_S(x) := \begin{cases} 1 & \text{when } x \in S \\ 0 & \text{when } x \in \Omega \setminus S \end{cases}.$$

Since the notation doesn't show the space (i.e, we don't write $\mathbf{1}_{S, \Omega}$), we sometimes write " $\mathbf{1}_S: \Omega \rightarrow \mathbb{R}$ " to emphasize the domain. For example: *What is the discontinuity-set of fnc $\mathbf{1}_{\mathbb{Q}}: \mathbb{R} \rightarrow \mathbb{R}$?* Answer: *All of \mathbb{R} .* But the discontinuity-set of $\mathbf{1}_{\mathbb{Q}}: \mathbb{Q} \rightarrow \mathbb{R}$ is empty; this fnc is constant-1, hence cts.

As another example, let J be the set of positive rationals whose square lies between 4 and 7. Let g mean $\mathbf{1}_J: \mathbb{Q} \rightarrow \mathbb{R}$, and f mean $\mathbf{1}_J: \mathbb{R} \rightarrow \mathbb{R}$. Use h for $\mathbf{1}_{[2, \sqrt{7}]}: \mathbb{R} \rightarrow \mathbb{R}$. Then

$$\text{DisCty}(g) = \{2\} \subset \mathbb{Q}, \quad \text{and}$$

$$\text{DisCty}(f) = [2, \sqrt{7}] \subset \mathbb{R}.$$

But $\text{DisCty}(h) = \{2\} \cup \{\sqrt{7}\}$, just a doubleton.

46: Prop'n. For a subset $E \subset \Omega$ of a topological space, $\text{DisCty}(\mathbf{1}_E) = \partial_\Omega(E)$. **Proof.** Exercise. \diamond

Ruler function. We are born grokking the **dyadic rationals**,

$$\mathbb{D} := \left\{ \frac{n}{2^e} \mid n \in \mathbb{Z} \text{ and } e \in \mathbb{N} \right\}.$$

Say that a fraction "n/d" is in **standard form** (LCTerms?) if $n \in \mathbb{Z}$ and $d \in \mathbb{Z}_+$, with $n \perp d$. (Std.form is unique. As a fraction, the std. form of 0 is 0/1.)

From a subset $S \subset \mathbb{Q}$, define the "**S-ruler function**" $\mathcal{R}_S: \mathbb{R} \rightarrow \mathbb{R}$ by

$$47: \quad \begin{aligned} \mathcal{R}_S\left(\frac{n}{d}\right) &:= \frac{1}{d}, & \text{for } \frac{n}{d} \in S \text{ in std.form;} \\ \mathcal{R}_S(x) &:= 0, & \text{for } x \in \mathbb{R} \setminus S. \end{aligned}$$

In the special case where $S := \mathbb{D}$, we call this just the **ruler function** $\mathcal{R} := \mathcal{R}_{\mathbb{D}}$. \square

Exer. 47.1: Ruler function \mathcal{R}_S is idempotent IFF the subset $S \subset \mathbb{Q}$ satisfies ... *What?*

The ruler fnc is interesting in that both its cty and its discty sets are dense in \mathbb{R} , as the next Observation shows.

48: Obs. For $S \subset \mathbb{Q}$ arbitrary, $\text{DisCty}(\mathcal{R}_S) = S$. \diamond

Proof of $\text{DisCty}(\mathcal{R}_S) \supset S$. Exercise. \diamond

Proof of $\text{DisCty}(\mathcal{R}_S) \subset S$. FTSOC, suppose a $\lambda \in \mathbb{R} \setminus S$ is a discty-point of \mathcal{R}_S . Then there exists a posint D and sequence $r_n \rightarrow \lambda$ with each $\mathcal{R}_S(r_n) \geq \frac{1}{D}$. So each r_n is in the set Q_D from (49), below. But (49) implies that Q_D has no cluster-points. Thus \mathbf{r} is eventually-constant, WLOG constant. So each r_n equals λ . Since $\mathcal{R}_S(\lambda) = 0$, this is an outrageous contradiction. \diamond

49: Lem HW1. For N a posint, let Q_N be the set of ratios $\frac{k}{\ell}$ with $k \in \mathbb{Z}$ and $\ell \in [1 .. N]$. Produce a posint P_N so that: For all distinct $x, y \in Q_N$, nec. $|x - y| \geq 1/P_N$. \diamond

Proof. Note that Q_1 is \mathbb{Z} , so $P_1 = 1$. It turns out that the value $P_N := N!$ works, but we can get a better formula when $(N \geq 2)$, which we henceforth consider.

Firstly, $\frac{1}{N-1} - \frac{1}{N} = \frac{1}{[N-1]N}$. So $P_N \geq [N-1]N$. Let's establish the reverse inequality, thus proving

$$49.1: \quad P_N = [N-1]N, \quad \text{for each } N \in [2.. \infty).$$

Write $x = \frac{\alpha}{k}$ and $y = \frac{\beta}{\ell}$ as ratios of integers, with k and ℓ in $[1..N]$. Setting $\mathbf{L} := \text{LCM}(k, \ell)$, observe that

$$x - y = \frac{m}{\mathbf{L}}, \quad \text{for some integer } m. \text{ This } m \neq 0, \text{ since } x \neq y.$$

Hence $|x - y| \geq \frac{1}{\mathbf{L}}$, so P_N is less-equal the max-value that \mathbf{L} can assume. Thus

$$49.2: \quad P_N \leq \text{Max} \{ \text{LCM}(k, \ell) \mid k, \ell \in [1..N] \}.$$

If $k = \ell$, then $\text{LCM}(k, \ell) \leq N$. Thus $\text{LCM}(k, \ell) \leq [N-1]N$, since $N-1 \geq 1$. Conversely, if $k < \ell$, then $\text{LCM}(k, \ell) \leq k \cdot \ell \leq [N-1]N$. In either case, we get the "reverse inequality", courtesy (49.2). Hence (49.1). \blacklozenge

50: Lem HW2. Consider $\lambda \in \mathbb{R}$ and integers $b_n > 0$ and a_n (not-nec coprime) such that $r_n \rightarrow \lambda$, where $r_n := \frac{a_n}{b_n}$, yet each $r_n \neq \lambda$. Then $b_n \rightarrow \infty$, as $n \nearrow \infty$. \blacklozenge

Piecewise-linear functions. Consider a closed interval $J := [a, b] \subset \mathbb{R}$ and a tuple \vec{p} of **cutpoints** of J ,

$$a = p_0 < p_1 < p_2 \dots < p_{N-1} < p_N = b.$$

Call the subinterval $B_k := [p_{k-1}, p_k]$ the " **k^{th} block**" of \vec{p} ". A function $g: J \rightarrow \mathbb{R}$ is "**piecewise linear**" on J "

i: if g is continuous and

ii: each restriction $g|_{B_k}$ has a straight-line graph.

Using the heights $h_k := g(p_k)$, here is the formula for $g(x)$ when $x \in B_4$:

$$g(x) := \left[\frac{x-p_4}{p_3-p_4} \cdot h_3 \right] + \left[\frac{x-p_3}{p_4-p_3} \cdot h_4 \right].$$

Turning this around, a cutpoint-tuple \vec{p} and a "height-tuple" $\vec{h} = (h_0, h_1, \dots, h_N)$ of reals, engenders a **PL** (piecewise linear) fnc. For $x \in B_k$,

$$51: \quad \text{PL}_{\vec{p}, \vec{h}}(x) := \left[\frac{x-p_k}{p_{k-1}-p_k} \cdot h_{k-1} \right] + \left[\frac{x-p_{k-1}}{p_k-p_{k-1}} \cdot h_k \right].$$

More generally, we can have $\text{PL}_{\vec{p}, \vec{h}}$ map interval J into a real **vectorspace** \mathbf{W} . Each h_k is a vector in \mathbf{W} , and each ratio, e.g $\frac{x-p_3}{p_4-p_3}$, is a scalar in \mathbb{R} . \square

Continuity and VSes. Given TSes X and Ω , let $\mathbf{C}(X \rightarrow \Omega)$ be the set of *continuous* functions $X \rightarrow \Omega$.

Usually Ω is a MS; suppose μ is its metric. We can define an extended-metric μ_{sup} on $\mathbf{C}(X \rightarrow \Omega)$ by:

$$52: \quad \mu_{\text{sup}}(f, g) := \sup_{x \in X} \mu(f(x), g(x)).$$

An $f \in \mathbf{C}(X \rightarrow \Omega)$ is **bounded** if $\text{Diam}(\text{Range}(f)) < \infty$. Use $\mathbf{C}_{\text{Bnd}}(X \rightarrow \Omega)$ for these; note that on this set, μ_{sup} is an actual metric.

When Ω is a real-VS \mathbf{W} , the set $\mathbf{V} := \mathbf{C}(X \rightarrow \mathbf{W})$ becomes a \mathbb{R} -VS under **pointwise operations**

$$[f+g](x) := f(x) + g(x), \text{ and } [5f](x) := 5f(x).$$

Putting a norm $\|\cdot\|$ on \mathbf{W} engenders the supremum-norm

$$\|f\|_{\text{sup}} := \sup_{x \in X} \|f(x)\|, \quad \text{on } \mathbf{V},$$

which is necessarily finite when X is compact, thus making $(\mathbf{V}, \|\cdot\|_{\text{sup}})$ a normed-VS. (When X non-compact, we can use $\mathbf{C}_{\text{Bnd}}(X \rightarrow \mathbf{W})$ as a normed-VS.) \square

53: P.L-approximation thm. Fix $J := [a, b] \subset \mathbb{R}$, normed-VS $(\mathbf{W}, \|\cdot\|)$, and continuous $f: J \rightarrow \mathbf{W}$. Then, given $\varepsilon > 0$, there exists a P.L function $g: J \rightarrow \mathbf{W}$ with $\|f - g\|_{\text{sup}} \leq \varepsilon$. \blacklozenge

Proof. For free, f is unif-cts since J is cpt. Pick posint N large enough that, with $\delta := \frac{b-a}{N}$: For all pairs $x, y \in J$,

$$|x - y| \leq \delta \implies \|f(x) - f(y)\| \leq \varepsilon/2.$$

Define cutpoints $p_k := a + k\delta$ and heights $h_k := f(p_k)$, for $k = 0, 1, \dots, N$. Is the $g := \text{PL}_{\vec{p}, \vec{h}}$ function ε -close to f ?

WELOG, fix an $x \in B_4$. Since $g|_{B_4}$ is linear,

$$\begin{aligned} \|g(p_4) - g(x)\| &\leq \|g(p_4) - g(p_3)\| \\ &= \|h_4 - h_3\| \stackrel{\text{Why?}}{\leq} \varepsilon/2. \end{aligned}$$

Now $\|x - p_4\| \leq \delta$, so $\|f(x) - f(p_4)\| \leq \varepsilon/2$. By the Δ nequality, difference $\|f(x) - g(x)\|$ is less-equal the sum

$$\begin{aligned} &\|f(x) - f(p_4)\| + \|f(p_4) - g(p_4)\| + \|g(p_4) - g(x)\| \\ &\leq \varepsilon/2 + \|h_4 - h_3\| + \varepsilon/2 = \varepsilon. \end{aligned} \quad \blacklozenge$$

Defn. A fnc $h := \text{PL}_{\vec{p}, \vec{h}}$ is a **rational-P.L function** if every cutpoint and height is rational. More generally, given an (open, closed, half-open) interval $I \subset [p_0, p_N]$, its *restriction* $f := h \downarrow_I$ is also called “ \mathbb{Q} -piecewise-linear”. This allows us to define “rational-P.L” on intervals whose endpoints are not rational. \square

54: Theorem. On a bounded interval $J \subset \mathbb{R}$, have \mathcal{Q} denote the set of \mathbb{Q} -piecewise-linear functions. Then \mathcal{Q} is countable. Moreover, \mathcal{Q} is $\|\cdot\|_{\text{sup}}$ -dense in the set of all P.L fncs on J .

When J is compact, then \mathcal{Q} is $\|\cdot\|_{\text{sup}}$ -dense in $\mathbf{C}(J \rightarrow \mathbb{R})$. Thus $\mathbf{C}(J \rightarrow \mathbb{R})$ becomes a CSD normed-VS. \diamond

Proof. Exercise. Use the P.L-approximation thm. \diamond

Uniform Convergence

Consider a TSes X and Ω , as well as functions $g, f_n: X \rightarrow \Omega$. Let \vec{f} denote this sequence (f_1, f_2, \dots) . Say that “Sequence \vec{f} **converges pointwise** to g ” if

$$\forall x \in X: f_n(x) \xrightarrow{n \rightarrow \infty} g(x).$$

Now suppose (Ω, μ) is a MS, and use μ_{sup} from (52) as a metric on fncs. If we have that

$$*: \mu_{\text{sup}}(f_n, g) \rightarrow 0, \text{ as } n \nearrow \infty,$$

then say that “sequence \vec{f} **converges uniformly** to g ”.

When Ω is a normed-VS $(\Omega, \|\cdot\|)$ then we can restate (*) as $\|f_n - g\|_{\text{sup}} \rightarrow 0$.

55: Uniform-convergence theorem. With notation from above: If each f_n is continuous, and $f_n \xrightarrow[n \rightarrow \infty]{\text{uniformly}} g$, then g is continuous.

Now suppose that (Ω, μ) is a complete metric-space (a CMS). Then $\Lambda := \mathbf{C}_{\text{Bnd}}(X \rightarrow \Omega)$ is complete with respect to the μ_{sup} metric. \diamond

Proof. Let m denote the metric μ_{sup} from (52).

Fix a point $P \in X$ and an $\varepsilon > 0$. Pick N large enough that $m(f_N, g) \leq \varepsilon$; WELOG, suppose $N = 7$.

Since f_7 is continuous at P , there exists an X -open set $U \ni P$ for which: If $x \in U$ then

$$\mu(f_7(x), f_7(P)) < 3\varepsilon.$$

For such an x , note that $\mu(g(x), g(P))$ is dominated by

$$\begin{aligned} &\mu(g(x), f_7(x)) + \mu(f_7(x), f_7(P)) + \mu(f_7(P), g(P)) \\ &\leq \varepsilon + 3\varepsilon + \varepsilon = 5\varepsilon. \end{aligned}$$

Completeness of Λ . Consider an m -Cauchy sequence $\vec{f} \subset \Lambda$. Fix a $z \in X$. For each pair of indices j and k ,

$$\mu(f_j(z), f_k(z)) \leq m(f_j, f_k);$$

so $n \mapsto f_n(z)$ is μ -Cauchy. Call its limit $g(z)$.

This defines a (not-nec cts) fnc $g: X \rightarrow \Omega$, which is the pointwise limit of \vec{f} . **Exer: Show $f_n \rightarrow g$ uniformly.**

To demonstrate that \vec{f} is Λ -convergent, we need to prove that the above g is in Λ , i.e. that g is continuous and bounded. The continuity follows from the uniform convergence. As for boundedness, pick N large enough that $m(f_N, g) < 17$. The Δ nequality then shows (Exer: exercise) that

$$\text{Diam}(\text{Range}(g)) \leq \text{Diam}(\text{Range}(f_N)) + 34. \quad \diamond$$

56: Weird Appl. of Unif-Conv. Suppose $f_n \xrightarrow{\text{unif.}} g$, for maps $g, f_n: X^{\text{TS}} \rightarrow (\Omega^{\text{MS}}, \mu)$. Consider points $y, z_k \in X$ with $z_k \rightarrow y$. If $y \in \text{Cty}(g)$ then

$$\ddagger: \lim_{\substack{n \rightarrow \infty \\ k \rightarrow \infty}} f_n(z_k) = g(y). \quad \diamond$$

Proof. Fix $\varepsilon > 0$. Choose an index N large enough that

$$\mu_{\text{sup}}(f_n, g) \leq 2\varepsilon, \quad \text{for each } n \geq N.$$

Since g is continuous at y , we can take K so that

$$\mu(g(z_k), g(y)) \leq \varepsilon, \quad \text{for each } k \geq K.$$

For all $n \geq N$ and $k \geq K$, then,

$$\begin{aligned} \ddagger: \mu(f_n(z_k), g(y)) &\leq \mu(f_n(z_k), g(z_k)) + \mu(g(z_k), g(y)) \\ &\leq 2\varepsilon + \varepsilon = 3\varepsilon. \end{aligned} \quad \diamond$$

Exer 2. Modify the proof of Uniform-convergence thm to show: Suppose $f_n \xrightarrow{\text{unif.}} g$, for maps $g, f_n: (X, d) \rightarrow (\Omega, \mu)$. If each f_n is uniformly continuous, then so is g . \square

57: Unif-conv Composition Lemma. Consider sets Z, Y and MSes X and Ω . For maps $f_n, g: Y \rightarrow X$, suppose $f_n \xrightarrow{\text{unif.}} g$, as $n \rightarrow \infty$. Then the following hold.

$$i: \text{For an arbitrary fnc } \beta: Z \rightarrow Y: [f_n \circ \beta] \xrightarrow[n \rightarrow \infty]{\text{unif.}} [g \circ \beta].$$

ii: Suppose map $\alpha: X \rightarrow \Omega$ is uniformly continuous. Then

$$[\alpha \circ f_n] \xrightarrow[n \rightarrow \infty]{\text{unif.}} [\alpha \circ g], \text{ as } n \rightarrow \infty. \quad \diamond$$

Proof. Exercise 3. \diamond

What does nesting give? Use “ $f_n \searrow g$ ” to mean, for each x , that $n \mapsto f_n(x)$ is decreasing, and decreases to $g(x)$.

58: Nested uniform-convergence thm (Nested UC). On a metric space X , suppose functions $g, f_n: X \rightarrow \mathbb{R}$ are continuous, and $f_n \searrow g$ pointwise. Then $f_n \xrightarrow[n \rightarrow \infty]{\text{uniformly}} g$, if either:

i: Space X is compact, or

ii: $\forall \varepsilon > 0, \exists$ an index K such that the set

$$\{x \in X \mid |f_K - g|(x) \geq \varepsilon\} \text{ is compact.} \quad \diamond$$

Preliminary reduction. Use $\|\cdot\|$ for $\|\cdot\|_{\text{sup}}$. Replace “ f_n ” by $f_n - g$ (which is continuous, since f_n and g are) and replace “ g ” by $\mathbf{0}$, the zero-function. By hypothesis,

$$f_1 \geq f_2 \geq f_3 \geq \dots \geq \mathbf{0}, \text{ pointwise.}$$

So $n \mapsto \|f_n\|$ is decreasing (non-increasing) and thus \vec{f} converges uniformly IFF $\forall \varepsilon, \exists N$ with $\|f_N\| \leq \varepsilon$.

In particular, ISTShow that some *subseq* of \vec{f} converges uniformly. \square

Proof of (i). FTSOC, suppose $\inf_n \|f_n\|$ dominates, say, 7. So there are points $y_n \in X$ with

$$\dagger: \quad f_n(y_n) \geq 6.$$

Since MS X is cpt, it is seq-cpt, so we can subsequence and renumber so that

$$\ddagger: \quad z := \lim_{n \rightarrow \infty} y_n \text{ exists in } X.$$

But $f_n(z) \xrightarrow{n \rightarrow \infty} 0$. WLOG $f_1(z) < 5$. Since f_1 is continuous at z , there is an open set $U \ni z$ on which $f_1 \downarrow_U < 5$.

But each $f_n \leq f_1$, so $f_1(y_n) \geq f_n(y_n) \geq 6$, by (\dagger). Thus no y_n point is in U . This is a grave insult to (\ddagger). \diamond

Pf of (ii). Fix $\varepsilon > 0$. Pick K st. $C := \{x \in X \mid f_K(x) \geq \varepsilon\}$ is compact. Part (i) tells us the restriction $f_n \downarrow_C$, as $n \rightarrow \infty$, converges *uniformly* to $\mathbf{0} \downarrow_C$. So we can pick an N large enough that $\|f_N \downarrow_C\| \leq \varepsilon$. We can also have taken $N \geq K$. Thus

$$\|f_N \downarrow_{[X \setminus C]}\| \leq \|f_K \downarrow_{[X \setminus C]}\| \leq \varepsilon.$$

Hence $\|f_N\| \leq \varepsilon$. \diamond

2nd proof of (i). Fix $\varepsilon > 0$. I’ll produce an N with $\|f_N\| \leq \varepsilon$.

Fix a $z \in X$. Since $f_n(z) \rightarrow 0$, there exists an index L with $f_L(z) < \varepsilon$; let L_z be the smallest such. Thus

$$U_z := \{x \in X \mid f_{L_z}(x) < \varepsilon\},$$

is an open set owning z .

Since $\{U_z \mid z \in X\}$ is an open cover of X , there exists a finite set $E \subset X$ with $\{U_z \mid z \in E\}$ covering X . I claim that

$$N := \text{Max}\{L_z \mid z \in E\}$$

satisfies $\|f_N\| \leq \varepsilon$. To see this, fix an arbitrary $y \in X$. There exists a $z \in E$ with $U_z \ni y$. Thus

$$0 \leq f_N(y) \leq f_{L_z}(y) \leq \varepsilon,$$

since \vec{f} is nested and $N \geq L_z$. \diamond

CEXes to Nested UC. On $X := \mathbb{R}$, let f_n be zero on $(-\infty, n]$, growing linearly from zero to three on $[n, n+1]$, and three on $[n+1, +\infty)$. So \vec{f} decreases pointwise to $\mathbf{0}$, but each $\|f_n\| = 3$. Ah!, but our X is not compact.

On compact $X := [5, 6]$, let f_n be piecewise-linear with cutpoints $(5, 6 - \frac{1}{n}, 6)$ and heights $(0, 0, 3)$. Although \vec{f} decreases pointwise to $g := 3 \cdot \mathbf{1}_{\{6\}}$, this \vec{f} does not converge uniformly. Oh!, but g is not continuous.

Keep $X := [5, 6]$. On $[5, 6)$, define h_n to be the above P.L. f_n , but define $h_n(6) := 0$. Now \vec{h} decreases pointwise to $\mathbf{0}$. Alas!, each h_n is not continuous. \square

Miscellaneous continuity/limit results

There are several elementary properties that we will use without proof, e.g, that a composition of cts fncs is continuous.

Composition notation. Consider fncs $A \xrightarrow{f} B$ and $B \xrightarrow{g} C$. The std notation for their composition is $g \circ f$, where $[g \circ f](a)$ means $g(f(a))$. It is sometimes convenient to have chiral versions of the composition operator. Define

$$59: \quad [f \triangleright g](a) := g(f(a)) \quad \text{and} \quad [g \triangleleft f](a) := g(f(a)).$$

So $g \triangleleft f$ is a synonym of $g \circ f$.

When a fnc maps a space to *itself*, $X \xrightarrow{f} X$, use f^{on} for the composition of n copies of f , the fnc $f \circ \dots \circ f$. \square

60: Prop'n. Suppose $f: X^{TS} \rightarrow \Omega^{TS}$ is cts. Let g denote the map f but with $\text{CoDom}(g) = f(X)$. Then g is cts.^{♡9} ♦

Proof. Fixing an $f(X)$ -open set U , there is an Ω -open set \widehat{U} st. $\widehat{U} \cap f(X) = U$. Now $f^{-1}(\widehat{U})$ is X -open, since f is cts. Thus $g^{-1}(U) = f^{-1}(\widehat{U})$ is X -open. ♦

61: Forward-inheritance Lemma. Consider a continuous map $f: X \rightarrow \Omega$ between TSes. Suppose X is compact or connected or path-connected. Then $f(X)$ has the same property. **Reduction.** WLOG, f is surjective. ♦

Pf of compactness. Let \mathcal{Y} be an open-cover of Ω . Its pull-back $\mathcal{C} := \{f^{-1}(P) \mid P \in \mathcal{Y}\}$ covers X . This is an X -open-cover, since f is cts. Compactness of X implies there exists a finite subset $\Phi \subset \mathcal{Y}$ for which $\{f^{-1}(P) \mid P \in \Phi\}$ covers X . Thus Φ covers Ω ; this, since f maps onto Ω . ♦

Pf of connectedness. Consider an Ω -open partition $\Omega = P \sqcup Q$ of Ω . The pull-backs $f^{-1}(P)$ and $f^{-1}(Q)$ form an X -open partition of X . Since X is connected, WLOG $f^{-1}(Q)$ is empty. Hence Q is empty, since f is surjective. ♦

Pf of path-connectedness. Fix points $\beta_0, \beta_1 \in \Omega$. Since f is onto, there exist points $b_i \in f^{-1}(\beta_i)$. And X is path-connected, so there is a cts map (a "path") $p: [0, 1] \rightarrow X$ with $p(0) = b_0$ and $p(1) = b_1$. Hence $p \triangleright f$ is a path from β_0 to β_1 . ♦

62: General limits. In MS (X, d) , centered at $q \in X$, the **punctured ball** of radius ε is

$$\text{PBal}_\varepsilon(q) := \{x \in X \mid 0 < d(x, q) < \varepsilon\}.$$

Consider a map $f: (X, d) \rightarrow (\Omega, \mu)$, points $q \in X$ and $\omega \in \Omega$. Analogous to (38) on P.12, we define

$$62.1: \quad \lim_{x \rightarrow q} f(x) = \omega.$$

to mean:

For each $\varepsilon > 0$ there exists $\delta > 0$ such that

$$62.2: \quad \text{PBal}_\delta(q) \subset f^{-1}(\text{Bal}_\varepsilon(\omega))$$

Extending this to general TSes X and Ω is routine. In the general case, (62.1) means the following.

For each Ω -open set $\Lambda \ni \omega$ there exists an

$$62.3: \quad X\text{-open set } U \ni q \text{ with} \quad U \setminus \{q\} \subset f^{-1}(\Lambda). \quad \square$$

^{♡9}This Prop'n is for convenience. It allows us to start some proofs with: "Our continuous function, WLOG, is surjective".

Miscellaneous connectedness results

In a TS Ω , the relation of two points being in the same *connected-component* is an equivalence relation. Also, *path-connected* is an equivalence relation.

63: Connected-interval Thm. Each interval J in \mathbb{R} is connected. (The interval can be infinite, or half-open, or ...) ♦

Proof. WELOG (exercise), $J = [3, 7]$, and let "open" mean J -open. Suppose we have an open-ptn $J = A \sqcup B$; so we have colored each point either Amber or Blue. WLOG, 3 is amber. To show there is no blue, we let

$$\dagger: \quad \alpha := J\text{-inf}(B) \stackrel{\text{note}}{\in} J.$$

Since 3 is in the interior of amber, there exists $\varepsilon > 0$ so that interval $[3, 3 + \varepsilon)$ is amber. Thus $\boxed{\alpha > 3}$.

Could α be blue? If yes, then since B is open there exists a posreal $\varepsilon < \alpha - 3$ so that interval $(\alpha - \varepsilon, \alpha]$ is blue. But this contradicts (\dagger) , so $\boxed{\alpha \text{ is amber}}$.

FTSOC, suppose $\alpha < 7$. Since A is open, there would exist an $\varepsilon > 0$ with $[\alpha, \alpha + \varepsilon)$ all amber. But $[\alpha, \alpha + \varepsilon)$ is amber, so this would force $J\text{-inf}(B) \geq \alpha + \varepsilon$, annoying (\dagger) .

The upshot: $\boxed{\alpha = 7}$ and consequently B is empty. ♦

§A Differentiability

Differentiability of a fnc $h: \mathbb{R} \rightarrow \mathbf{E}$, where $(\mathbf{E}, \|\cdot\|)$ is a normed-VS, is our goal. The Reader is to modify the discussion accordingly when the domain is just some (punctured) interval in \mathbb{R} , or when we are taking 1-sided derivatives; or when the domain is some subset of \mathbb{C} , with \mathbf{E} a complex normed-VS.

Suppose that h is defined in a nbhd of a point P in $\text{Dom}(h)$. Suppose (using (62)) the following limit exists:

$$64: \quad h'(P) := \lim_{x \rightarrow P} \frac{h(x) - h(P)}{x - P} \stackrel{\text{note}}{\in} \mathbf{E}.$$

Then we say that h is **differentiable** at P , and its derivative is the vector $h'(P)$. So h' is a vector-valued fnc just like h is, but with a possibly smaller domain.

How discontinuous can a derivative be? Extending, by continuity, the function

$$h(x) := x^2 \cdot \exp(-1/x^{57}),$$

is a simple example of an everywhere-differentiable fnc whose derivative is not cts, h' is not cts at the origin.

But certain kinds of discontinuities are ruled out.

65: Deriv-cty Lemma. Suppose $h: (a, c] \rightarrow \mathbb{R}$ is continuous, with h differentiable on (a, c) , and $L := \lim_{x \nearrow c} h'(x)$ exists. Then $h(\cdot)$, at c , has a lefthand derivative, which equals L . \diamond

Proof. Fixing an $\varepsilon > 0$, IStEstablish (65'), below. Pick $b \in (a, c)$ close enough to c that, letting $J := [b, c)$, the values of $h' \downarrow_J$ lie within ε of L .

For each $x \in J$, the MVT asserts a point $\dot{x} \in (x, c)$ with

$$\frac{h(c) - h(x)}{c - x} \stackrel{\text{MVT}}{=} h'(\dot{x}) \stackrel{\varepsilon}{\approx} L.$$

Consequently,

$$65': \quad \limsup_{x \nearrow c} \left| \frac{h(c) - h(x)}{c - x} - L \right| \leq \varepsilon. \quad \diamond$$

67: Appl. Fnc $h: [0, \infty) \rightarrow \mathbb{R}$ is diff'able on $J := (0, \infty)$.

i: Our h is continuous, with $h(0) = 0$. And...

ii: $\exists M \geq 0$ such that $\forall x \in J: |h'(x)| \leq M \cdot |h(x)|$.

Then h is constant-zero. \diamond

Pf. Fnc $g(x) := h(\frac{x}{M})$ fulfills (ii) for $M=1$. So WLOG

$$67: \quad \forall x \in J: |h'(x)| \leq |h(x)|.$$

And $\lim_{x \searrow 0} h(x) = h(0) = 0$, so (67) forces $h'(x) \rightarrow 0$. Thus by (65), h' is diff'able at the origin, and $\boxed{h'(0) = 0}$.

FTSOC, suppose $\{x \in J \mid h(x) \neq 0\}$ is non-void; let B be its infimum. By cty from the left, necessarily $h(B) = 0$. Therefore, replacing h by its translate $x \mapsto h(x - B)$, now

67: There are numbers $y > 0$, as small as one pleases, with $h(y) \neq 0$.

The Bound. I'll henceforth assume that $0 \leq h' \leq h$ on $[0, \infty)$; the hard-working Reader can put in the abs-value signs so as to make a complete proof.

Since $h'(0) = 0 < 1$, there exists a number $C > 0$ so, for each $x \in [0, C]$, that $0 \leq h(x) \leq x$. So we've shown exponent 1 to be *good*... where: A posint N is *good* if

67: for each $x \in [0, C]$, we have $0 \leq h(x) \leq x^N$.

Let's show that $[N \text{ good}] \implies [(N+1) \text{ good}]$. Fix an $x \in [0, C]$. Then by the Fund. Thm of Calculus,

$$\begin{aligned} h(x) &= h(x) - h(0) \stackrel{\text{FTC}}{=} \int_0^x h'(t) dt \\ &\leq \int_0^x t^N dt \\ &= \frac{1}{N+1} \cdot [x^{N+1} - 0^{N+1}], \end{aligned}$$

which is less-equal x^{N+1} .

Each posint is good, so (67) tells us that $h(x) = 0$ whenever $0 \leq x < \text{Min}(C, 1)$. But this offends (67). \diamond

End: Potential H-problem.

Weighted averages. Consider a point L in \mathbf{E} , a normed vectorspace. Given two vectors close to L , we seek a condition implying that all appropriate weighted-averages of these vectors are also close to L . For generality, we'll allow our weights, v_j , to be complex numbers. When applying (68'), below, we will typically send $\varepsilon \searrow 0$; hence the particular constant $2[1 + \mathcal{U}]$ is usually irrelevant.

68: Weighted-average lemma. Fix a bound $\mathcal{U} \in \mathbb{R}_+$ and $L \in \mathbf{E}$. Given $\varepsilon > 0$, suppose we have vectors $R_1, R_2 \in \mathbf{E}$

with each $\|R_j - L\| \leq \varepsilon$. Suppose we have (possibly complex) weights $v_1 + v_2 = 1$. Then

$$68': \quad \left\| [v_1 R_1 + v_2 R_2] - L \right\| \leq 2[1+\mathcal{U}] \cdot \varepsilon,$$

if, for at least one value of j , we have $|v_j| \leq \mathcal{U}$. \diamond

Proof. WLOG $|v_1| \leq \mathcal{U}$. So $|v_2| \leq 1+\mathcal{U}$, since $v_1 + v_2 = 1$. Thus each $|v_j| \leq C := 1+\mathcal{U}$.

Note that $L = v_1 L + v_2 L$. Thus LhS(68') is less-equal

$$|v_1| \cdot \|R_1 - L\| + |v_2| \cdot \|R_2 - L\| \stackrel{\text{note}}{\leq} C\varepsilon + C\varepsilon. \quad \blacklozenge$$

69: Deriv-sample lemma. Fix a normed-VS \mathbf{E} , an upper-bound $\mathcal{U} \in \mathbb{R}_+$ and a point $P \in \mathbb{R}$. Suppose $h: \mathbb{R} \rightarrow \mathbf{E}$ is differentiable at P .

Then, given ε there exists δ so that for each pair of distinct "sample points" y and z that are δ -close to P :

$$\dagger: \quad \frac{h(y)-h(z)}{y-z} \text{ is } \varepsilon\text{-close to } h'(P),$$

$$\text{as long as } \frac{\text{Min}(|y-P|, |z-P|)}{|y-z|} \leq \mathcal{U}. \quad \blacklozenge$$

Pf. Let $L := h'(P)$. WLOG, neither y nor z equals P . In light of (68'), let $\alpha := \frac{\varepsilon}{2[1+\mathcal{U}]}$ and take δ small enough that:

$$\text{If } x \in \text{PBal}_\delta(P) \text{ then } \left\| \frac{h(x)-h(P)}{x-P} - L \right\| < \alpha.$$

Setting $v_1 := \frac{y-P}{y-z}$ and $v_2 := \frac{P-z}{y-z}$, POFA^{♥10} informs us that

$$\ddagger: \quad \frac{h(y)-h(z)}{y-z} = v_1 \cdot \frac{h(y)-h(P)}{y-P} + v_2 \cdot \frac{h(z)-h(P)}{z-P}.$$

Since $v_1 + v_2 = 1$ and $\text{Min}(|v_1|, |v_2|) \leq \mathcal{U}$, the Weighted-average lemma applies. It insists that

$$\left\| \text{RhS}(\ddagger) - L \right\| \leq 2[1+\mathcal{U}] \cdot \alpha = \varepsilon.$$

Hence $\left\| \text{LhS}(\ddagger) - L \right\| \leq \varepsilon$, which is (\dagger). \blacklozenge

^{♥10}Plain Old-Fashioned Algebra.

vdW's no-where differentiable fnc. Let $J := [0, 1]$. We will define van der Waerden's fnc $\mathcal{W}: \mathbb{R} \rightarrow J$ and prove that it does not even have a *one-sided* derivative, anywhere.

Let $\varphi(\cdot)$ be the *distance-to-nearest-integer function*,^{♥11}

$$\varphi(x) := \text{Min}(x - \lfloor x \rfloor, \lceil x \rceil - x).$$

Its graph looks like $\cdots \wedge \wedge \wedge \wedge \wedge \wedge \wedge \wedge \wedge \wedge \cdots$.

For $n \in \mathbb{N}$, let $f_n(x) := \frac{1}{2^n} \varphi(2^n x)$. Thus

$$\text{SetOfZeros}(f_n) = \frac{1}{2^n} \cdot \mathbb{Z}.$$

Each f_n is continuous, since φ is. Hence each partial sum

$$g_K := \sum_{n \in [0..K]} f_n$$

is cts. Since $\|f_n\|_{\text{sup}} = 1/2^{n+1}$, and $\text{seq } n \mapsto 1/2^{n+1}$ is summable, sequence $(g_k)_{k=1}^\infty$ is $\|\cdot\|_{\text{sup}}$ -Cauchy. By the Uniform-convergence thm, then, \vec{g} converges uniformly to a continuous fnc

$$\mathcal{W} := \sum_{n=0}^\infty f_n.$$

To show its nondifferentiability, we will evaluate \mathcal{W} at *dyadic rationals*, elements of the set

$$\mathbb{D} := \left\{ \frac{\ell}{2^n} \mid \ell \in \mathbb{Z} \text{ and } n \in \mathbb{N} \right\}.$$

70: vdW-function thm. At each point $P \in \mathbb{R}$, van der Waerden's fnc \mathcal{W} has no onesided-derivative. \blacklozenge

Proof (Due to Patrick Billingsley). FTSOC, suppose $\mathcal{W}()$ has a righthand derivative at $P \in \mathbb{R}$. Fix posint K . Take the unique integer ℓ st.

$$*: \quad \frac{\ell-1}{2^K} \leq P < \overbrace{\frac{\ell}{2^K}}^y < \overbrace{\frac{\ell+1}{2^K}}^z.$$

For each $n \geq K$, note that $f_n(y) = 0 = f_n(z)$. Thus

$$**: \quad \frac{\mathcal{W}(z) - \mathcal{W}(y)}{z - y} = \sum_{n=0}^{K-1} \underbrace{\frac{f_n(z) - f_n(y)}{z - y}}_{s_n}.$$

But y and z are *consecutive* order- K dyadic rationals, and $n < K$, so each slope s_n , above, must be ± 1 .

^{♥11}There doesn't seem to be a std name for this beast. It is related to the *fractional part* fnc, $x - \lfloor x \rfloor$; and *that* name isn't great, since the "fractional part" need not be rational.

In (*), rename the y, z points to y_K, z_K . Equality (**) tells us that ratio

$$r_K := \frac{W(z_K) - W(y_K)}{z_K - y_K}$$

is a sum of K many instances of ± 1 . It follows that the difference $r_{K+1} - r_K$ is odd. Therefore sequence \vec{r} is not convergent.

This contradicts the Deriv-sample lemma, (69), which insists that \vec{r} converge to $W'(P)$. And the lemma indeed applies, with bound $\mathcal{U} := 1$, since (*) forces $\frac{|y-P|}{|y-z|} \leq 1$. ♦

Product rule. Here are several examples of bilinear maps:
 Multiplication, $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$. Scalar-vector-mult, $\mathbb{R} \times \mathbf{V} \rightarrow \mathbf{V}$.
 Inner-product, $\mathbf{V} \times \mathbf{V} \rightarrow \mathbb{R}$. Cross-product, $\mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$.
 Matrix-multiplication, $\text{MAT}(3, 5) \times \text{MAT}(5, 2) \rightarrow \text{MAT}(3, 2)$.

71: Product-rule thm. On normed VSes we have a bilinear map $\langle\langle \cdot \rangle\rangle: \mathbf{A} \times \mathbf{B} \rightarrow \mathbf{W}$ which is (jointly) continuous. Suppose maps $F: \mathbb{R} \rightarrow \mathbf{A}$ and $G: \mathbb{R} \rightarrow \mathbf{B}$ are differentiable at a point $p \in \mathbb{R}$. Then, for $t \in \mathbb{R}$, the map

$$\dagger: t \mapsto \langle\langle F(t), G(t) \rangle\rangle: \mathbb{R} \rightarrow \mathbf{W}$$

is differentiable at $t=p$. And its derivative, there, is

$$\ddagger: \langle\langle F'(p), G(p) \rangle\rangle + \langle\langle F(p), G'(p) \rangle\rangle. \quad \diamond$$

Pf. For brevity, use p^F for $F(p)$, etc. So $\langle\langle t^F, t^G \rangle\rangle - \langle\langle p^F, p^G \rangle\rangle$ equals

$$\begin{aligned} & \langle\langle t^F, t^G \rangle\rangle - \langle\langle p^F, t^G \rangle\rangle + \langle\langle p^F, t^G \rangle\rangle - \langle\langle p^F, p^G \rangle\rangle \\ &= \langle\langle t^F - p^F, t^G \rangle\rangle + \langle\langle p^F, t^G - p^G \rangle\rangle. \end{aligned}$$

Dividing both sides by $t - p$ gives

$$\left\langle\left\langle \frac{t^F - p^F}{t - p}, t^G \right\rangle\right\rangle + \left\langle\left\langle p^F, \frac{t^G - p^G}{t - p} \right\rangle\right\rangle.$$

Sending $t \rightarrow p$ gives (\ddagger), using cty of F, G and $\langle\langle \cdot, \cdot \rangle\rangle$. ♦

Total derivative

Consider a map $f: (\mathbf{V}, \|\cdot\|) \rightarrow (\mathbf{E}, \|\cdot\|)$ between normed VSes, Near a point $p \in \mathbf{V}$ we can try to approximate f with a linear map $L: \mathbf{V} \rightarrow \mathbf{E}$, by examining the **error term**,

$$72: \text{Err}_L(x) := [f(x+p) - f(p)] - L(x) \stackrel{\text{note}}{\in} \mathbf{E}.$$

Unsurprisingly, say that f is “differentiable at p ” if there exists such a linear map (evidently unique) for which

$$72': \frac{\text{Err}_L(x)}{\|x\|} \rightarrow \mathbf{0}_E, \quad \text{as } x \rightarrow \mathbf{0}_V.$$

Equivalently, in terms of the two norms,

$$72'': \forall \varepsilon > 0, \exists \delta > 0 \text{ st. } \forall x \in \mathbf{V}, \quad \begin{aligned} & \text{if } \|x\| \leq \delta \text{ then: } \frac{\|\text{Err}_L(x)\|}{\|x\|} \leq \varepsilon. \end{aligned}$$

73: Lemma. (Notation from above.) There is at most one linear map with zero-going error term. ♦

Pf. Contemplate two such linear approximators, L and M . Fixing a (WLOG non-zero) vector $\mathbf{v} \in \mathbf{V}$, our goal is to show that the difference vector, $\mathbf{d} := L(\mathbf{v}) - M(\mathbf{v})$, equals $\mathbf{0}_E$.

With $x := \alpha \mathbf{v}$, for a positive scalar α , linearity implies $L(x) - M(x) = \alpha \cdot \mathbf{d}$. Dividing by $\|x\| \stackrel{\text{note}}{=} \alpha \cdot \|\mathbf{v}\|$ yields

$$*: \frac{L(\alpha \mathbf{v}) - M(\alpha \mathbf{v})}{\|\alpha \mathbf{v}\|} = \frac{\mathbf{d}}{\|\mathbf{v}\|}.$$

Note $L(x) - M(x) = \text{Err}_M(x) - \text{Err}_L(x)$. Hence (72') implies, as $\alpha \searrow 0$, that LhS(*) $\rightarrow \mathbf{0}_E$. But RhS(*) doesn't change with α . Thus \mathbf{d} has secretly been $\mathbf{0}_E$ all along. ♦

Defn. Courtesy uniqueness, we call the linear L from (72'), “the **total derivative** of f at a point p ”. We write this L either as $\mathbf{D}_{p,f}$ or $\mathbf{D}_p[f]$ or $\mathbf{D}^f(p)$, depending on what we wish to emphasize. To evaluate this linear map at vector \mathbf{v} , we write $\mathbf{D}_{p,f}(\mathbf{v})$ or $\mathbf{D}_p[f](\mathbf{v})$ or $\mathbf{D}^f(p)(\mathbf{v})$.

Henceforth, we use $\|\cdot\|$ for the norm on all of our normed VSes.

Equality (74 \ddagger), stated further below, is written to resemble this “Calc 1” version of the Chain rule:

$$74\ddagger: [g \circ f]'(p) = g'(f(p)) \cdot f'(p). \quad \square$$

74: Basic derivative thm. Consider maps $f, \tilde{f}: \mathbf{V} \rightarrow \mathbf{E}$ and $g: \mathbf{E} \rightarrow \mathbf{W}$ between normed vectorspaces, a scalar α and a point $p \in \mathbf{V}$. Then

a: Differentiation is linear: $D_p[f + \tilde{f}] = D_p[f] + D_p[\tilde{f}]$
and $D_p[\alpha f] = \alpha D_p[f]$.

b: Chain rule:

$$74\ddagger: D^{g \circ f}(p) = D^g(f(p)) \circ D^f(p). \quad \diamond$$

In alternate notation: $D_p[g \circ f] = D_{f(p)}[g] \circ D_p[f]$.

Proof. Exer: Exercise. ♦

In finite dim'al spaces. In applications, often \mathbf{V} and \mathbf{E} have finite dimension; say K and N , respectively. Fixing ordered bases, each linear map $\mathbf{V} \rightarrow \mathbf{E}$ is represented by an $N \times K$ matrix. So formula (74 \ddagger) becomes

$$D^{g \circ f}(p) = D^g(f(p)) \bullet D^f(p).$$

where the “ \bullet ” is denoting matrix-multiplication. □

§B Riemann Integration

We employ the word “partition” (abbrev. “ptn”) in the specialized way^{♥12} it is used in RI.

Initially, we’ll discuss the 1-dimensional case, integrating over an interval $J := [a, b]$. A “**partition** P of J ” will be determined by a tuple of cutpoints

$$a = p_0 < p_1 < p_2 \dots < p_k < \dots < p_N = b.$$

Call the closed subinterval $B_k := [p_{k-1}, p_k]$, the “ k^{th} **block** of P ”. We’ll use P to also denote the *set* of P -blocks, e.g we might write $\sum_{B \in P} \text{Diam}(B) < 5$.

The **mesh(size)** of P is

$$\text{Mesh}(P) := \text{Max} \left\{ \text{Diam}(B_k) \mid k \in [1 .. N] \right\}.$$

$$75a: \quad \text{Use } \#P := \#\{\text{Set of } P\text{-blocks}\} \stackrel{\text{note}}{=} N \quad \text{and} \\ \text{CutPts}(P) := (p_0, p_1, \dots, p_N).$$

We say that ptn Q **refines** P , written $Q \succcurlyeq P$, if each P -block is a union of Q -blocks. Equivalently, in our 1-dim case, $\text{CutPts}(Q) \supset \text{CutPts}(P)$ [interpreted as sets, not tuples].

A pair of ptns $\{P, Q\}$ has a smallest common refinement

$$75b: \quad R := P \vee Q,$$

called “the **join** of P and Q ”, whose cutpoint set is $\text{CutPts}(P) \cup \text{CutPts}(Q)$.

Sample points. A **pointed partition** P (also called a “tagged ptn”) is a partition together with **tags** (x_1, \dots, x_N) , also called **sample points**, such that each $x_k \in B_k$. Use notation

$$75c: \quad \text{Tags}(P) := (x_1, \dots, x_N).$$

Given a function $f: J \rightarrow \mathbb{R}$, our **pptn** (“pointed partition”) gives a **Riemann sum**

$$75d: \quad \text{RS}^f(P) := \sum_{k=1}^N [f(x_k) \cdot \text{Size}(B_k)].$$

But wait?! Why a vague word like “size”? Well, in the 1-dim case, “size” will mean *length*, whereas for 2-dim integrals, “size” will mean *area*.

Treating the 1-dimensional integral, below, $\text{Size}(B_k)$ will mean the *unsigned*^{♥13} length $|p_k - p_{k-1}|$. From now

^{♥12}In set theory, a **partition** of a set Ω is a pairwise-disjoint collection, P , of Ω -subsets whose union, $\bigsqcup(P)$, is all of Ω . The elements of P are called “the **atoms** of P ”. Usually one assumes that the atoms of P are non-empty, and that there are only finitely many atoms in a partition.

^{♥13}Later, we will extend to integrating over an **oriented** interval, and then $\text{Size}(B_k)$ will mean the “signed length” $p_k - p_{k-1}$.

on, \widehat{B} to abbreviate $\text{Size}(B)$. In the 1-dim'al case, \widehat{B} equals $\text{Diam}(B)$. The general case simply needs

75e: $\forall \varepsilon > 0, \exists \delta > 0$ st. for each set B that can be the block of a ptn: If $\text{Diam}(B) < \delta$, then $\widehat{B} < \varepsilon$.

Analogous to $\text{Mesh}(\mathcal{P})$, define

$$\text{MaxSiz}(\mathcal{P}) := \text{Max} \{ \widehat{B} \mid B \in \mathcal{P} \} \quad \square$$

Standing convention: Henceforth, $J=[a, b]$ is a closed bounded positive-length interval. And $f: J \rightarrow \mathbb{R}$ is a function, not necessarily integrable.

Oscillation/Variation. The *f-variation* of a block B , is

75f: $\sup_{x,y \in B} [f(x) - f(y)]$. (Irrelevant whether we use brackets or absolute-values.)

Write this as $\text{Var}^f(B)$ or $\text{Var}(B)$. The quantity that we are really interested in is the *f-oscillation*¹⁴ of a block B :

75g: $\text{Osc}(B) = \text{Osc}^f(B) := \widehat{B} \cdot \text{Var}^f(B)$.

Define the “*f-oscillation* of a partition \mathcal{P} ” to be

75h: $\text{Osc}(\mathcal{P}) = \text{Osc}^f(\mathcal{P}) := \sum_{B \in \mathcal{P}} \text{Osc}^f(B)$.

Analogously, $\text{Var}^f(\mathcal{P}) := \sum_{B \in \text{Blks}(\mathcal{P})} \text{Var}^f(B)$. □

76: Osc lemma. Consider partitions $\mathcal{P}, \mathcal{Q}, \mathcal{R}$:

- ① If $\text{Osc}^f(\mathcal{P}) < \infty$ then $|f|$ is bnded.
- ② For $\mathcal{P} \leq \mathcal{Q}$ ptns: $\text{Osc}^f(\mathcal{P}) \geq |\text{RS}^f(\mathcal{P}) - \text{RS}^f(\mathcal{Q})|$.
- ③ If $\mathcal{Q} \leq \mathcal{R}$ then $\text{Osc}^f(\mathcal{Q}) \geq \text{Osc}^f(\mathcal{R})$. *Exer: Exercise.*
- ④ Suppose $U := \sup_{x \in J} |f(x)|$ is finite. Suppose we split one \mathcal{Q} -block C to get a partition \mathcal{R} , i.e $\#\mathcal{R} = 1 + \#\mathcal{Q}$ and $\mathcal{Q} \leq \mathcal{R}$. Then

$$\text{Osc}^f(\mathcal{Q}) \leq \text{Osc}^f(\mathcal{R}) + 2U \cdot \text{MaxSiz}(\mathcal{Q}).$$

When 1-dim'al, $\text{Osc}^f(\mathcal{Q}) \leq \text{Osc}^f(\mathcal{R}) + 2U \cdot \text{Mesh}(\mathcal{Q})$.[◇]

¹⁴So variation is average oscillation; it is oscillation-per-length.

Pf of ①. Were f unbded, then there is a \mathcal{P} -block B on which $f|_B$ is unbded; so $\text{Osc}^f(B)$ is already infinite. ◆

Pf of ②. Focus on some \mathcal{P} -block B and its tag x_B , and let

$$S = S_B := \sup_{x,y \in B} |f(x) - f(y)|.$$

This B equals a union of (consecutive) \mathcal{Q} -blocks, say

$$B = C_5 \cup C_6 \cup C_7 \cup C_8 \cup C_9,$$

which overlap only at their endpoints. Adding sizes,

$$\dagger: \quad \widehat{B} = \sum_{k=5}^9 \widehat{C}_k.$$

Use y_k for the \mathcal{Q} -tag of C_k ; so $|f(x_B) - f(y_k)| \leq S$, since $B \ni x_B, y_k$. Thus

$$-S \cdot \widehat{C}_k \leq f(x_B)\widehat{C}_k - f(y_k)\widehat{C}_k \leq S \cdot \widehat{C}_k.$$

Summing over k ,

$$\ddagger: \quad -S_B \cdot \widehat{B} \leq f(x_B)\widehat{B} - \sum_5^9 f(y_k)\widehat{C}_k \leq S_B \cdot \widehat{B}.$$

Summing this over all \mathcal{P} -blocks B yields the desired inequality that $-\text{Osc}(\mathcal{P}) \leq \text{RS}(\mathcal{P}) - \text{RS}(\mathcal{Q}) \leq \text{Osc}(\mathcal{P})$. ◆

Pf of ④. The largest value that $\text{Osc}^f(C)$ can assume is $[U - -U] \cdot \widehat{C}$, which is upper-bnded by $2U \cdot \text{MaxSiz}(\mathcal{Q})$. ◆

Riemann integral. We define the “proper” Riemann integral, which is only useful for bounded fncs. Later, we’ll extend to “improper” integrals.

A partition \mathcal{P} is “ δ -small” if $\text{Mesh}(\mathcal{P}) \leq \delta$. A function $f: J \rightarrow \mathbb{R}$ is (*Riemann*) *integrable*¹⁵, with integral $V \in \mathbb{R}$, if:

77a: $\forall \varepsilon > 0, \exists \delta > 0$ st. for each pointed-partition \mathcal{P} which is δ -small: $|\text{RS}^f(\mathcal{P}) - V| \leq \varepsilon$.

Trivially, if such a number V exists then it is unique. We may write this V as

$$\int_J f \quad \text{or} \quad \int_{[a,b]} f \quad \text{or} \quad \int_J f(t) dt.$$

We do not yet use symbol “ \int_a^b ”, since it presupposes an *orientation* of J , allowing us to distinguish \int_b^a from \int_a^b . □

77: Integrability-equivalence Lemma. *Integrability of f is equivalent to each of the following. The first is a kind of “Cauchy condition”.*

b: For each $\varepsilon > 0$, $\exists \delta > 0$ so that for each two δ -small pptns P and Q : $|\text{RS}^f(P) - \text{RS}^f(Q)| \leq \varepsilon$.

c: $\forall \varepsilon > 0, \exists \delta > 0$ so for each two pptns $P \leq Q$ with the P partition δ -small: $|\text{RS}^f(P) - \text{RS}^f(Q)| \leq \varepsilon$.

d: $\forall \varepsilon, \exists \delta$ st. $\forall \delta$ -small partitions P : $\text{Osc}^f(P) \leq \varepsilon$.

e: $\forall \varepsilon, \exists$ a partition P such that $\text{Osc}^f(P) \leq \varepsilon$. \diamond

Pf $[\exists V \in \mathbb{R}$ st. (77a)] \Leftrightarrow (b). Implication (\Rightarrow) is immediate, so we establish (\Leftarrow) . To this end, take a sequence $(Q_m)_{m=1}^\infty$ of pointed-ptns with $\lim_m \text{Mesh}(Q_m) = 0$.

Condition (b) implies that $m \mapsto \text{RS}^f(Q_m)$ is a Cauchy sequence of reals. Hence this limit exists:

$$V := \lim_{m \rightarrow \infty} \text{RS}^f(Q_m) \in \mathbb{R}.$$

Using (b) again shows that V fulfills (77a). \diamond

Pf (b) \Leftrightarrow (c). *Exer. 3: Prove (\Leftarrow) , the non-trivial direction.*

[Hint: Fix $\varepsilon > 0$. Take $\delta = \delta(\varepsilon/2)$ from (c). Given unrelated δ -small pptns P and Q , consider their join, $R := P \vee Q$. Now...] \diamond

Pf (c) \Leftrightarrow (d). Dir (\Leftarrow) follows from (76②), Osc lemma.

For (\Rightarrow) , let P and Q be the same *partition*, but let the tags of each vary over all possibilities. The supremum of $|\text{RS}^f(P) - \text{RS}^f(Q)|$ over all tags is precisely $\text{Osc}^f(P)$. \diamond

Pf (d) \Leftrightarrow (e). Since *some* partition has finite oscillation, (76①) says that our f is bounded; $\text{WLOG } 3 \geq |f|$.

To establish the non-trivial (\Leftarrow) , fix some ptn P with

$$\text{Osc}^f(P) \leq \frac{\varepsilon}{2}.$$

With $N := \#P$, let δ be $\overset{\heartsuit 16}{\frac{\varepsilon}{2}} / 6N$. Given a δ -small Q , the refinement $R := P \vee Q$ is obtained by splitting *fewer than* N blocks of Q . Applying (76④) at most N times yields

$$\text{Osc}(Q) \leq \text{Osc}(R) + N \cdot [2 \cdot 3 \cdot \delta] \leq \text{Osc}(R) + \frac{\varepsilon}{2}.$$

And (76③) courteously gives $\text{Osc}(R) \leq \text{Osc}(P) \leq \frac{\varepsilon}{2}$. Hence $\text{Osc}(Q) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, as requested. \diamond

^{\heartsuit 16}When J is not 1-dim'al, there is an extra step. We pick δ small enough that every δ -small partition Q has $\text{MaxSiz}(Q) \leq \frac{\varepsilon}{2} / 6N$.

78: Basic RI Thm. *(For improper integrals, this needs to be altered.) Consider interval $J := [a, b]$ and fnc $f: J \rightarrow \mathbb{R}$. Then*

i: *If f continuous, then f is integrable.*

ii: *If f monotonic, then f is integrable. (For discontinuous R-Stieltjes integrators, this is false.)* \diamond

Proof of (i). Fix an $\varepsilon > 0$. Since f is uniformly-cts (being continuous on a compact set) there is a $\delta > 0$ such that

$$\forall x, y \in J: |x - y| \leq \delta \implies |f(x) - f(y)| \leq \varepsilon / \widehat{J}.$$

This implies that $\text{Osc}^f(P) \leq \varepsilon$, whenever P is a δ -small partition. Hence (77d). \diamond

Proof of (ii). We use (77e). WLOG $J = [0, 1]$. WLOG, f is increasing (i.e. non-decr). For each subinterval $B := [x, y]$, then, $\text{Var}^f(B) = f(y) - f(x)$.

Given posint N , let partition P_N cut J into N equal-length blocks, with j^{th} -block $B_j := [\frac{j-1}{N}, \frac{j}{N}]$. So

$$\begin{aligned} \text{Osc}^f(P_N) &= \sum_{j=1}^N \frac{1}{N} \cdot \text{Var}^f(B_j) \\ &= \frac{1}{N} \cdot \sum_{j=1}^N \left[f\left(\frac{j}{N}\right) - f\left(\frac{j-1}{N}\right) \right] \\ &= \frac{1}{N} \cdot [f(1) - f(0)]. \end{aligned}$$

And this latter goes to zero, as $N \nearrow \infty$. \diamond

Closure properties of RI

Let $\text{RI}(J \rightarrow \mathbb{R})$ denote the *set* of Riemann-integrable functions $J \rightarrow \mathbb{R}$.

As an example of *non-integrability*, let h be $\mathbf{1}_Q$, but restricted to J . For each partition P , then, $\text{Osc}^h(P) = 1 \cdot \widehat{J}$. So h is a non-RI fnc with the peculiar property that $h \circ h$ is integrable, since $h \circ h \equiv 1$.

79: Integration-is-Linear lemma. **W** := $\text{RI}(J \rightarrow \mathbb{R})$ is an \mathbb{R} -vectorspace. The map $[h \mapsto \int_J h]$, from $\mathbf{W} \rightarrow \mathbb{R}$, is a positive (non-negative) \mathbb{R} -linear-functional. Consequently, for integrable f and g : $[f \geq g] \implies [\int_J f \geq \int_J g]$. \diamond

Proof. That \mathbf{W} is a VS follows from observing that

$$\begin{aligned} \text{RS}^{5f}(Q) &= 5 \cdot \text{RS}^f(Q) \quad \text{and} \\ \text{RS}^{f+g}(Q) &= \text{RS}^f(Q) + \text{RS}^g(Q), \end{aligned}$$

for arbitrary fncs $f, g: J \rightarrow \mathbb{R}$, scalar $5 \in \mathbb{R}$ and pptn \mathbf{Q} . This also shows (exercise) that $[h \mapsto \int_J h]$ is a linear functional.

When $h \in \text{RI}$ is non-negative, then $\text{RS}^h(\mathbf{Q}) \geq 0$ for each pptn \mathbf{Q} ; so $\int_J h \geq 0$. Now apply this to $h := f - g$. \blacklozenge

Pos/Neg parts. We define the positive/negative parts of a function. The “**positive part** of f ” is

80.1: $f^+ := \text{Max}(f, \mathbf{0})$, ie. $f^+(x) = \text{Max}(f(x), 0)$. And $f^- := -\text{Min}(f, \mathbf{0})$ is the **negative part** of f .

Easily, each of f^+ and f^- is non-negative, and

80.2:
$$\begin{aligned} f^+ + f^- &= |f| \quad \text{and} \\ f^+ - f^- &= f. \end{aligned}$$

For a pair of functions, one verifies that

80.3:
$$\begin{aligned} \text{Max}(f, g) &= [f+g + |f-g|]/2 \quad \text{and} \\ \text{Min}(f, g) &= [f+g - |f-g|]/2. \end{aligned} \quad \square$$

80: AbsValue RI Thm. Suppose $f, g: J \rightarrow \mathbb{R}$ are integrable. Then each of $f^+, f^-, |f|, \text{Max}(f, g)$ and $\text{Min}(f, g)$ is integrable. Finally

80*:
$$\left| \int_J f \right| \leq \int_J |f|. \quad \blacklozenge$$

Proof. The f^+ -oscillation of each partition \mathbf{P} is upper-bounded by its f -oscillation; so f^+ is RI, by (77d). Ditto f^- is RI; hence so is $|f|$, their sum. Consequently, functions (80.3) are integrable.

For (80*), note $\int f = \int f^+ - \int f^-$. By the Δ nequality,

$$\begin{aligned} \left| \int f \right| &\leq \left| \int f^+ \right| + \left| \int f^- \right| \\ &= \int f^+ + \int f^- = \int |f|. \end{aligned} \quad \blacklozenge$$

81: Product-RI Thm. If $f, g \in \text{RI}(J \rightarrow \mathbb{R})$. then $f \cdot g \in \text{RI}$. \blacklozenge

Pf. WLOG $|f| \leq 2$ and $|g| \leq 3$. Let $h := f \cdot g$. ISTE establish

$$\text{Osc}^h(\mathbf{P}) \stackrel{?}{\leq} 3 \cdot \text{Osc}^f(\mathbf{P}) + 6 \cdot \text{Osc}^g(\mathbf{P})$$

for each partition \mathbf{P} . Fixing \mathbf{P} and a \mathbf{P} -block B , our goal is

81a:
$$\text{Osc}^h(B) \stackrel{?}{\leq} 3 \cdot \text{Osc}^f(B) + 6 \cdot \text{Osc}^g(B)$$

Fix pts $x, y \in B$. Define numbers Φ, v, Γ, w by $\Phi := f(x)$, $\Phi + v := f(y)$, $\Gamma := g(x)$ and $\Gamma + w := g(y)$. Subtracting,

$$\begin{aligned} h(y) - h(x) &= [\Phi + v][\Gamma + w] - \Phi\Gamma \\ &= v\Gamma + \Phi w + vw \\ &\leq |v| \cdot 3 + 2|w| + [2 - -2] \cdot |w| \\ &= 3|v| + 6|w| \\ &\leq 3 \cdot \text{Var}^f(B) + 6 \cdot \text{Var}^g(B). \end{aligned}$$

Multiply by \widehat{B} , then take $\sup_{x, y \in B}$ to obtain (81a). \blacklozenge

Exer. 7: Prove: Suppose $f \in \text{RI}(J \rightarrow \mathbb{R})$ and $L > 0$, where $L := \left[\inf_{x \in J} |f(x)| \right]$. Prove that $1/f$ is integrable.

Exer. 4: Dis/Prove: Suppose $f, g: [0, 1] \rightarrow [0, 1]$ are (Riemann) integrable fncs. Then $h := g \circ f$ is integrable.

False. Let f be the ruler-function $\mathcal{R}_{\mathbb{Q}}$. So $f(\frac{p}{q}) := \frac{1}{q}$, when $p \perp q$ are integers with $q > 0$. And $f(\text{irrational})$ is 0.

Let $g := \mathbf{1}_{(0,1]}$. Then $g \circ f$ is the indicator-fnc of the rationals; this is not Riemann-integrable.

In contrast, the reverse composition $f \circ g$ is RI, indeed continuous. Indeed, $f \circ g$ is the constant-1 function. \blacklozenge

Exer. 5: (Does t.fol hold for R-Stieltjes integration?)

Dis/Prove: On compact sets $K, J \subset \mathbb{R}$, with J an interval, we have an integrable $f: J \rightarrow K$ function and continuous $g: K \rightarrow \mathbb{R}$. Then $h := g \circ f$ is integrable.

True. WLOG $|g| \leq 3$. WLOG $\widehat{J} \leq 2$. I'll use “partition” to mean a partition of J .

Fix $\eta > 0$. I will produce $\delta > 0$ st. for each δ -small ptn \mathbf{P} :

‡a:
$$\text{Osc}^{g \circ f}(\mathbf{P}) \stackrel{?}{\leq} 8\eta.$$

Bad blocks. The uniform-continuity of g produces an $(\varepsilon \leq \eta)$ such that

‡b: $\forall a, b \in K: |a - b| \leq \varepsilon \implies |g(a) - g(b)| \leq \eta.$

Since $f \in \text{RI}$, we can take δ so small that each δ -small partition \mathbf{P} has

‡c:
$$\text{Osc}^f(\mathbf{P}) \leq \varepsilon^2.$$

Define the set of “good” blocks

‡d:
$$\mathcal{G} := \{B \in \text{Blks}(\mathbf{P}) \mid \text{Var}^f(B) < \varepsilon\}.$$

Define the “bad” blocks $\mathcal{B} := \text{Blks}(\mathcal{P}) \setminus \mathcal{G}$. From $(\ddagger c)$ and $(\ddagger d)$,

$$\begin{aligned} \varepsilon^2 &\geq \sum_{B \in \mathcal{B}} \text{Osc}^f(B) \\ &\geq \sum_{B \in \mathcal{B}} \varepsilon \cdot \widehat{B} = \varepsilon \cdot \widehat{\mathcal{B}}. \end{aligned}$$

Dividing by ε yields $\widehat{\varepsilon \geq \widehat{\mathcal{B}}}$. For each block B ,

$$\text{Osc}^{g \circ f}(B) \leq [3 - 3] \cdot \widehat{B} = 6 \cdot \widehat{B},$$

by our bound on $|g|$. Summing over the bad blocks,

$$\ddagger e: \quad \text{Osc}^{g \circ f}(\mathcal{B}) \leq 6 \cdot \widehat{\mathcal{B}} \leq 6\varepsilon \leq 6\eta.$$

Good blocks. Fix $B \in \mathcal{G}$ and $x, y \in B$. By $(\ddagger d)$, then $(\ddagger b)$, the oscillation $\text{Osc}^{g \circ f}(B) \leq \eta \cdot \widehat{B}$. Summing over good blocks,

$$\text{Osc}^{g \circ f}(\mathcal{G}) \leq \eta \cdot \widehat{\mathcal{G}} \leq \eta \cdot \widehat{J} = 2\eta.$$

Adding this to the $(\ddagger e)$ inequality, yields $(\ddagger a)$. ♦

Exer. 6: Dis/Prove: On compact intervals $K, J \subset \mathbb{R}$, we have a continuous $f: J \rightarrow K$ and an integrable $g: K \rightarrow \mathbb{R}$. Then $h := g \circ f$ is integrable.

82: Closure-RI Thm. Fix an integrable $f: J \rightarrow \mathbb{R}$. Then for each closed subinterval $I \subset J$, the restriction $f|_I$ is integrable.

Conversely, consider a fnc $g: J \rightarrow \mathbb{R}$ and a point $y \in J$. If g is integrable on $[a, y]$ and on $[y, b]$, then g is integrable. ♦

Proof. Fix $\varepsilon > 0$ and take δ from (77d) applied to f on J . Given a δ -small ptn \mathcal{P} of I , extend this \mathcal{P} to create a δ -small ptn \mathcal{P}' of J . Thus $\text{Osc}^f(\mathcal{P}) \leq \text{Osc}^f(\mathcal{P}') \leq \varepsilon$.

Conversely, fixing ε there exist ptns \mathcal{Q} of $[a, y]$ and \mathcal{R} of $[y, b]$ each with oscillation less than $\frac{\varepsilon}{2}$. Glue them together to get a ptn of J with oscillation less than ε . ♦

Oriented integral. We may write an integral on $J = [a, b]$ as

$$\int_J f \quad \text{or} \quad \int_{[a,b]} f \quad \text{or} \quad \int_a^b f \quad \text{or} \quad \int_a^b f(t) dt.$$

Reversing the “limits of integration”, define

$$\int_b^a f := - \int_{[a,b]} f.$$

So our 1-dim’al integral is an *oriented integral*. □

83: Lemma. For $a, b, c \in \mathbb{R}$, and function f :

$$\int_a^c f = \left[\int_a^b f \right] + \left[\int_b^c f \right],$$

as soon as f is integrable on the interval from $\text{Min}(a, b, c)$ to $\text{Max}(a, b, c)$. **Proof.** Exer: ♦

The Fundamental Theorem of Calculus

For an integrable (not-necessarily cts) function $f: J \rightarrow \mathbb{R}$, recall that $\mathcal{U} := \sup_{t \in J} |f(t)|$ is finite. And

$$\varphi(x) := \int_{[a,x]} f, \quad \text{as a map } \varphi: J \rightarrow \mathbb{R},$$

is well-defined, thanks to (82). This φ is sometimes called an *antiderivative* of f .

84: FTC. With $f(), \mathcal{U}, \varphi()$ from above: This φ is Lipschitz continuous, with \mathcal{U} a Lipschitz bound. Moreover, at each f -continuity point $z \in J$, our φ is differentiable and

$$84a: \quad \varphi'(z) = f(z).$$

Conversely, each fnc $\psi \in \mathcal{C}^1(J \rightarrow \mathbb{R})$ has $\psi' \in \text{RI}$ and

$$84b: \quad \int_{[a,b]} \psi' = \psi(b) - \psi(a). \quad \diamond$$

Pf of (84a). For $x < y$ in J , note, $\varphi(y) - \varphi(x) = \int_{[x,y]} f$. So

$$|\varphi(y) - \varphi(x)| = \left| \int_{[x,y]} f \right| \leq \int_{[x,y]} |f| \leq [y - x] \cdot \mathcal{U},$$

by (80*) and (79). Hence φ is \mathcal{U} -Lipschitz.

At an f -continuity-point z . WELLOG, z is not an endpoint of J . WLOG $f(z) = 4$. Fixing an $\varepsilon > 0$, the continuity of f at z asserts an open interval $I \ni z$ st.

$$\ddagger: \quad 4 - \varepsilon \leq f|_I \leq 4 + \varepsilon.$$

Consider a small non-zero “bump” $\beta \in \mathbb{R}$ with $z + \beta \in I$. WELLOG $\beta > 0$; let $B := [z, z + \beta]$. Courtesy (79), integrating (\ddagger) over B yields, since \widehat{B} equals β , that

$$\ddagger: \quad [4 - \varepsilon] \cdot \beta \leq \int_B f \leq [4 + \varepsilon] \cdot \beta.$$

But the integral equals $\varphi(z + \beta) - \varphi(z)$. Thus the difference-quotient satisfies

$$4 - \varepsilon \leq \frac{\varphi(z + \beta) - \varphi(z)}{\beta} \leq 4 + \varepsilon.$$

This holds for every small-enough non-zero β . Thus φ is differentiable at z , and $\varphi'(z) = 4 = f(z)$. ♦

Pf of (84b). Firstly, since ψ' is cts, it is RI. Define φ by

$$\varphi(x) := \int_{[a,x]} \psi', \quad \text{as a map } \varphi: J \rightarrow \mathbb{R},$$

Thus $\int_J \psi' \stackrel{\text{def}}{=} \varphi(b) - \varphi(a) = \varphi(b) - \varphi(a)$, since $\varphi(a) = 0$.

By (84a), $\varphi' = \psi'$. This means (**Exer: By what thm?**) that $\psi - \varphi$ is a constant-fnc. Thus $\psi(b) - \psi(a)$ equals $\varphi(b) - \varphi(a)$, which equals $\int_J \psi'$. \blacklozenge

Measuring the size of sets

Fix a metric space X and a way of measuring the size of open balls; we'll use "ball" to mean "non-empty open ball". At a "center" $c \in X$, we use $\text{Bal}_r(c)$ for the (open) ball of radius r .

Fix μ , a "measure on open balls": For each center $c \in X$ and radius $r \in \mathbb{R}_+$, this μ assigns a "mass"

$$85a: \quad \mu(\text{Bal}_r(c)) \in [0, \infty).$$

Henceforth, in this section, let **cover** mean a cover by open balls. For a set $K \subset X$, let "C is a K-cover" mean that each $B \in \mathcal{C}$ is an open ball, and $\bigcup(\mathcal{C}) \supset K$. Agree to use $\mu(\mathcal{C})$ to mean

$$\mu(\mathcal{C}) := \sum_{B \in \mathcal{C}} \mu(B).$$

Defining two measures. To measure a set $E \subset X$, we let \mathcal{C} vary over all covers of E ; *finite* covers for **Jordan mass**, $\mathcal{J}()$, and *countable* covers for **Lebesgue mass**, $\lambda()$:

$$85b: \quad \mathcal{J}(E) := \inf_{\mathcal{C} \text{ finite}} \mu(\mathcal{C}). \quad \lambda(E) := \inf_{\mathcal{C} \text{ countable}} \mu(\mathcal{C}).$$

We impose the following requirements on μ .

M1: Each ball B is \mathcal{J} -measurable and λ -measurable, and

$$\mathcal{J}(B) = \lambda(B) = \mu(B).$$

M2: For each $c \in X$: $\lim_{r \searrow 0} \mu(\text{Bal}_r(c)) = 0$.

Occasionally we will want some of these conditions.

M3: The function $r \mapsto \mu(\text{Bal}_r(c))$ is continuous.

M4: The function $r \mapsto \mu(\text{Bal}_r(c))$ is strictly increasing. \square

86: Basic measure lemma. For all sets $A, B, E \in \mathcal{P}(X)$:

$$i: \lambda(E) \leq \mathcal{J}(E).$$

ii: If $A \subset B$, then $\mathcal{J}(A) \leq \mathcal{J}(B)$ and $\lambda(A) \leq \lambda(B)$.

iii: $\mathcal{J}(\emptyset) = 0 = \lambda(\emptyset)$.

iv: If $A_1 \cup A_2 \supset B$ then $\mathcal{J}(A_1) + \mathcal{J}(A_2) \geq \mathcal{J}(B)$. Ditto for $\lambda()$.

v: If $[\bigcup_1^\infty A_n] \supset B$ then $[\sum_1^\infty \lambda(A_n)] \geq \lambda(B)$. \blacklozenge

Remark. In contrast to λ , Jordan-measure is not countably-subadditive: Enumerate $Q := \mathbb{Q} \cap [0, 1]$, and let A_n comprise the first n rationals in Q . Then $\sum_{n=1}^\infty \mathcal{J}(A_n) = 0$, but $\mathcal{J}(Q) = 1$. \square

87: Prop'n. Consider $E, K \in \mathcal{P}(X)$, with K compact. Then

a: $\mathcal{J}(K) < \infty$. (**Exer: .**)

b: $\mathcal{J}(K) = \lambda(K)$.

c: Suppose (M3). Then $\mathcal{J}(\text{Cl}(E)) = \mathcal{J}(E)$. (This fails for Jordan-measure replaced by $\lambda()$: Let $X := \mathbb{R}$ and $E := \mathbb{Q}$.) \blacklozenge

Pf of (b). Fix $\varepsilon > 0$. Since $\lambda(K) < \infty$, we can find a countable K -cover \mathcal{C} with $\mu(\mathcal{C}) \leq \varepsilon + \lambda(K)$. The compactness of K asserts a finite subcover $\mathcal{F} \subset \mathcal{C}$. Thus

$$\mathcal{J}(K) \leq \mu(\mathcal{F}) \leq \mu(\mathcal{C}) \leq \varepsilon + \lambda(K).$$

For each ε this holds, so $\mathcal{J}(K) \leq \lambda(K)$. \blacklozenge

Pf of (c). Fix $\varepsilon > 0$. WLOG $\mathcal{J}(E) < \infty$, so there is a finite E -cover $\{\text{Bal}_{r_j}(c_j)\}_{j=1}^N$, with

$$\ddagger: \quad \sum_{j=1}^N \mu(\text{Bal}_{r_j}(c_j)) \leq 2\varepsilon + \mathcal{J}(E).$$

Take a posreal δ sufficiently small that for each $j \in [1..N]$,

$$\ddot{:} \quad \mu(\text{Bal}_{\delta+r_j}(c_j)) - \mu(\text{Bal}_{r_j}(c_j)) \leq \varepsilon/N;$$

possibly, since there are only *finitely* many balls under consideration, and each map $r \mapsto \mu(\text{Bal}_r(c))$ is cts.

Automatically, collection $\{\text{Bal}_{\delta+r_j}(c_j)\}_{j=1}^N$ covers

$$*: \quad \text{Bal}_\delta\left(\bigcup_{j=1}^N \text{Bal}_{r_j}(c_j)\right) \stackrel{\text{note}}{\supset} \text{Bal}_\delta(E) \stackrel{\text{note}}{\supset} \text{Cl}(E).$$

Inequalities ($\ddot{:}$) and (\ddagger) justify

$$\begin{aligned} \sum_1^N \mu(\text{Bal}_{\delta+r_j}(c_j)) &\leq \varepsilon + \sum_1^N \mu(\text{Bal}_{r_j}(c_j)) \\ &\leq 3\varepsilon + \mathcal{J}(E). \end{aligned}$$

From (*), then, $\mathcal{J}(\text{Cl}(E)) \leq 3\varepsilon + \mathcal{J}(E)$. Now send $\varepsilon \searrow 0$. \blacklozenge

A condition for Integrability

Let's examine the discontinuity set of an $f: J \rightarrow \mathbb{R}$.

Fixing $\varepsilon > 0$, define two sets $C, K \subset J$ to comprise those $x \in J$ such that for each posreal δ :

88a: For C : $\exists y \in \text{Bal}_\delta(x)$ with $|f(y) - f(x)| \geq \varepsilon$.

88b: For K : $\exists y_1, y_2 \in \text{Bal}_\delta(x)$ with $|f(y_1) - f(y_2)| \geq \varepsilon$.

Both C and K are ε -approximations to $\text{DisCty}(f)$. While C is simpler to describe, it need not be closed.¹⁷ In contrast, K is closed (Exer:), hence compact.

Rematerializing the ε , easily $C_\varepsilon \subset K_\varepsilon$ and, by the Δ inequality, $K_{2\varepsilon} \subset C_\varepsilon$. Redefining, let C_n denote (88a) where $\varepsilon := \frac{1}{n}$. Make the analogous defn for K_n . Thus

$$C_1 \subset C_2 \subset \dots \quad \text{and} \quad K_1 \subset K_2 \subset \dots$$

88c: $\text{and } C_n \subset K_n \subset C_{2n}$.

$$\text{Thus } \bigcup_{n=1}^{\infty} K_n = \bigcup_{n=1}^{\infty} C_n \stackrel{\text{def}}{=} \text{DisCty}(f).$$

Complexity of sets. We need names for two types of sets. A subset E of X is said to be a " \mathcal{G}_δ -set" if it can be written as a countable intersection of open sets. A subset is an " \mathcal{F}_σ -set" if it equals some countable union of closed sets.¹⁸ On a topological space X ,

A decomposition $A \sqcup B = X$ has: $A \in \mathcal{F}_\sigma \Leftrightarrow B \in \mathcal{G}_\delta$.

The last line of (88c) shows the following.

88d: For a function $f: X \rightarrow Y$ between two metric spaces, its discontinuity set is always an \mathcal{F}_σ , and $\text{Cty}(f)$ is always a \mathcal{G}_δ .

Staying in metric spaces, here is a nice exercise:

Exer. 8: Suppose $K \subset X^{\text{MS}}$ is closed. Then K is $X\text{-}\mathcal{G}_\delta$. More generally, $\mathcal{G}_\delta \cap \mathcal{F}_\sigma \supset \text{CLD}(X) \cup \text{OPN}(X)$.

Alas, this can fail in general topological spaces. □

¹⁷To make an example, let $S := [3, 5] \cap \mathbb{Q}$. Define $f: \mathbb{R} \rightarrow \mathbb{R}$ so that f is the indicator-fnc $\mathbf{1}_S$ **except that** $f(5) := 1/2$. Then for $\varepsilon := 1$, the corresponding C_ε set is the half-open $[3, 5)$, which is neither open nor closed. Yet $K_\varepsilon = [3, 5]$, which is closed.

This f has closed discty set, since $\text{DisCty}(f) = K_\varepsilon$, for $\varepsilon=1$. As a contrasting example, $\text{DisCty}(\text{Ruler}_\mathbb{Q})$ is \mathbb{Q} , which is neither open nor closed. But \mathbb{Q} is indeed an \mathcal{F}_σ -set.

¹⁸The "F" is from the French word *fermé*, "closed", and the "σ" is from the German word *Summe*, sum, here meaning "union".

The "G" is from *Gebiet* (German, "area"), here meaning "open set". And the "δ" is from the German *Durchschnitt*, meaning intersection.

89: Integrability Theorem. On interval $J = [a, b]$, consider a subset $S \subset J$.

a: The map $\mathbf{1}_S: J \rightarrow \mathbb{R}$ is Riemann-integrable IFF $\mathcal{J}(\partial_J(S)) = 0$.

[Recall exercise (46P.13) that $\partial(S) = \text{DisCty}(\mathbf{1}_S)$.]

b: A function $f: J \rightarrow \mathbb{R}$ is RI IFF f is bounded and $\lambda(\text{DisCty}(f)) = 0$. ♦

Pf of (a). The discontinuity set of an indicator-fnc is closed; hence is compact, since J is. Thus its Jordan-mass equals its Lebesgue-mass. So (a) is implied by (b). ♦

Direct proof of (a(⇒)). Fix $\varepsilon > 0$. Take a ptn P with $\text{Osc}^{\mathbf{1}_S}(P) \leq \varepsilon$. Let \mathcal{G} comprise those "good" P -blks B on which $\mathbf{1}_S$ is constant. Use \mathcal{B} be the remaining "bad" blocks. By defn,

$$\varepsilon \geq \text{Osc}^{\mathbf{1}_S}(P) = 0 \cdot \widehat{\mathcal{G}} + 1 \cdot \widehat{\mathcal{B}}.$$

I.e., $\widehat{\mathcal{B}} \leq \varepsilon$. But $\bigcup(\mathcal{B}) \supset \text{DisCty}(\mathbf{1}_S)$; so $\varepsilon \geq \mathcal{J}(\partial(S))$. ♦

Pf of (b(⇒)). Earlier work shows that f is bounded.

Write $\Delta := \text{DisCty}(f)$ as $\bigcup_{n=1}^{\infty} C_n$, from (88c): A point x is in C_n IFF there is a seq \vec{y} converging to x , so that each point $y \in \vec{y}$ has $|f(y) - f(x)| \geq \frac{1}{n}$.

ISTShow, for each n , that $\lambda(C_n) = 0$, since then the countable subadditivity of (86v) shows that Δ is a nullset.

Fix n , let $C := C_n$ and $\varepsilon := \frac{1}{n}$. Fix an arbitrary $\delta > 0$ and take a ptn P with $\text{Osc}(P) < \varepsilon\delta$. A P -block B is "bad" if $\text{Var}(B) \geq \varepsilon$. So $\text{Osc}(B) \geq \varepsilon \widehat{B}$. Summing over the bads,

$$\varepsilon\delta \geq \text{Osc}(P) \geq \sum_{B \text{ bad}} \varepsilon \cdot \widehat{B}.$$

Thus $\delta \geq \lambda(L)$, where $L := \bigcup_{B \text{ bad}} B$.

Consider a point $z \in C$. If z is in the interior of a block B , then automatically B is bad, so $z \in L$. Thus

$$C \subset \text{CutPts}(P) \cup L.$$

Hence $\lambda(C) \leq 0 + \delta = \delta$. This holds for all $\delta > 0$, so C is a Lebesgue nullset. ♦

Brillo's proof of (b(⇐)). WLOG $|f| \leq \frac{3}{2}$. WLOG $\widehat{J} = 1$. Write $J = \Delta \sqcup \Gamma$, where $\Delta := \text{DisCty}(f)$ and $\Gamma := \text{Cty}(f)$.

Fix $\varepsilon > 0$; we'll produce a partition P with

$$\forall: \quad \text{Osc}^f(P) \leq 8\varepsilon.$$

Since $\lambda(\Delta) = 0$, there exists a countable cover \mathcal{D} of Δ with $\mu(\mathcal{D}) \leq \varepsilon$.

For each $z \in \Gamma$, there an open interval $I_z \ni z$ with $\text{Var}^f(I_z) \leq 5\varepsilon$. Thus $\mathcal{C} := \{I_z\}_{z \in \Gamma}$ covers Γ and so the union $\mathcal{U} := \mathcal{D} \cup \mathcal{C}$ is an open cover of J . Hence \mathcal{U} has a Lebesgue number δ ; this, since J is compact.

Consider a ptn P with $\text{Mesh}(P) < \delta$; necessarily, each P -block lies inside of some \mathcal{U} -patch. Call P -block B "good" if there exists a patch $I \in \mathcal{C}$ with $I \supset B$. So

$$\sum_{B \text{ good}} \text{Osc}^f(B) \leq 5\varepsilon \cdot \widehat{J} = 5\varepsilon.$$

Each "bad" B is covered by some \mathcal{D} -patch. Thus

$$\sum_{B \text{ bad}} \text{Osc}^f(B) \leq \text{Var}^f(J) \cdot \sum_{U \in \mathcal{D}} \widehat{U} \leq \left[\frac{3}{2} - \frac{-3}{2}\right] \cdot \varepsilon = 3\varepsilon.$$

Adding these together yields that $\text{Osc}^f(P) \leq 8\varepsilon$. ♦

Interchange of limit-operations

An exercise that could have been stated earlier.

90: Obs. Take fncs $g, f_n: X^{\text{Set}} \rightarrow \Omega^{\text{MS}}$ with each f_n bounded. If $f_n \xrightarrow{\text{uniformly}} g$, then g is bounded. **Pf. Exer:** ♦

More interestingly.

91: Prop'n. Suppose $b, f: J \rightarrow \mathbb{R}$ are bounded functions; set $\varepsilon := \|b\|_{\text{sup}}$. Then for each partition P ,

$$|\text{Osc}^{f+b}(P) - \text{Osc}^f(P)| \leq 2\varepsilon \cdot \widehat{J}. \quad (\text{Exer :}) \quad \diamond$$

92: Integral-Convergence Theorem. Consider functions $g, f_n: J \rightarrow \mathbb{R}$, with $f_n \in \text{RI}$. Suppose $f_n \xrightarrow[n \rightarrow \infty]{\text{uniformly}} g$. Then g is RI. Moreover, $\int_J f_n \rightarrow \int_J g$, as $n \nearrow \infty$. ♦

Pf. Take $\varepsilon > 0$ and take n large enough that $\|g - F\|_{\text{sup}} \leq \varepsilon$, where $F := f_n$. Now take a ptn P st. $\text{Osc}^f(P) \leq \varepsilon$. By (91),

$$|\text{Osc}^g(P) - \text{Osc}^F(P)| \leq 2\varepsilon \cdot \widehat{J}.$$

So $\text{Osc}^g(P) \leq \text{Osc}^F(P) + 2\varepsilon \cdot \widehat{J} \leq [1 + 2 \cdot \widehat{J}] \varepsilon$. This holds for each ε , so (77e) tells us that g is integrable.

Being integrable, we can replace f_n by $f_n - g$, and replace g by $g - g$, to say WLOG $f_n \xrightarrow[n \rightarrow \infty]{\text{unif}} \mathbf{0}$. But

$$\left| \int_J f_n \right| \leq \|f_n\|_{\text{sup}} \cdot \widehat{J},$$

so $\left[\int_J f_n \right] \xrightarrow[n \rightarrow \infty]{} 0$. And, indeed, $0 = \int_J \mathbf{0}$. ♦

93: DUC Thm (Derivative uniform-convergence). We have functions $f_n \in \mathbf{C}^1(J \rightarrow \mathbb{R})$ whose derivative-sequence $(f'_n)_1^\infty$ is sup-norm Cauchy. Thus function

$$93a: \quad \Delta := \text{unif-lim}_{n \rightarrow \infty} f'_n$$

exists. Suppose there is a point $A \in J$ such that

$$93b: \quad \lim_{n \rightarrow \infty} f_n(A) \text{ exists in } \mathbb{R}.$$

Then for each x , the limit $g(x) := \lim_{n \rightarrow \infty} f_n(x)$ exists in \mathbb{R} . Moreover, g is differentiable and $g' = \Delta$. ♦

Proof. Fix an $x \in J$. Thanks to (92), and FTC applied to each f_n ,

$$\int_A^x \Delta \stackrel{\text{by (92)}}{=} \lim_{n \rightarrow \infty} \int_A^x f'_n = \lim_{n \rightarrow \infty} [f_n(x) - f_n(A)].$$

So (93b) tells us that $\lim_{n \rightarrow \infty} f_n(x)$ exists in \mathbb{R} .

Restating, the map $g: J \rightarrow \mathbb{R}$ is well-defined, and

$$g(x) = g(A) + \int_A^x \Delta().$$

By hypothesis, each f'_n is cts; thus Δ is cts, by (55), P.15. By FTC, P.25, the map $x \mapsto \int_A^x \Delta$ is differentiable, and its derviative equals Δ . So g is differentiable, and $g' = 0 + \Delta$. ♦

Series and Sequences

In a normed-VS \mathbf{V} , a series $\vec{s} \subset \mathbf{V}$ is convergent if the “sequence \vec{p} of partial sums” converges in \mathbf{V} , where

94a:
$$p_k := \sum_{n=1}^k s_n .$$

Series $\vec{s} \subset \mathbf{V}$ is **absolutely convergent** if $\sum_{n=1}^{\infty} \|s_n\|$ is finite.

94b: Lemma. *Suppose $\vec{s} \subset \mathbf{V}$, where \mathbf{V} is a complete normed-VS. If \vec{s} is absolutely convergent, then \vec{s} is convergent.* \diamond

Proof. Let $p_k := \sum_{n=1}^k s_n$. Our goal is to show \vec{p} Cauchy. Fix $\varepsilon > 0$. Take K_0 st. for all pairs $L > K$ exceeding K_0 ,

*:
$$\sum_{n \in (K .. L]} \|s_n\| \leq \varepsilon .$$

By Δ nequality, LhS(*) dominates the norm of

$$\sum_{n \in (K .. L]} s_n \stackrel{\text{note}}{=} p_L - p_K .$$

Thus $\varepsilon \geq \|p_L - p_K\|$. \diamond

Exer. 9:Dis/Prove: Even in a *non-complete* normed-VS, abs-convergence implies convergence.

Defn. A sequence $\vec{s} \subset \mathbb{R}$ is a function, so use \vec{s}^+ to mean the corresponding *positive-part* sequence, from (80.1), and use \vec{s}^- for the seq of *negative parts*. These two sequences are non-negative, and satisfy that

94c:
$$\begin{aligned} s_n^+ + s_n^- &= |s_n| \quad \text{and} \\ s_n^+ - s_n^- &= s_n . \end{aligned} \quad \square$$

95: Reordering Thm. *Suppose sequence $\vec{s} \subset \mathbb{R}$ satisfies*

- i: Terms $s_k \rightarrow 0$, as $k \nearrow \infty$.*
- ii: Sum $\sum_{n=1}^{\infty} s_n^+ = \infty$. And $\sum_{n=1}^{\infty} s_n^- = \infty$.*

Then for each pair of values $A \leq B$ in $[-\infty, +\infty]$, there exists a reordering, \vec{y} , of \vec{s} for which

$$\left[\limsup_{K \rightarrow \infty} \sum_{n \in [1 .. K]} y_n \right] = B \quad \text{and} \quad \left[\liminf_{K \rightarrow \infty} \sum_{n \in [1 .. K]} y_n \right] = A . \quad \diamond$$

Pf(Sketch). Let $b_1 \geq b_2 \geq \dots > 0$ be an enumeration of the positive elts of \vec{s} . Let $a_1 \leq a_2 \leq \dots \leq 0$ be an enumeration of the non-positive elts of \vec{s} . From (i) and (ii),

95①:
$$b_n \searrow 0 \quad \text{and} \quad a_n \nearrow 0, \quad \text{as } n \rightarrow \infty .$$

95②:
$$\sum_{k=1}^{\infty} b_k = +\infty \quad \text{and} \quad \sum_{\ell=1}^{\infty} a_{\ell} = -\infty .$$

Think of \vec{y} as initially being an empty “stack”, into which we “pop” the elts of \vec{b} and \vec{a} , also viewed as stacks. We leave to the Reader the case where either A or B is $\pm\infty$.

Pop the \vec{b} -stack until the running-sum

$$p_{K_1} := \left[\sum_{n \in [1 .. K_1]} s_n \right] \text{ exceeds } B .$$

Now pop the \vec{a} -stack until the first time $K_2 > K_1$ that the running-sum has $p_{K_2} < A$. Return to popping the \vec{b} -stack, stopping at the first time $K_3 > K_2$ that $p_{K_3} > B$. Etc.

Condition (95②) says that the procedure never stops; so the $\limsup \geq B$ and the $\liminf \leq A$. Condition (95①) implies that $\limsup \leq B$ and $\liminf \geq A$. \diamond

Appendix: *Menagerie of Strange Functions*

(Being revised)

Pf of (ii). The foregoing showed $\text{Range}(\vec{c}) \subset \text{DisCty}(V)$. For the opposite, we fix a $z \in J \setminus \text{Range}(\vec{c})$ and show that $V()$ is left-cts at z .

Unfinished: as of 1Feb2022 ♦

Prelims. For a fnc $f: \mathbb{R} \rightarrow \mathbb{R}$ and point $z \in \mathbb{R}$, let

$$\text{A1: } f(z^+) := \lim_{x \searrow z} f(x) \text{ and } f(z^-) := \lim_{x \nearrow z} f(x),$$

when these limits exist. Use “ f is **right-continuous** at z ” to mean that $f(z^+) = f(z)$. Define **left-continuous** analogously.

Strictly increasing fnc, with Cty and DisCty dense.

Let $J := [0, 1]$. Mapping from $J \rightarrow \mathbb{R}$, we define a fnc $V = V_{\vec{c}, \vec{h}}$ determined by a **placement sequence** $\vec{c} \subset J$, and a **height sequence** $\vec{h} \subset \mathbb{R}_+$. The place-seq \vec{c} must be dense in J , and have distinct values. The height-seq must have $\sum(\vec{h})$ finite. We typically

A2: Normalize $\sum_{n=1}^{\infty} h_n = 1$, and have $\vec{c} \not\ni 0$.

Our definition, for each $x \in J$ is:

$$\text{A3: } V(x) = V_{\vec{c}, \vec{h}}(x) := \sum \left(\left\{ h_k \mid \begin{array}{l} k \in \mathbb{Z}_+ \text{ and} \\ c_k \leq x \end{array} \right\} \right).$$

Courtesy (A2), we have $V(1) = \sum(\vec{h}) = 1$ and $V(0) = \sum(\emptyset) = 0$.

A4: Jag-fnc Thm. Consider a $V = V_{\vec{c}, \vec{h}}$ which is normalized, (A2). Then $V(1) = 1$ and $V(0) = 0$. Further, V is strictly-increasing and maps $J \hookrightarrow J$. Moreover

i: Function V is right-continuous. And for each N :
 $f(c_N) - f(c_N^-) = h_N$.

ii: $\text{DisCty}(V) = \text{Range}(\vec{c})$. ♦

Sketch of right-cty. Fix $z \in [0, 1)$. For each $x \in (z, 1]$, let R_x be the set of indices k with $c_k \in (z, x]$. Thus

$$V(x) - V(z) = \sum_{k \in R_x} h_k.$$

And R_x decreases to the void-set, as $x \searrow z$, so the sum goes to zero. (Exer: Fill in the details, and for next paragraph too.)

Fix N . For $x \in [0, c_N)$, let L_x comprise those k with $c_k \in (x, c_N]$. So $V(c_N) - V(x)$ equals $\sum_{k \in L_x} h_k$. Sending $x \nearrow z$ makes the L_x sets decrease down to the singleton $\{c_N\}$. ♦

§Index for ADVANCED-CALC NOTES

I use these NOTES in my Advanced Calc and Modern Analysis courses.

CLD(metric-OR-space), 6

OPN(metric-OR-space), 5

1-point compactification of \mathbb{R} , 5

2-point compactification of \mathbb{R} , 5

Lip
 \succ , 13

Topo, Cau
 \succ , \succ , 5

\succ , \preccurlyeq , refining ptns, 21

f^+ , f^- , positive/negative part, 24

\circ , \triangleleft , \triangleright , *see* composition

absolutely convergent, 29

Addition-is-continuous thm, 4

antiderivative, 25

ball, punctured, 17

Banach space, 6

block, 21

Bolzano, 3

boundary, 7

bufferable, 11

circular reasoning, *see* tautology

closure point, 6

cluster point, 7

cluster-pt cpt, 8

CMS = “complete MS”, 15

compactness properties, 8

complete, 3, 15

composition, 16, 24

continuity set, $\text{Cty}(f)$, 12

continuous, 12, 14

uniformly, 12

$C(X)$, $C_{\text{Bnd}}(X)$, 6

$C(X \rightarrow \Omega)$, $C_{\text{Bnd}}(X \rightarrow \Omega)$, 14

convergence

of \vec{x} , 2, 6

pointwise, uniform, 15

countable local-base, 7

cover, 8, 26

finite, 8

Lebesgue number, 9, 12

open, 8

patch, 8

CSD = “Countably self-dense”, 15

cutpoints, 14, 21

\mathbf{D} , differentiation, 20

Discontinuity set, $\text{DisCty}(f)$, 12, 13

approximations, 27

\mathbb{D} , dyadic rationals, 13, 19

error term, 20

extended reals, 3, 5

finite cover, *see* cover

first-countable, *see* space, LCG

\mathcal{F}_σ -set, 27

function(al), *see* map

\mathcal{G}_δ -set, 27

grok, *see* STRANGER IN A STRANGE LAND, 13

HS = “Hausdorff space”, 6, 7

indicator function, $1_S(\cdot)$, 13

induced topology, 7

integrable, 22

interior point, 7

IPVS = “inner-product space”, 1

join of a partition, 21

Jordan mass, 26

Lebesgue mass, 26

Lebesgue number, 9, 28

limit-point, 7

linear functional, 1

Lip-equiv, 13

Lipschitz, 13, 25

ℓ^p , $\|\cdot\|_p$, 5

LUBP, Least upper-bound property, 3

map

antiderivative, 25

bilinear, 20

biLipschitz, 13

homeomorphism, 7

isometry, 7

linear, -functional, 1, 23

nearest-integer, distance, 19

ruler function, 13, 24

van der Waerden, 19

mesh size, 21

metric, 5

arclength, chordal, 5

arctan, stereographic, 5

Equivalence

topologically/Cauchy, 5

Lipschitz, 13

Mr. Rogers, *see* neighborhood

MS = “metric space”, 5

Multiplication-is-cts thm, 4

neighborhood, nbhd, 7

norm, 4, 6, 18

Euclidean, 10

$\|\cdot\|_p$, $\|\cdot\|_{\text{sup}}$, 5

NVS = “normed vectorspace”, 4

NP = “north pole”, 5

open cover, *see* cover

oscillation, 22

P.L, *see* piecewise-linear

panic

don't, 6

partition, 17, 21

atom, 21

block, 14, 21

join, 21

pointed, 21

refinement, 21

piecewise-linear fnc, $\text{PL}_{\vec{p}, \vec{h}}$, 14, 16

rational, 15

POFA, 19

pointwise operations, 14

positive/negative part, **24**, **29**
 pptn = “pointed partition”, **21**
 precompact, **10**
 prejudice, **6**
 Proof
 by contradiction, **9**, **10**, **12**
 circular, *see* circular reasoning
 provisional, **11**
 ptn, *see* partition

 relative topology, **7**
 ruler function, **13**, **24**

 \circ
 \mathbb{S}, \mathbb{S} , (punctured) circle, **5**
 sample point, tag, **21**
 seq-cpt, **8**
 sequence
 Cauchy, **2**, **5**
 tail of, **2**
 set
 inescapable, **7**
 open, **6**
 space/property
 cluster-point compact, **8**
 compact, **6**, **8**
 cover-positive, **9**
 Hausdorff, **6**
 LCG, **7**
 metric/topological, **6**
 metrizable, **6**
 sequentially compact, **8**
 totally-bounded, TB, **9**
 space/property Euclidean, **10**
 standard form, **13**

 tag, sample point, **21**
 $\text{Tail}_N(\vec{x})$, **2**
 tautology, *see* Proof, circular
 TB = “totally bounded”, **9**
 Theorems
 Addition-is-continuous, **4**
 Multiplication-is-cts, **4**
 topo-equiv, **5**
 topology, **6**
 TOS = “totally-ordered space”, **2**
 total derivative, **20**

TS = “Topological space”, **6**

 unif. convergence, *see* convergence
 uniform continuity, *see* continuous
 upper-bound, **3**

 van der Waerden function, **19**
 variation, **22**
 vectorspace, **6**
 VS = “vector space”, **1**

 $\mathbf{0}, \widehat{\mathbf{0}}$, zero-vector, **1**