Modern Fractal Geometry

Volume I: Dimensions, Measures, and Self-Similarity

Working reader's edition — revised 14 September 2026

Reader's Guide

This volume begins the common geometric-measure-theory and iterated-function-system kernel of the series. It assumes the real-analysis material in Anthony W. Knapp's *Basic Real Analysis* and *Advanced Real Analysis*, together with the separately declared harmonic-analysis, additive-combinatorics, and foundational dynamical-systems companions. Every chapter begins with its fractal-geometric definitions and conventions. These are part of the book, not prerequisites to be found in working files. Definitions repeated in the development serve as reminders at their point of use.

Names of the permitted prerequisites. The following abbreviations identify sources, not additional mathematical assumptions. A-RA denotes the permitted real-analysis background in Knapp's two books. The source code ARA-KB1 means Knapp's *Basic Real Analysis* [Kna05b], and ARA-KA1 means his *Advanced Real Analysis* [Kna05a]. A-SRA denotes individually cited results from Stein and Shakarchi's *Real Analysis: Measure Theory, Integration, and Hilbert Spaces* [SS05]; its source code is ASRA-SS1. The codes A-HA, A-AC, and A-DS denote the harmonic-analysis, additive-combinatorics, and foundational dynamical-systems companion books, respectively. Such a code does not by itself license an unspecified theorem: an external proof step must identify the result it uses.

The printed identifiers preserve the series dependency graph. A result marked [F] has its proof in the chapter or an earlier chapter. A result marked [G] gives the declared guided architecture and identifies the omitted research block. A result marked [B] is an exact cited interface whose proof is intentionally external. A node marked [Q] is a question or a later-owner boundary and is never a proof premise. Publication currency is separate from proof depth.

Read the chapter introduction first for its main questions and its connection to earlier material. The definitions section gives precise quantifiers and conventions; the following development explains their consequences. Linked identifiers lead to the relevant statements and definitions within the book. The notation D-I.5.02, for example, refers to Definition I.5.02.

Each chapter keeps intuition physically distinct from formal mathematical work. Intuition sections explain why the definitions and proof mechanisms are useful, but no later proof may cite them. The exercises are followed by an identifier-matched solution compendium at the end of the volume.

This working reader's edition is not independently reviewed. The graduate-student, structure, and three expert reviews are scheduled after all seven volumes have passed their internal gates.

Contents

Chapter I.1

Covers, Gauges, and Hausdorff Measure

How can one measure a set whose size lies between that of a point and that of a curve? We begin with covers of small diameter. Gauges allow the cost of a cover to depend on scale in more ways than a single power; power gauges then lead to Hausdorff dimension. The main tasks are to make this construction into a measure, establish its behavior under mappings, and turn a growth bound for a carried measure into a lower bound for dimension. The examples at the end distinguish dimension from closure and other elementary set operations.

The definitions below fix this chapter's conventions. Geometric intuition and proof strategy are discussed separately after the main development; exercises and their solutions provide a second route through the material.

Definitions and notation

All gauges and Hausdorff measures in this chapter use **diameter normalisation**: a covering set of diameter r costs $g(r)$. No Euclidean volume normalising constant is inserted.

Definition — I.1.01: Metric conventions

Let (X, d) be a metric space. For $x \in X$, $r > 0$, and $A, B \subset X$, set

$$B(x,r) = \{y \in X : d(x,y) < r\}, \qquad \overline{B}(x,r) = \{y \in X : d(x,y) \le r\},$$

$$d(x,A) = \inf_{a \in A} d(x,a), \qquad \operatorname{dist}(A,B) = \inf\{d(a,b) : a \in A,\ b \in B\}.$$

The infimum of the empty set is $+\infty$; hence $\operatorname{dist}(\emptyset, B) = \operatorname{dist}(A, \emptyset) = +\infty$.

Definition — I.1.02: Diameter

For $U \subset X$,

$$\operatorname{diam} U = \sup\{d(x, y) : x, y \in U\} \in [0, +\infty],$$

with $\operatorname{diam}(\emptyset) = 0$.

Definition — I.1.03: Countable delta-cover

Let $E \subset X$ and $\delta > 0$. A **countable cover** of E is a finite or countably infinite family $(U_i)_{i \in I}$ of subsets of X such that $E \subset \cup_{i \in I} U_i$. It is a **delta-cover** if $\operatorname{diam}(U_i) \leq \delta$ for every i. The empty family covers only the empty set.

Definition — I.1.04: Gauge

A **gauge** is a nondecreasing function

$$g : [0, +\infty) \longrightarrow [0, +\infty)$$

such that $g(0) = 0$ and $\lim_{r \downarrow 0} g(r) = 0$. No continuity away from zero and no strict positivity are required. For gauges g, h, write $g \preccurlyeq_0 h$ if there exist $0 < C < \infty$ and $r_0 > 0$ such that

$$g(r) \leq C h(r) \qquad (0 < r \leq r_0).$$

Write $g = o_0(h)$ if for every $\varepsilon > 0$ there exists $r_\varepsilon > 0$ such that

$$g(r) \leq \epsilon h(r) \qquad (0 < r \leq r_\epsilon).$$

Definition — I.1.05: Delta-level gauge content

For $E \subset X$ and $\delta > 0$, define

$$\mathscr{H}^g_\delta(E) = \inf\left\{ \sum_{i \in I} g(\operatorname{diam} U_i) : (U_i)_{i \in I} \text{ is a countable } \delta\text{-cover of } E \right\}.$$

The infimum of the empty collection of admissible covers is $+\infty$, and the sum over the empty cover is zero.

Definition — I.1.06: Outer measure

An **outer measure** on X is a function $\mu^* : \mathscr{P}(X) \to [0, +\infty]$ satisfying

$$\mu^*(\emptyset) = 0,$$

$$A \subset B \Longrightarrow \mu^*(A) \leq \mu^*(B),$$

and, for every sequence $(A_n)_{n\geq 1}$ of subsets of X,

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n).$$

Definition — I.1.07: Metric outer measure

An outer measure μ^* on (X, d) is a **metric outer measure** if

$$\operatorname{dist}(A, B) > 0 \quad \Longrightarrow \quad \mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$$

for all $A, B \subset X$.

Definition — I.1.08: Carathéodory measurability and Borel regularity

Given an outer measure μ^*, a set $E \subset X$ is **Carathéodory measurable** if, for every $T \subset X$,

$$\mu^*(T) = \mu^*(T \cap E) + \mu^*(T \setminus E).$$

An outer measure on a metric space is **Borel regular** if every $A \subset X$ has a Borel superset $B \supset A$ satisfying $\mu^*(B) = \mu^*(A)$.

Definition — I.1.09: Gauge Hausdorff measure

The **Hausdorff measure with gauge** g is

$$\mathscr{H}^g(E) = \sup_{\delta>0} \mathscr{H}^g_\delta(E).$$

The derived identity $\mathscr{H}^g(E) = \lim_{\delta\downarrow 0} \mathscr{H}^g_\delta(E)$ is L-I.1.01 and is not part of the logical content of the definition.

Definition — I.1.10: Power gauges and Hausdorff dimension

For $s > 0$, let $g_s(r) = r^s$ and abbreviate $\mathscr{H}^s_\delta = \mathscr{H}^{g_s}_\delta$ and $\mathscr{H}^s = \mathscr{H}^{g_s}$. Define

$$\dim_{\mathrm{H}} \varnothing = 0,$$

and, for nonempty $E \subset X$,

$$\dim_{\mathrm{H}} E = \inf\{s > 0 \;:\; \mathscr{H}^s(E) = 0\},$$

where $\inf(\varnothing) = +\infty$.

Definition — I.1.11: Lipschitz and bi-Lipschitz maps

For metric spaces (X, d_X) and (Y, d_Y), a map $f : E \to Y$, $E \subset X$, is **L-Lipschitz**, $0 \le L < \infty$, if

$$d_Y(f(x), f(y)) \le L d_X(x, y) \qquad (x, y \in E).$$

It is **bi-Lipschitz** if it is injective and there exist $0 < c \le C < \infty$ such that

$$c d_X(x, y) \le d_Y(f(x), f(y)) \le C d_X(x, y) \qquad (x, y \in E).$$

Definition — I.1.12: Carried measure and gauge growth bound

Let μ be a finite Borel measure on X and let $E \subset X$ be Borel. The measure μ is **carried by** E if $\mu(X \setminus E) = 0$. It satisfies a (g, C, r_0) **growth bound on** E if $C > 0$, $r_0 > 0$, and

$$\mu(\overline{B}(x, r)) \le C g(r) \qquad (x \in E,\ 0 < r \le r_0).$$

Definition — I.1.13: Graphs and product metric

For $A \subset \mathbb{R}^m$ and $f : A \to \mathbb{R}^n$,

$$\operatorname{graph}(f) = \{(x, f(x)) : x \in A\} \subset \mathbb{R}^{m+n}.$$

For metric spaces (X, d_X) and (Y, d_Y), the **max product metric** on $X \times Y$ is

$$d_\infty((x, y), (x', y')) = \max\{d_X(x, x'), d_Y(y, y')\}.$$

All measures use diameter normalisation.

I.1.1 Covers and gauges

Throughout Sections 1–4, (X, d) is a metric space. For subsets $A, B \subset X$,

$$d(x, A) = \inf_{a \in A} d(x, a), \qquad \operatorname{dist}(A, B) = \inf_{a \in A,\, b \in B} d(a, b),$$

with the infimum of the empty set equal to $+\infty$. For $U \subset X$,

$$\operatorname{diam} U = \sup_{x,y \in U} d(x, y), \qquad \operatorname{diam} \emptyset = 0.$$

Let $\delta > 0$. A countable δ-cover of $E \subset X$ is a finite or countably infinite family $(U_i)_{i \in I}$ such that

$$E \subset \bigcup_{i \in I} U_i, \qquad \operatorname{diam} U_i \le \delta \quad (i \in I).$$

A gauge is a nondecreasing function $g : [0, +\infty) \to [0, +\infty)$ for which

$$g(0) = 0, \qquad \lim_{r\downarrow 0} g(r) = 0.$$

Define

$$\mathscr{H}^g_\delta(E) = \inf\left\{\sum_{i\in I} g(\operatorname{diam} U_i) \;:\; (U_i)_{i\in I} \text{ is a countable } \delta\text{-cover of } E\right\}. \tag{1.1}$$

The infimum is $+\infty$ if no such cover exists. The empty family covers the empty set and has cost zero.

Lemma — 1.1 (L-I.1.01: monotonicity in scale)

If $0 < \delta' \le \delta$, then

$$\mathscr{H}^g_{\delta'}(E) \ge \mathscr{H}^g_\delta(E). \tag{1.2}$$

Therefore

$$\mathscr{H}^g(E) := \sup_{\delta>0} \mathscr{H}^g_\delta(E) = \lim_{\delta\downarrow 0} \mathscr{H}^g_\delta(E). \tag{1.3}$$

Proof

Every δ'-cover is a δ-cover, so the infimum in (1.1) is taken over a smaller class at scale δ'. Thus (1.2) holds. A monotone nondecreasing function of the directed variable $\delta \downarrow 0$ has limit equal to its supremum, with values allowed in $[0, +\infty]$. This proves (1.3). $\square$

Lemma — 1.2 (L-I.1.02: outer-measure property)

For every gauge g, the set function $\mathscr{H}^g$ is an outer measure: it vanishes at the empty set, is monotone, and is countably subadditive.

Proof

The empty cover gives $\mathscr{H}^g_\delta(\emptyset) = 0$ for every $\delta > 0$, hence $\mathscr{H}^g(\emptyset) = 0$.

If $A \subset B$, every cover of B covers A; consequently

$$\mathscr{H}^g_\delta(A) \le \mathscr{H}^g_\delta(B) \quad (\delta > 0),$$

and taking the supremum over δ proves monotonicity.

Let $(E_n)_{n\ge 1}$ be subsets of X. If $\sum_n \mathscr{H}^g(E_n) = +\infty$, countable subadditivity is automatic. Suppose instead that this series is finite. Fix $\delta > 0$ and $\varepsilon > 0$. For each n, choose a countable δ-cover $(U_{n,j})_j$ of E_n such that

$$\sum_j g(\operatorname{diam} U_{n,j}) \le \mathscr{H}^g_\delta(E_n) + \epsilon 2^{-n}. \tag{1.4}$$

Such a cover exists because $\mathscr{H}^g_\delta(E_n) \le \mathscr{H}^g(E_n) < +\infty$. The doubly indexed family covers $\cup_n E_n$ at scale δ, and hence

$$\begin{aligned}\mathscr{H}^g_\delta\left(\bigcup_{n=1}^\infty E_n\right) &\le \sum_{n=1}^\infty \sum_j g(\operatorname{diam} U_{n,j}) \\ &\le \sum_{n=1}^\infty \mathscr{H}^g_\delta(E_n) + \epsilon \qquad (1.5) \\ &\le \sum_{n=1}^\infty \mathscr{H}^g(E_n) + \epsilon.\end{aligned}$$

Letting $\varepsilon \downarrow 0$ and then taking the supremum over $\delta > 0$ gives

$$\mathscr{H}^g\left(\bigcup_{n=1}^\infty E_n\right) \le \sum_{n=1}^\infty \mathscr{H}^g(E_n).$$

Thus $\mathscr{H}^g$ is an outer measure. □

Lemma — 1.3 (L-I.1.03: metric additivity)

If $A, B \subset X$ and $\rho = \operatorname{dist}(A, B) > 0$, then

$$\mathscr{H}^g(A \cup B) = \mathscr{H}^g(A) + \mathscr{H}^g(B). \qquad (1.6)$$

Proof

Fix $0 < \delta < \rho$. No set of diameter at most δ can meet both A and B. Given any countable δ-cover $(U_i)_i$ of $A \cup B$, partition the indices that meet $A \cup B$ into

$$I_A = \{i : U_i \cap A \neq \varnothing\}, \qquad I_B = \{i : U_i \cap B \neq \varnothing\}.$$

These sets of indices are disjoint. The corresponding subfamilies cover A and B, respectively, so

$$\sum_i g(\operatorname{diam} U_i) \ge \mathscr{H}^g_\delta(A) + \mathscr{H}^g_\delta(B).$$

Taking the infimum over covers yields

$$\mathscr{H}^g_\delta(A \cup B) \ge \mathscr{H}^g_\delta(A) + \mathscr{H}^g_\delta(B).$$

Letting $\delta \downarrow 0$ gives the inequality $\ge$ in (1.6). The reverse inequality follows from Lemma 1.2. □

An outer measure satisfying (1.6) whenever $\operatorname{dist}(A, B) > 0$ is called a metric outer measure. Thus Lemmas 1.2–1.3 prove that $\mathscr{H}^g$ is a metric outer measure.

I.1.2 Carathéodory measurability and Borel regularity

Let μ^* be an outer measure on a set X. A set $E \subset X$ is Carathéodory measurable if

$$\mu^*(T) = \mu^*(T \cap E) + \mu^*(T \setminus E) \qquad (T \subset X). \qquad (2.1)$$

Write M_{μ^*} for the collection of such sets.

Lemma — 2.1 (L-I.1.04: Carathéodory theorem)

The collection M_{μ^*} is a sigma-algebra, $\mu^*|M_{\mu^*}$ is a countably additive complete measure, and every subset of a μ^*-null set is in M_{μ^*}.

Proof

The empty set satisfies (2.1), and (2.1) is unchanged when E is replaced by $X \setminus E$; hence M_{μ^*} contains the empty set and is closed under complements.

Let $E, F \in M_{\mu^*}$ and $T \subset X$. Splitting first by E and then splitting $T \setminus E$ by F gives

$$\mu^*(T) = \mu^*(T \cap E) + \mu^*((T \setminus E) \cap F) + \mu^*(T \setminus (E \cup F)). \tag{2.2}$$

By outer subadditivity,

$$\mu^*(T \cap (E \cup F)) \le \mu^*(T \cap E) + \mu^*((T \setminus E) \cap F).$$

Equation (2.2) therefore supplies the $\ge$ direction in (2.1) for $E \cup F$; the reverse direction is outer subadditivity. Thus M_{μ^*} is an algebra.

Now let $(E_n)_{n\ge1}$ be pairwise disjoint members of M_{μ^*}, and put $F_N = \cup_{n=1}^N E_n$ and $F = \cup_{n=1}^\infty E_n$. Repeated use of (2.1) gives, for every N,

$$\mu^*(T) = \sum_{n=1}^N \mu^*(T \cap E_n) + \mu^*(T \setminus F_N). \tag{2.3}$$

Because $T \setminus F \subset T \setminus F_N$, (2.3) implies

$$\mu^*(T) \ge \sum_{n=1}^\infty \mu^*(T \cap E_n) + \mu^*(T \setminus F) \ge \mu^*(T \cap F) + \mu^*(T \setminus F). \tag{2.4}$$

Outer subadditivity gives the reverse inequality, so $F \in M_{\mu^*}$. An arbitrary countable union is reduced to the disjoint case by replacing E_n with

$$E_n \setminus \bigcup_{j<n} E_j,$$

which remains in the algebra. Therefore M_{μ^*} is a sigma-algebra.

If the measurable sets E_n are pairwise disjoint, outer subadditivity gives

$$\mu^*(F) \le \sum_{n=1}^\infty \mu^*(E_n),$$

while finite additivity and monotonicity give

$$\mu^*(F) \ge \mu^*(F_N) = \sum_{n=1}^N \mu^*(E_n) \qquad (N \ge 1).$$

Letting N tend to infinity proves countable additivity.

Finally, suppose $N \in M_{\mu^*}$, $\mu^*(N) = 0$, and $A \subset N$. For every $T \subset X$,

$$\mu^*(T \cap A) = 0, \qquad \mu^*(T \setminus A) \le \mu^*(T).$$

These inequalities and outer subadditivity imply

$$\mu^*(T) = \mu^*(T \cap A) + \mu^*(T \setminus A).$$

Thus $A \in M_{\mu^*}$ and the restricted measure is complete. □

Lemma — 2.2 (L-I.1.05: metric outer measures measure Borel sets)

Let (X, d) be a metric space and let μ^* be a metric outer measure on X. Every closed subset of X lies in M_{μ^*}. Consequently every Borel subset of X lies in M_{μ^*}.

Proof

Fix a closed set $F \subset X$ and a test set $T \subset X$. If $F = \emptyset$, then $\mu^*(T) = \mu^*(T \cap F) + \mu^*(T \setminus F) = 0 + \mu^*(T)$; assume from now on that F is nonempty. If $\mu^*(T) = +\infty$, outer subadditivity forces

$$+\infty = \mu^*(T \cap F) + \mu^*(T \setminus F),$$

so (2.1) holds. Assume henceforth that $\mu^*(T) < +\infty$.
The distance function is 1-Lipschitz:

$$|d(x, F) - d(y, F)| \le d(x, y) \qquad (x, y \in X). \tag{2.5}$$

For $k \ge 0$, put $r_k = 3^{-k}$ and

$$A_k = \{x \in T : r_{k+1} \le d(x, F) < r_k\}. \tag{2.6}$$

If $\ell \ge k + 2$, $x \in A_k$, and $y \in A_\ell$, then (2.5) gives

$$d(x, y) \ge d(x, F) - d(y, F) > r_{k+1} - r_{k+2} > 0. \tag{2.7}$$

Every finite family of even-indexed shells is therefore pairwise positively separated, as is every finite family of odd-indexed shells. Finite induction using metric additivity and monotonicity yields

$$\sum_{\substack{0 \le k \le N \\ k \text{ even}}} \mu^*(A_k) \le \mu^*(T), \qquad \sum_{\substack{0 \le k \le N \\ k \text{ odd}}} \mu^*(A_k) \le \mu^*(T).$$

Hence

$$\sum_{k=0}^{\infty} \mu^*(A_k) \le 2\mu^*(T) < +\infty. \tag{2.8}$$

Define

$$T_N = \{x \in T \,:\, d(x, F) \geq r_N\}.$$

Since $\operatorname{dist}(T \cap F, T_N) \geq r_N$, metric additivity and monotonicity give

$$\mu^*(T) \geq \mu^*((T \cap F) \cup T_N) = \mu^*(T \cap F) + \mu^*(T_N). \tag{2.9}$$

Because F is closed, every point of $T \setminus F$ has positive distance from F, and

$$T \setminus F \subset T_N \cup \bigcup_{k=N}^{\infty} A_k.$$

Equations (2.8)–(2.9) imply

$$\begin{aligned} \mu^*(T \setminus F) &\leq \mu^*(T_N) + \sum_{k=N}^{\infty} \mu^*(A_k) \\ &\leq \mu^*(T) - \mu^*(T \cap F) + \sum_{k=N}^{\infty} \mu^*(A_k). \end{aligned}$$

Let N tend to infinity. The series tail tends to zero, so

$$\mu^*(T) \geq \mu^*(T \cap F) + \mu^*(T \setminus F).$$

Outer subadditivity gives the opposite inequality. Thus F is Carathéodory measurable. Lemma 2.1 and the fact that closed sets generate the Borel sigma-algebra prove the final assertion. □

Lemma — 2.3 (closed covers)

For every $E \subset X$ and $\delta > 0$, the infimum in (1.1) is unchanged if every covering set is required to be closed.

Proof

For every $U \subset X$,

$$\operatorname{diam} \overline{U} = \operatorname{diam} U. \tag{2.10}$$

Indeed, one inequality follows from $U \subset \overline{U}$; for the other, approximate any two points of $\overline{U}$ by points of U and use continuity of the metric. Replacing every member of a cover by its closure preserves the cover, all diameters, and therefore its cost. Since closed covers are a subclass of all covers, the two infima agree. □

Lemma — 2.4 (L-I.1.06: Borel regularity)

Assume that (X, d) is separable. For every $A \subset X$ there is a Borel set $B \supset A$ such that

$$\mathscr{H}^g(B) = \mathscr{H}^g(A). \tag{2.11}$$

Proof

Separability guarantees a countable δ-cover of X for every $\delta > 0$: if D is a countable dense set, the balls $(B(q, \delta/4))_{q \in D}$ cover X and have diameter at most $\delta/2$.

If $\mathscr{H}^g(A) = +\infty$, take $B = X$; then (2.11) follows from monotonicity. Suppose $\mathscr{H}^g(A) < +\infty$. For every $n \geq 1$, Lemma 2.3 permits a countable closed $1/n$-cover $(F_{n,j})_j$ of A satisfying

$$\sum_j g(\operatorname{diam} F_{n,j}) \leq \mathscr{H}^g_{1/n}(A) + 2^{-n}. \tag{2.12}$$

Let

$$B_n = \bigcup_j F_{n,j}, \qquad B = \bigcap_{n=1}^{\infty} B_n. \tag{2.13}$$

Each B_n is F_σ, so B is Borel, and $A \subset B$. Fix $\delta > 0$. For every sufficiently large n, $1/n \leq \delta$, and the family in (2.12) is a δ-cover of $B \subset B_n$. Consequently

$$\mathscr{H}^g_\delta(B) \leq \mathscr{H}^g_{1/n}(A) + 2^{-n} \leq \mathscr{H}^g(A) + 2^{-n}.$$

Arbitrarily large such n may be chosen, so $\mathscr{H}^g_\delta(B) \leq \mathscr{H}^g(A)$. Taking the supremum over δ gives $\mathscr{H}^g(B) \leq \mathscr{H}^g(A)$. The reverse inequality follows from $A \subset B$ and monotonicity. ☐

Theorem — 2.5 (T-I.1.01: metric Hausdorff-measure construction)

Let (X, d) be a separable metric space and let g be a gauge. Then $\mathscr{H}^g$ is a Borel-regular metric outer measure. Its restriction to its Carathéodory sigma-algebra is a complete measure, and this sigma-algebra contains every Borel subset of X.

Proof

Lemmas 1.2 and 1.3 give the metric outer-measure assertion, Lemmas 2.1 and 2.2 give completeness and Borel measurability, and Lemma 2.4 gives Borel regularity. ☐

I.1.3 Gauge comparison and Hausdorff dimension

For gauges g, h, write $g \preccurlyeq_0 h$ if there are $0 < C < +\infty$ and $r_0 > 0$ such that

$$g(r) \leq C h(r) \qquad (0 < r \leq r_0), \tag{3.1}$$

and write $g = o_0(h)$ if for every $\varepsilon > 0$ there is $r_\varepsilon > 0$ such that

$$g(r) \leq \epsilon h(r) \qquad (0 < r \leq r_\epsilon). \tag{3.2}$$

Theorem — 3.1 (T-I.1.02: gauge comparison)

Let $E \subset X$.

(1) If (3.1) holds, then $\mathscr{H}^g(E) \leq C\mathscr{H}^h(E)$.

(2) If (3.2) holds and $\mathscr{H}^h(E) < +\infty$, then $\mathscr{H}^g(E) = 0$.

Proof

If $0 < \delta \le r_0$, every δ-cover $(U_i)_i$ satisfies

$$\sum_i g(\operatorname{diam} U_i) \le C \sum_i h(\operatorname{diam} U_i).$$

Taking infima gives

$$\mathscr{H}^g_\delta(E) \le C\mathscr{H}^h_\delta(E).$$

Letting $\delta \downarrow 0$ proves part 1. Under (3.2), the same argument gives

$$\mathscr{H}^g(E) \le \epsilon \mathscr{H}^h(E) \qquad (\epsilon > 0).$$

The right side tends to zero because $\mathscr{H}^h(E)$ is finite. This proves part 2. □

For $s > 0$, define the power gauge $g_s(r) = r^s$, write $\mathscr{H}^s$ for $\mathscr{H}^{g_s}$, set $\dim_{\mathrm{H}}(\varnothing) = 0$, and, for nonempty E, define

$$\dim_{\mathrm{H}} E = \inf\{s > 0 \ : \ \mathscr{H}^s(E) = 0\}, \tag{3.3}$$

with $\inf(\varnothing) = +\infty$.

Corollary — 3.2 (C-I.1.03: power-gauge critical jump)

If $0 < s < t$, then

$$\mathscr{H}^s(E) < +\infty \Longrightarrow \mathscr{H}^t(E) = 0, \tag{3.4}$$

and

$$\mathscr{H}^t(E) > 0 \Longrightarrow \mathscr{H}^s(E) = +\infty. \tag{3.5}$$

Proof

Since

$$\frac{r^t}{r^s} = r^{t-s} \longrightarrow 0 \qquad (r \downarrow 0),$$

Theorem 3.1(2), with $g = g_t$ and $h = g_s$, proves (3.4). Statement (3.5) is the contrapositive of (3.4). □

Theorem — 3.3 (T-I.1.04: critical exponent)

Let E be nonempty and $d = \dim_{\mathrm{H}}(E)$. Then

$$\mathscr{H}^s(E) = +\infty \quad (0 < s < d), \qquad \mathscr{H}^s(E) = 0 \quad (s > d), \tag{3.6}$$

and

$$d = \sup\{s > 0 \,:\, \mathscr{H}^s(E) = +\infty\}, \tag{3.7}$$

where $\sup(\emptyset) = 0$.

Proof

Let $Z = \{s > 0 \,:\, \mathscr{H}^s(E) = 0\}$. If $d < +\infty$ and $s > d = \inf Z$, the definition of the infimum gives $t \in Z$ with $t < s$. If $d = +\infty$, there is no $s > d$. Corollary 3.2, or its proof with a zero rather than merely finite $\mathscr{H}^t(E)$, gives $\mathscr{H}^s(E) = 0$.

Now suppose $0 < s < d$. If $\mathscr{H}^s(E) < +\infty$, choose t with $s < t < d$. Corollary 3.2 would give $\mathscr{H}^t(E) = 0$, hence $d \le t$, a contradiction. Thus $\mathscr{H}^s(E) = +\infty$. The two conclusions in (3.6), including the cases $d = 0$ and $d = +\infty$, give (3.7). No conclusion follows at $s = d$: the critical measure may be zero, positive and finite, or infinite. □

Theorem — 3.4 (T-I.1.05: countable stability)

For every sequence $(E_n)_{n\ge1}$ of subsets of X,

$$\dim_{\mathrm{H}}\left(\bigcup_{n=1}^{\infty} E_n\right) = \sup_{n\ge1} \dim_{\mathrm{H}} E_n. \tag{3.8}$$

Proof

Put $E = \cup_n E_n$ and $a = \sup_n \dim_{\mathrm{H}}(E_n)$. Monotonicity under inclusion, which follows directly from (3.3) and monotonicity of $\mathscr{H}^s$, gives

$$\dim_{\mathrm{H}} E \ge a.$$

If $a = +\infty$, this is equality. Suppose $a < +\infty$ and take $s > a$. Theorem 3.3 gives $\mathscr{H}^s(E_n) = 0$ for every nonempty E_n; the same equality holds for an empty E_n by Lemma 1.2. Thus Lemma 1.2 gives

$$\mathscr{H}^s(E) \le \sum_{n=1}^{\infty} \mathscr{H}^s(E_n) = 0.$$

Thus $\dim_{\mathrm{H}}(E) \le s$ for every $s > a$, so $\dim_{\mathrm{H}}(E) \le a$. □

I.1.4 Lipschitz maps and the mass-distribution principle

A map $f \,:\, E \to Y$ between metric spaces is L-Lipschitz if

$$d_Y(f(x), f(y)) \le L d_X(x, y) \qquad (x, y \in E). \tag{4.1}$$

It is bi-Lipschitz if it is injective and there are $0 < c \le C < +\infty$ such that

$$cd_X(x,y) \le d_Y(f(x), f(y)) \le Cd_X(x,y) \qquad (x, y \in E). \tag{4.2}$$

Theorem — 4.1 (T-I.1.06: Lipschitz measure inequality)

Let $s > 0$ and let $f : E \to Y$ be L-Lipschitz. If $L > 0$, then

$$\mathscr{H}^s(f(E)) \le L^s \mathscr{H}^s(E). \tag{4.3}$$

If $L = 0$, then $\mathscr{H}^s(f(E)) = 0$.

Proof

Suppose $L > 0$. If $(U_i)_i$ is a δ-cover of E, then

$$V_i = f(E \cap U_i)$$

covers $f(E)$ and satisfies

$$\operatorname{diam} V_i \le L \operatorname{diam} U_i \le L\delta.$$

Therefore

$$\mathscr{H}^s_{L\delta}(f(E)) \le L^s \sum_i (\operatorname{diam} U_i)^s.$$

Taking the infimum over δ-covers and then letting $\delta \downarrow 0$ proves (4.3). If $L = 0$, (4.1) makes $f(E)$ empty or a singleton, which has zero s-dimensional Hausdorff measure because it is covered by a set of diameter zero. □

Corollary — 4.2 (C-I.1.07: Lipschitz monotonicity and bi-Lipschitz invariance)

If $f : E \to Y$ is Lipschitz, then

$$\dim_{\mathrm{H}} f(E) \le \dim_{\mathrm{H}} E. \tag{4.4}$$

If f is bi-Lipschitz, equality holds.

Proof

If $\dim_{\mathrm{H}}(E) = +\infty$, (4.4) is automatic. Otherwise, for every $s > \dim_{\mathrm{H}}(E)$, Theorem 3.3 and Theorem 4.1 give

$$\mathscr{H}^s(E) = 0 \quad \Longrightarrow \quad \mathscr{H}^s(f(E)) = 0.$$

This proves (4.4). Under (4.2), the inverse $f^{-1} : f(E) \to E$ is c^{-1}-Lipschitz. Applying (4.4) to f and to f^{-1} gives equality. □

Let μ be a finite Borel measure on X, and let $E \subset X$ be Borel. We say that μ is carried by E if $\mu(X \setminus E) = 0$.

Theorem — 4.3 (T-I.1.08: mass-distribution principle)

Let g be a gauge. Suppose that μ is a finite Borel measure carried by a Borel set E, that $0 < \mu(E) < +\infty$, and that there are $C > 0$ and $r_0 > 0$ such that

$$\mu(\overline{B}(x,r)) \le Cg(r) \qquad (x \in E,\ 0 < r \le r_0). \tag{4.5}$$

Then

$$\mathscr{H}^g(E) \ge \frac{\mu(E)}{C} > 0. \tag{4.6}$$

If $g(r) = r^s$, it follows that $\dim_{\mathrm{H}}(E) \ge s$.

Proof

First, for each $x \in E$, (4.5) and $g(r) \to 0$ give

$$0 \le \mu(\{x\}) \le \mu(\overline{B}(x,r)) \le Cg(r) \quad (0 < r \le r_0),$$

and hence $\mu(\{x\}) = 0$.

Fix $0 < \delta \le r_0$. If no countable closed δ-cover exists, then $\mathscr{H}^g_\delta(E) = +\infty$ and the desired lower bound at this scale already holds. Otherwise, let $(F_i)_i$ be an arbitrary countable closed δ-cover of E. For each i with $F_i \cap E$ nonempty, choose $x_i \in F_i \cap E$ and put $r_i = \operatorname{diam}(F_i)$. Since μ is carried by E,

$$\mu(F_i) = \mu(F_i \cap E).$$

If $r_i > 0$, then $F_i \subset \overline{B}(x_i, r_i)$, so (4.5) gives

$$\mu(F_i) \le Cg(r_i).$$

If $r_i = 0$, the nonempty set F_i is a singleton and the same inequality holds because both sides are zero. Sets disjoint from E have μ-measure zero. Countable subadditivity of μ now gives

$$\mu(E) \le \sum_i \mu(F_i) \le C \sum_i g(\operatorname{diam} F_i). \tag{4.7}$$

By Lemma 2.3, closed covers yield the same infimum as arbitrary covers. Taking the infimum in (4.7) gives

$$\mathscr{H}^g_\delta(E) \ge \frac{\mu(E)}{C}.$$

Letting $\delta \downarrow 0$ proves (4.6). In the power-gauge case, $\mathscr{H}^s(E) > 0$; Theorem 3.3 then rules out $\dim_{\mathrm{H}}(E) < s$. □

I.1.5 Stress tests

Proposition — 5.1 (P-I.1.09: intervals)

If $a < b$, then, with diameter normalisation,

$$\mathscr{H}^1([a,b]) = b - a, \qquad \dim_{\mathrm{H}}[a,b] = 1. \tag{5.1}$$

Proof

Fix $\delta > 0$ and choose an integer N with $(b-a)/N \le \delta$. Partition $[a,b]$ into N closed intervals of length $(b-a)/N$. Their total diameter is $b-a$, so

$$\mathscr{H}^1_\delta([a,b]) \le b - a. \tag{5.2}$$

Conversely, let $(U_i)_i$ be any countable cover of $[a,b]$. Each nonempty $U_i \subset R$ is contained in an interval of length $\mathrm{diam}(U_i)$. By monotonicity and countable subadditivity of Lebesgue outer measure λ^*,

$$b - a = \lambda^*([a,b]) \le \sum_i \lambda^*(U_i) \le \sum_i \mathrm{diam}\, U_i. \tag{5.3}$$

Thus $\mathscr{H}^1_\delta([a,b]) \ge b-a$ for every $\delta > 0$. Equations (5.2)–(5.3) prove the first assertion. Corollary 3.2 gives

$$\mathscr{H}^s([a,b]) = +\infty \quad (0 < s < 1), \qquad \mathscr{H}^s([a,b]) = 0 \quad (s > 1),$$

which proves the dimension assertion. □

Proposition — 5.2 (P-I.1.10: countable dense sets)

Every countable subset E of a metric space has Hausdorff dimension zero. Nevertheless, Hausdorff dimension is not invariant under closure; in particular,

$$\dim_{\mathrm{H}}(\mathbb{Q} \cap [0,1]) = 0, \qquad \dim_{\mathrm{H}} \overline{\mathbb{Q} \cap [0,1]} = 1. \tag{5.4}$$

Proof

A singleton has $\mathscr{H}^s$-measure zero for every $s > 0$, because it has a cover of diameter zero. Countable subadditivity gives $\mathscr{H}^s(E) = 0$ for every $s > 0$. Thus $\dim_{\mathrm{H}}(E) = 0$. The set $Q \cap [0,1]$ is countable and dense in $[0,1]$, so (5.4) follows from Proposition 5.1. □

Proposition — 5.3 (P-I.1.11: Lipschitz graphs)

Let $A \subset \mathbb{R}^m$ and let $f : A \to \mathbb{R}^n$ be L-Lipschitz for Euclidean metrics. Then

$$\dim_{\mathrm{H}} \mathrm{graph}(f) = \dim_{\mathrm{H}} A. \tag{5.5}$$

Proof

The graph map $G : A \to \mathbb{R}^{m+n}$, $G(x) = (x, f(x))$, satisfies

$$|G(x) - G(y)|^2 = |x - y|^2 + |f(x) - f(y)|^2 \le (1 + L^2)|x - y|^2.$$

Thus G is $\sqrt{1 + L^2}$-Lipschitz. The coordinate projection $\pi : \mathbb{R}^{m+n} \to \mathbb{R}^m$ is 1-Lipschitz and restricts to the inverse of G on the graph. Corollary 4.2 applied in both directions proves (5.5). □

Proposition — 5.4 (P-I.1.12: elementary products)

Let $E \subset X$ and $F \subset Y$ be nonempty, and equip $X \times Y$ with

$$d_\infty((x, y), (x', y')) = \max\{d_X(x, x'), d_Y(y, y')\}. \tag{5.6}$$

Then

$$\dim_{\mathrm{H}}(E \times F) \ge \max\{\dim_{\mathrm{H}} E, \dim_{\mathrm{H}} F\}. \tag{5.7}$$

For every $y_0 \in Y$,

$$\dim_{\mathrm{H}}(E \times \{y_0\}) = \dim_{\mathrm{H}} E. \tag{5.8}$$

Proof

The coordinate projections from $E \times F$ onto E and F are 1-Lipschitz and surjective. Corollary 4.2 therefore gives both lower bounds in (5.7). The map $x \mapsto (x, y_0)$ is an isometry from E to $E \times y_0$ under (5.6), which proves (5.8). □

I.1.5.1 Boundary of the product discussion

There is no general conclusion

$$\dim_{\mathrm{H}}(E \times F) = \dim_{\mathrm{H}} E + \dim_{\mathrm{H}} F$$

from the results proved here. The upper and lower product estimates that are valid in general require packing or box dimensions and are owned by Chapter I.2. A particularly sharp warning already comes from countability:

$$\dim_{\mathrm{H}}((\mathbb{Q} \cap [0, 1])^2) = 0, \qquad \overline{(\mathbb{Q} \cap [0, 1])^2} = [0, 1]^2.$$

Density in a high-dimensional ambient set is therefore not a dimension lower bound.

I.1.6 Intuition and strategy

The following discussion explains the geometric ideas behind the proofs.

I.1.6.1 The chapter's central idea

Hausdorff measure asks how cheaply a set can be covered when every covering piece of diameter r is charged a price $g(r)$. Two operations are essential:

1. minimise the bill at a fixed resolution δ; and
2. force δ to zero.

The second operation is not cosmetic. Without it, a single large covering set can join geometrically separate pieces and destroy the metric additivity that makes the construction a genuine Borel measure.

I.1.6.2 Why gauges precede dimensions

The power price r^s gives the familiar Hausdorff dimensions, but a single number s cannot describe all critical-scale behaviour. At the eventual critical exponent the power measure can be zero, finite positive, or infinite. Logarithmic corrections and other gauges later distinguish sets having the same Hausdorff dimension. Beginning with g therefore prevents the reader from mistaking dimension for the whole theory.

I.1.6.3 The exponent as a phase transition

At small scales and for $s < t$,

$$r^t \ll r^s.$$

Thus a cover affordable at exponent s becomes arbitrarily cheap at exponent t. The entire critical-exponent theory is this observation plus exact control of the order of the two operations: first minimise at scale δ, then let δ tend to zero.

I.1.6.4 The two directions of a dimension calculation

An upper bound is a construction: exhibit economical covers. A lower bound is an obstruction: show that every sufficiently fine cover must be expensive. The mass-distribution principle implements the obstruction. A measure whose mass in every small ball is at most $Cg(r)$ cannot be concentrated inside a cover whose total g-cost is less than $\mu(E)/C$.

This is why the later Frostman lemma is so powerful. The principle proved here says what a well-distributed measure would imply; Frostman's lemma in Chapter I.3 constructs such a measure from a Hausdorff-dimension hypothesis.

I.1.6.5 Mapping intuition

A Lipschitz map cannot enlarge a covering piece by more than a fixed factor. Power costs therefore change by at most the corresponding power of that factor. Bi-Lipschitz maps have controlled distortion in both directions, so Hausdorff dimension becomes an invariant of metric geometry rather than of a particular coordinate description.

The graph example illustrates both directions of this argument. The graph map gives one Lipschitz direction and coordinate projection gives the other; the equality of dimensions is then structural, not a special graph-covering calculation.

I.1.6.6 Three warnings

1. **Closure is dangerous.** A countable dense set has dimension zero even when its closure is an interval.
2. **Critical measure is not prescribed.** Knowing $\dim_{\mathrm{H}}(E) = s$ does not by itself determine $\mathscr{H}^s(E)$.
3. **Products need more than Hausdorff dimension.** Coordinate projections provide an elementary lower bound, but a correct general upper bound needs the packing/box machinery of I.2. The elementary lower bound does not imply an additive product formula.

I.1.6.7 Understanding the shell argument

In Lemma 2.2, the complement of a closed set is divided into layers according to distance from that set. Adjacent layers may touch, so metric additivity cannot simply be applied to all of them at once. Separating the layers into even and odd families leaves positive gaps between any finitely many layers in each family. This makes finite metric additivity applicable. The finite outer measure of the test set then bounds both series of layer masses, and their tails tend to zero. The exact radii and separation estimates in the proof are what justify this geometric picture.

I.1.7 Exercises

All exercises use only definitions and results available in Chapter I.1 together with the permitted real-analysis background.

I.1.7.1 Exercise 1 — Similarities

Let $f : E \to Y$ satisfy

$$d_Y(f(x), f(y)) = \lambda d_X(x, y) \qquad (x, y \in E)$$

for some $\lambda > 0$. Prove, for every $s > 0$,

$$\mathscr{H}^s(f(E)) = \lambda^s \mathscr{H}^s(E).$$

Deduce again that $\dim_{\mathrm{H}}(f(E)) = \dim_{\mathrm{H}}(E)$.

I.1.7.2 Exercise 2 — Hölder maps

Suppose $0 < \theta \le 1$ and

$$d_Y(f(x), f(y)) \le L d_X(x, y)^\theta \qquad (x, y \in E),$$

where $L > 0$. Prove that, for every $t > 0$,

$$\mathscr{H}^t(f(E)) \le L^t \mathscr{H}^{\theta t}(E),$$

and hence

$$\dim_{\mathrm{H}} f(E) \le \theta^{-1} \dim_{\mathrm{H}} E.$$

I.1.7.3 Exercise 3 — Snowflaked metrics

Let $0 < \theta \le 1$ and define $d_\theta(x, y) = d(x, y)^\theta$.

1. Prove that d_θ is a metric.
2. Prove that
$$\mathscr{H}^s_{d_\theta,\delta}(E) = \mathscr{H}^{\theta s}_{d,\delta^{1/\theta}}(E).$$
3. Deduce
$$\dim_{\mathrm{H}}^{d_\theta} E = \theta^{-1} \dim_{\mathrm{H}}^{d} E.$$

The superscript on the dimension names the ambient metric, not a power.

I.1.7.4 Exercise 4 — Open subsets of the line

Prove that every nonempty open set $U \subset R$ has

$$\dim_{\mathrm{H}} U = 1.$$

I.1.7.5 Exercise 5 — A measure need only live on the set

In Theorem 4.3, replace the assumption that μ is carried by E with the weaker assumptions

$$0 < \mu(E) < +\infty, \qquad \mu(E \cap \overline{B}(x, r)) \le C g(r)$$

for $x \in E$ and $0 < r \le r_0$. Prove that the same conclusion $\mathscr{H}^g(E) \ge \mu(E)/C$ holds.

I.1.7.6 Exercise 6 — Positive subsets inherit a lower bound

Under the hypotheses of Theorem 4.3, let $A \subset E$ be Borel with $\mu(A) > 0$. Prove

$$\mathscr{H}^g(A) \ge \frac{\mu(A)}{C}.$$

I.1.7.7 Exercise 7 — Countable modifications

Let E be a subset of a metric space and let N be countable. Prove

$$\dim_{\mathrm{H}}(E \cup N) = \dim_{\mathrm{H}} E.$$

Show by an example that replacing “countable” by “closure of a countable set” makes the assertion false.

I.1.7.8 Exercise 8 — Products with countable sets

Let $E \subset X$ be nonempty and let $F \subset Y$ be nonempty and countable. Give $X \times Y$ the max metric. Prove

$$\dim_{\mathrm{H}}(E \times F) = \dim_{\mathrm{H}} E.$$

I.1.7.9 Exercise 9 — Distance sets as a first Lipschitz application

Fix x_0 in a metric space (X, d) and define

$$\Delta_{x_0}(E) = \{d(x_0, x) : x \in E\} \subset [0, +\infty).$$

Prove

$$\dim_{\mathrm{H}} \Delta_{x_0}(E) \leq \dim_{\mathrm{H}} E.$$

I.1.7.10 Exercise 10 — The middle-thirds Cantor set

Let

$$C = \left\{ \sum_{n=1}^{\infty} \frac{2\omega_n}{3^n} : \omega_n \in \{0, 1\} \right\}$$

and put $\alpha = \log 2 / \log 3$. Let μ be the pushforward to C of the fair Bernoulli product measure on $\{0, 1\}^{N}$ under the displayed coding map.

1. Cover C by its 2^n level-n basic intervals and prove $\mathscr{H}^{\alpha}(C) \leq 1$.
2. Prove that, for every $x \in C$ and $0 < r \leq 1$,

$$\mu(\overline{B}(x, r)) \leq 2 \cdot 6^{\alpha} r^{\alpha}.$$

3. Apply the mass-distribution principle to prove

$$0 < \mathscr{H}^{\alpha}(C) \leq 1, \qquad \dim_{\mathrm{H}} C = \frac{\log 2}{\log 3}.$$

The existence of the Bernoulli product measure is an A-RA probability fact; the cylinder masses 2^{-n} fix the normalisation used here.

Chapter I.2

Packing, Box Dimensions, and Products

Counting balls at a fixed scale is often easier than optimizing covers of many different sizes. This chapter develops that counting viewpoint and explains why upper and lower box dimensions can disagree. Packing dimension repairs the lack of countable stability of upper box dimension. Digit-restriction examples make these distinctions concrete, and Cartesian products show why different notions of dimension naturally occur together in one estimate.

The definitions below fix this chapter's conventions. Geometric intuition and proof strategy are discussed separately after the main development; exercises and their solutions provide a second route through the material.

Definitions and notation

The conventions below are controlling. They deliberately distinguish equal-scale separated sets from variable-radius packings and distinguish packing premeasure from packing outer measure.

Definition — I.2.01: Total boundedness and scales

A subset E of a metric space (X, d) is **totally bounded** if, for every $r > 0$, it is contained in a finite union of open balls of radius r. Box dimensions are defined below only for nonempty totally bounded sets. All logarithmic scale limits are taken through $0 < r < 1$ with $r \downarrow 0$.

Definition — I.2.02: Covering and separated-set numbers

For a nonempty totally bounded set $E \subset X$ and $r > 0$, define

$$N_r(E) = \min \left\{ m \in \mathbb{N} \,:\, E \subset \bigcup_{j=1}^{m} \overline{B}(x_j, r), \quad x_j \in E \right\}.$$

A set $S \subset E$ is r-**separated** if

$$x, y \in S,\ x \neq y \quad \Longrightarrow \quad d(x,y) > r.$$

Define

$$M_r(E) = \max\{\#S \,:\, S \subset E \text{ is } r\text{-separated}\}.$$

Total boundedness makes both numbers finite, and the integer supremum in the definition of $M_r(E)$ is attained. Put $N_r(\varnothing) = M_r(\varnothing) = 0$ when an empty-set convention is needed outside logarithms.

Definition — I.2.03: Lower and upper box dimensions

For a nonempty totally bounded set E, define

$$\underline{\dim}_{\mathrm{B}} E = \liminf_{r \downarrow 0} \frac{\log N_r(E)}{-\log r}, \qquad \overline{\dim}_{\mathrm{B}} E = \limsup_{r \downarrow 0} \frac{\log N_r(E)}{-\log r}.$$

If the two values agree, their common value is the **box dimension** $\dim_{\mathrm{B}} E$. Set both box dimensions of the empty set equal to zero. No box dimension is assigned here to a nonempty set that is not totally bounded.

Definition — I.2.04: Modified upper box dimension

For $E \subset \mathbb{R}^d$, define $\overline{\dim}_{\mathrm{MB}}(\varnothing) = 0$. If E is nonempty, define its **modified upper box dimension** by

$$\overline{\dim}_{\mathrm{MB}} E = \inf\left\{ \sup_{j \geq 1} \overline{\dim}_{\mathrm{B}} E_j \,:\, E \subset \bigcup_{j=1}^{\infty} E_j,\ E_j \subset \mathbb{R}^d \text{ nonempty and bounded} \right\}.$$

Finite covers may be repeated to produce a sequence. The infimum is therefore over a nonempty class, even when E is unbounded.

Definition — I.2.05: Centred packings and packing premeasure

Let (X,d) be a separable metric space, $E \subset X$, $s \geq 0$, and $\delta > 0$. A **centred** δ**-packing of** E is a finite or countably infinite family

$$(\overline{B}(x_i, r_i))_{i \in I}$$

of pairwise disjoint closed balls with $x_i \in E$ and $0 < r_i \leq \delta$. Its s-cost is

$$\sum_{i \in I} (2r_i)^s.$$

The empty packing is allowed and has cost zero. The **packing premeasure at scale** δ and the **zero-scale packing premeasure** are

$$\mathscr{P}^s_\delta(E) = \sup\left\{\sum_{i\in I}(2r_i)^s \;:\; (\overline{B}(x_i,r_i))_{i\in I} \text{ is a centred } \delta\text{-packing of } E\right\},$$

$$\mathscr{P}^s_0(E) = \inf_{\delta>0}\mathscr{P}^s_\delta(E).$$

The cost uses twice the declared radius. In Euclidean space this is the diameter of the ball; in a general metric space the actual diameter can be smaller, and the formula still uses $2r_i$. At $s=0$, every ball costs one.

Definition — I.2.06: Packing outer measure

For $E \subset X$ and $s \geq 0$, define

$$\mathscr{P}^s(E) = \inf\left\{\sum_{j=1}^{\infty}\mathscr{P}^s_0(E_j) \;:\; E\subset\bigcup_{j=1}^{\infty}E_j,\, E_j\subset X\right\}.$$

The empty sequence covers only the empty set and has cost zero. The symbol P^s_0 always denotes the premeasure; P^s always denotes the outer measure.

Definition — I.2.07: Packing dimension

Set $\dim_{\mathrm{P}}(\varnothing)=0$. For nonempty $E\subset X$, define

$$\dim_{\mathrm{P}} E = \inf\{s\geq 0 \;:\; \mathscr{P}^s(E)=0\},$$

with $\inf(\varnothing)=+\infty$.

Definition — I.2.08: Product metrics

For metric spaces (X,d_X) and (Y,d_Y), the max product metric is

$$d_\infty((x,y),(x',y')) = \max\{d_X(x,x'),d_Y(y,y')\}.$$

In Euclidean products we also use

$$d_2((x,y),(x',y')) = \left(d_X(x,x')^2+d_Y(y,y')^2\right)^{1/2}.$$

The inequalities $d_\infty\leq d_2\leq\sqrt{2}d_\infty$ make these metrics bi-Lipschitz equivalent. Unless stated otherwise, products in internal proofs carry d_∞.

Definition — I.2.09: Digit-restriction Cantor sets

For $S \subset N$, define

$$K_S = \left\{ \sum_{n=1}^{\infty} \frac{2\varepsilon_n}{3^n} \,:\, \varepsilon_n \in \{0,1\}\ (n \in S), \quad \varepsilon_n = 0\ (n \notin S) \right\} \subset [0,1].$$

Write

$$s_n = \#(S \cap \{1,\dots,n\}), \qquad \underline{d}(S) = \liminf_{n\to\infty} \frac{s_n}{n}, \qquad \overline{d}(S) = \limsup_{n\to\infty} \frac{s_n}{n}.$$

These sets are used only for explicit comparisons of dimension notions.

Internal proofs use the max product metric. Every ball occurring in a packing is closed and centred on the set being packed.

I.2.1 Covering numbers and box dimensions

Let (X,d) be a metric space. A set $E \subset X$ is totally bounded if, for each $r > 0$, finitely many open balls of radius r cover E. For a nonempty totally bounded set define

$$N_r(E) = \min\left\{ m \in \mathbb{N} \,:\, E \subset \bigcup_{j=1}^{m} \overline{B}(x_j, r),\ x_j \in E \right\}. \tag{1.1}$$

An r-separated subset S of E satisfies $d(x,y) > r$ whenever $x, y \in S$ are distinct. Put

$$M_r(E) = \max\{\#S \,:\, S \subset E \text{ is } r\text{-separated}\}. \tag{1.2}$$

Total boundedness makes $M_r(E)$ finite: an r-separated set has at most one point in each member of any cover by open balls of radius $r/2$. Since the possible cardinalities are integers, their finite supremum is attained.

Lemma — 1.1 (L-I.2.01: covering–separation sandwich)

For every $r > 0$,

$$M_{2r}(E) \le N_r(E) \le M_r(E). \tag{1.3}$$

Proof

Let $S \subset E$ be $2r$-separated and let $E \subset \cup_{j=1}^{m} \overline{B}(x_j, r)$. Two points of S cannot lie in the same covering ball, since their mutual distance would then be at most $2r$. Assigning to each point of S one ball containing it gives $\#S \le m$. Maximising over S and minimising over the covers proves the first inequality.

Let S be an r-separated subset of cardinality $M_r(E)$. It is maximal by inclusion: otherwise one could add a point and obtain a larger r-separated set. Hence every $x \in E$ has distance at most r from some point of S, for a point farther than r from all of S could be added. Thus

$$E \subset \bigcup_{y \in S} \overline{B}(y, r),$$

which proves $N_r(E) \le M_r(E)$. □

Define

$$\underline{\dim}_{\mathrm{B}} E = \liminf_{r \downarrow 0} \frac{\log N_r(E)}{-\log r}, \qquad \overline{\dim}_{\mathrm{B}} E = \limsup_{r \downarrow 0} \frac{\log N_r(E)}{-\log r}. \tag{1.4}$$

Only $0 < r < 1$ enters these limits.

Lemma — 1.2 (L-I.2.02: bounded-ratio discrete scales)

Suppose (r_k) decreases to zero and

$$0 < c \le r_{k+1}/r_k \le 1. \tag{1.5}$$

Both limits in (1.4) are unchanged when evaluated only at $r = r_k$. They are also unchanged when $N_r(E)$ is replaced by $M_r(E)$.

Proof

If $r_{k+1} < r \le r_k$, monotonicity of covering numbers gives

$$N_{r_k}(E) \le N_r(E) \le N_{r_{k+1}}(E). \tag{1.6}$$

Moreover,

$$0 \le (-\log r_{k+1}) - (-\log r_k) = -\log(r_{k+1}/r_k) \le -\log c. \tag{1.7}$$

Since $-\log r_k$ tends to infinity, the two denominators in (1.7) have ratio tending to one. Divide (1.6) by $-\log r$, use $-\log r_k \le -\log r < -\log r_{k+1}$, and take lower and upper limits. The continuous-scale and discrete-scale liminf and limsup coincide.

By Lemma 1.1,

$$M_{2r}(E) \le N_r(E) \le M_r(E). \tag{1.8}$$

Multiplying the scale by the fixed constant 2 changes $-\log r$ by the constant $-\log 2$; the preceding bounded-ratio argument therefore shows that M_{2r}, N_r, and M_r have identical logarithmic lower and upper growth rates. □

Theorem — 1.3 (T-I.2.03: structural properties of box dimensions)

For nonempty totally bounded sets, lower and upper box dimensions are monotone under inclusion and invariant under closure. Upper box dimension is finitely stable:

$$\overline{\dim}_{\mathrm{B}}(E \cup F) = \max\{\overline{\dim}_{\mathrm{B}} E, \overline{\dim}_{\mathrm{B}} F\}. \tag{1.9}$$

If $f : E \to Y$ is Lipschitz, then

$$\underline{\dim}_{\mathrm{B}} f(E) \leq \underline{\dim}_{\mathrm{B}} E, \qquad \overline{\dim}_{\mathrm{B}} f(E) \leq \overline{\dim}_{\mathrm{B}} E. \tag{1.10}$$

Both values are invariant under bi-Lipschitz maps. Neither value is countably stable.

Proof

The inequality $\underline{\dim}_{\mathrm{B}} E \leq \overline{\dim}_{\mathrm{B}} E$ is the elementary inequality between liminf and limsup.

Suppose $A \subset E$. A radius-r cover of E may have centres outside A. Discard every ball missing A; from each remaining ball choose $a_j \in A$ inside it. The triangle inequality gives

$$A \subset \bigcup_j \overline{B}(a_j, 2r), \qquad N_{2r}(A) \leq N_r(E). \tag{1.11}$$

Fixed rescaling does not change logarithmic growth, so (1.11) proves monotonicity of both box dimensions.

Every finite union of closed balls is closed. Hence a finite radius-r cover of E also covers $\overline{E}$, and

$$N_r(\overline{E}) \leq N_r(E). \tag{1.12}$$

Monotonicity applied to $E \subset \overline{E}$ supplies the reverse dimension inequalities. Thus both dimensions are closure invariant.

For bounded E, F in one metric space,

$$N_r(E \cup F) \leq N_r(E) + N_r(F) \leq 2\max\{N_r(E), N_r(F)\}. \tag{1.13}$$

After logarithms, the term $\log 2/(-\log r)$ tends to zero. Equation (1.13) gives the $\leq$ direction of (1.9), and monotonicity gives the other direction.

Let f be L-Lipschitz. If $L = 0$, then $f(E)$ is a singleton and both its box dimensions are zero. If $L > 0$, a cover in (1.1) maps to

$$f(E) \subset \bigcup_{j=1}^{N_r(E)} \overline{B}(f(x_j), Lr),$$

so

$$N_{Lr}(f(E)) \leq N_r(E). \tag{1.14}$$

This also proves that $f(E)$ is totally bounded. Fixed scale factors do not affect (1.4), so (1.10) follows. Applying (1.10) to a bi-Lipschitz map and to its Lipschitz inverse gives equality.

Finally let $D = Q \cap [0, 1]$. It is countable, but its closure is $[0, 1]$. By closure invariance it is enough to compute the dimensions of the interval. A cover of $[0, 1]$ by N closed radius-r balls gives a cover by sets of total diameter at most $2rN$; P-I.1.09 yields

$$1 = \mathscr{H}^1([0, 1]) \leq 2rN,$$

and hence $N_r([0, 1]) \geq 1/(2r)$. Equally spaced centres give $N_r([0, 1]) \leq \lceil 1/r \rceil + 1$. Both logarithmic limits are one. Thus

$$\underline{\dim}_{\mathrm{B}} D = \overline{\dim}_{\mathrm{B}} D = 1,$$

although D is a countable union of singletons, each of box dimension zero. □

I.2.2 Packing premeasure and packing outer measure

Let (X, d) now be separable. For $s \geq 0$ and $\delta > 0$, a centred δ-packing of $E \subset X$ is a finite or countable pairwise disjoint family

$$(\overline{B}(x_i, r_i))_{i \in I}, \qquad x_i \in E, \quad 0 < r_i \leq \delta.$$

Define

$$\mathscr{P}^s_\delta(E) = \sup \sum_{i \in I} (2r_i)^s, \qquad \mathscr{P}^s_0(E) = \inf_{\delta > 0} \mathscr{P}^s_\delta(E), \tag{2.1}$$

where the supremum is over all such packings. Define the covering regularisation

$$\mathscr{P}^s(E) = \inf \left\{ \sum_{j=1}^{\infty} \mathscr{P}^s_0(E_j) \; : \; E \subset \bigcup_{j=1}^{\infty} E_j \right\}. \tag{2.2}$$

Lemma — 2.1 (L-I.2.04: packing-premeasure calculus)

If $0 < \delta' \leq \delta$, then

$$\mathscr{P}^s_{\delta'}(E) \leq \mathscr{P}^s_\delta(E), \qquad \mathscr{P}^s_0(E) = \lim_{\delta \downarrow 0} \mathscr{P}^s_\delta(E). \tag{2.3}$$

The premeasures are monotone in E. If $\operatorname{dist}(A, B) > 0$, then

$$\mathscr{P}^s_0(A \cup B) = \mathscr{P}^s_0(A) + \mathscr{P}^s_0(B). \tag{2.4}$$

Proof

Every δ'-packing is a δ-packing, which proves the first inequality in (2.3). A monotone decreasing function along $\delta \downarrow 0$ has limit equal to its infimum, including the value $+\infty$. If $A \subset B$, every centred packing of A is one of B, so monotonicity is immediate.

Put $\rho = \operatorname{dist}(A, B) > 0$ and choose $0 < \delta < \rho/2$. Every centred packing of $A \cup B$ splits, according to its centres, into a packing of A and a packing of B. Consequently

$$\mathscr{P}^s_\delta(A \cup B) \leq \mathscr{P}^s_\delta(A) + \mathscr{P}^s_\delta(B). \tag{2.5}$$

Conversely, take one centred δ-packing of A and one of B. A ball from the first family and a ball from the second cannot intersect: if they intersected, the triangle inequality would give

$$d(a, b) \leq r_a + r_b \leq 2\delta < \rho.$$

Their union is therefore a packing of $A \cup B$. Taking the two suprema gives the reverse inequality in (2.5), also when a supremum is infinite. Thus equality holds in (2.5) for every $\delta < \rho/2$. Letting δ decrease to zero proves (2.4). □

The regularisation in (2.2) is essential: P_0^s need not be countably subadditive.

Theorem — 2.2 (T-I.2.05: metric packing outer measure)

For each $s \geq 0$, P^s is a metric outer measure. All Borel subsets of X are therefore Carathéodory measurable, and P^s restricts to a complete measure on its Carathéodory sigma-algebra.

Proof

Write $q(E) = P_0^s(E)$. The empty packing and monotonicity give $q(\emptyset) = 0$ and $A \subset B \Longrightarrow q(A) \leq q(B)$.

The empty cover in (2.2) gives $P^s(\emptyset) = 0$; monotonicity of P^s follows because every cover of a set covers each of its subsets. To prove countable subadditivity, let $E = \cup_n E_n$. If $\sum_n P^s(E_n) = +\infty$, there is nothing to prove. Otherwise fix $\varepsilon > 0$. For each n, choose sets $(E_{n,k})_{k\geq 1}$ covering E_n such that

$$\sum_{k=1}^{\infty} q(E_{n,k}) \leq \mathscr{P}^s(E_n) + \epsilon 2^{-n}. \tag{2.6}$$

The doubly indexed family covers E, and (2.2) gives

$$\mathscr{P}^s(E) \leq \sum_{n,k} q(E_{n,k}) \leq \sum_n \mathscr{P}^s(E_n) + \epsilon.$$

Letting $\varepsilon \downarrow 0$ proves outer subadditivity.

It remains to prove metric additivity. Suppose $\operatorname{dist}(A, B) > 0$. Outer subadditivity gives

$$\mathscr{P}^s(A \cup B) \leq \mathscr{P}^s(A) + \mathscr{P}^s(B). \tag{2.7}$$

For the reverse inequality, let (E_j) be any countable cover of $A \cup B$. By monotonicity of q and Lemma 2.1,

$$\begin{aligned} q(E_j) &\geq q((E_j \cap A) \cup (E_j \cap B)) \\ &= q(E_j \cap A) + q(E_j \cap B). \end{aligned} \tag{2.8}$$

The two families $(E_j \cap A)$ and $(E_j \cap B)$ cover A and B, respectively. Summing (2.8) and using (2.2) yields

$$\sum_j q(E_j) \geq \mathscr{P}^s(A) + \mathscr{P}^s(B).$$

Taking the infimum over all covers proves the reverse of (2.7). Thus P^s is metric.

By L-I.1.05, every Borel set is measurable for a metric outer measure. By L-I.1.04, the Carathéodory measurable sets form a sigma-algebra on which the outer measure is a complete measure. □

Corollary — 2.3 (C-I.2.06: packing critical jump)

If $0 \le s < t$, then

$$\mathscr{P}^s(E) < +\infty \quad \Longrightarrow \quad \mathscr{P}^t(E) = 0. \tag{2.9}$$

Consequently $P^t(E) > 0$ implies $P^s(E) = +\infty$.

Proof

For every centred δ-packing,

$$\sum_i (2r_i)^t \le (2\delta)^{t-s} \sum_i (2r_i)^s.$$

Taking suprema gives

$$\mathscr{P}^t_\delta(A) \le (2\delta)^{t-s} \mathscr{P}^s_\delta(A). \tag{2.10}$$

If $P^s_0(A) < +\infty$, there is $\delta_0 > 0$ with $P^s_{\delta_0}(A) < +\infty$; otherwise the infimum in (2.1) would be infinite. For $0 < \delta \le \delta_0$, monotonicity in scale and (2.10) give

$$0 \le \mathscr{P}^t_\delta(A) \le (2\delta)^{t-s} \mathscr{P}^s_{\delta_0}(A).$$

The right side tends to zero, so $P^t_0(A) = 0$.

Now suppose $P^s(E) < +\infty$. By (2.2), there is a cover (E_j) of E such that

$$\sum_j \mathscr{P}^s_0(E_j) < +\infty.$$

Every summand is finite, so the preceding paragraph gives $P^t_0(E_j) = 0$ for every j. Therefore (2.2) yields $P^t(E) = 0$. The final claim is the contrapositive in $[0, +\infty]$. □

Define $\dim_{\mathrm{P}}(\varnothing) = 0$ and, for nonempty E,

$$\dim_{\mathrm{P}} E = \inf\{s \ge 0 \,:\, \mathscr{P}^s(E) = 0\}, \tag{2.11}$$

with $\inf(\varnothing) = +\infty$.

Theorem — 2.4 (T-I.2.07: critical exponent and countable stability)

If E is nonempty and $p = \dim_{\mathrm{P}} E$, then

$$\mathscr{P}^s(E) = +\infty \quad (0 \le s < p), \qquad \mathscr{P}^s(E) = 0 \quad (s > p). \tag{2.12}$$

For every sequence (E_j),

$$\dim_{\mathrm{P}}\left(\bigcup_{j=1}^{\infty} E_j\right) = \sup_{j \ge 1} \dim_{\mathrm{P}} E_j. \tag{2.13}$$

Proof

Let $Z = \{s \geq 0 \,:\, P^s(E) = 0\}$. If $s > p = \inf Z$, the definition of an infimum gives $u \in Z$ with $u < s$; Corollary 2.3 gives $P^s(E) = 0$. If Z is empty, then $p = +\infty$ and this assertion has no case.

If $0 \leq s < p$ and $P^s(E) < +\infty$, choose t with $s < t < p$. Corollary 2.3 gives $P^t(E) = 0$, contradicting $p \leq t$. Hence $P^s(E) = +\infty$. This proves (2.12), with no assertion at $s = p$.

Put $E = \cup_j E_j$ and $a = \sup_j \dim_{\mathrm{P}} E_j$. Monotonicity gives $\dim_{\mathrm{P}} E \geq a$. If $a = +\infty$, equality follows. If $a < +\infty$, take $s > a$. Equation (2.12) gives $P^s(E_j) = 0$ for every j, and outer subadditivity yields

$$\mathscr{P}^s(E) \leq \sum_j \mathscr{P}^s(E_j) = 0.$$

Thus $\dim_{\mathrm{P}} E \leq s$ for every $s > a$, proving (2.13). Empty terms cause no change because their packing dimension and packing outer measure are zero. □

I.2.3 Packing dimension as modified upper box dimension

For $E \subset \mathbb{R}^d$, define

$$\overline{\dim}_{\mathrm{MB}} E = \inf\left\{\sup_j \overline{\dim}_{\mathrm{B}} E_j \,:\, E \subset \bigcup_j E_j, \quad E_j \text{ nonempty and bounded}\right\}, \tag{3.1}$$

and set the value equal to zero for the empty set.

Lemma — 3.1 (L-I.2.08: premeasure–box comparison)

Let E be nonempty and bounded in $\mathbb{R}^d$, and let $s \geq 0$.
(1) If $\overline{\dim}_{\mathrm{B}} E < s$, then $P_0^s(E) = 0$.
(2) If $\overline{\dim}_{\mathrm{B}} E > s$, then $P_0^s(E) = +\infty$.

Proof

Assume first that $\overline{\dim}_{\mathrm{B}} E < s$. Choose u with

$$\overline{\dim}_{\mathrm{B}} E < u < s. \tag{3.2}$$

Lemma 1.2, applied to the dyadic scales, gives an integer k_0 such that

$$M_{2^{-k}}(E) \leq 2^{ku} \qquad (k \geq k_0). \tag{3.3}$$

Fix $K \geq k_0$ and consider a centred 2^{-K}-packing $(\overline{B}(x_i, r_i))_i$ of E. Put

$$I_k = \{i \,:\, 2^{-k-1} < r_i \leq 2^{-k}\}, \qquad k \geq K.$$

If $i, j \in I_k$ are distinct, disjointness of the two closed Euclidean balls implies

$$|x_i - x_j| > r_i + r_j > 2^{-k}.$$

Thus their centres are 2^{-k}-separated, so (3.3) gives $\#I_k \le 2^{ku}$. Hence

$$\begin{aligned}\sum_i (2r_i)^s &\le \sum_{k=K}^{\infty} \#I_k\, 2^{s(1-k)} \\ &\le 2^s \sum_{k=K}^{\infty} 2^{-k(s-u)}.\end{aligned} \tag{3.4}$$

The right side is independent of the packing and tends to zero as $K \to \infty$. Taking the supremum and then the zero-scale infimum proves $P_0^s(E) = 0$.

Now suppose $\overline{\dim}_{\mathrm{B}} E > s$ and choose

$$s < t < \overline{\dim}_{\mathrm{B}} E. \tag{3.5}$$

By Lemma 1.2 there are scales $q_n \downarrow 0$ such that

$$M_{q_n}(E) > q_n^{-t}. \tag{3.6}$$

Choose a q_n-separated set of cardinality $M_{q_n}(E)$. The closed balls of radius $q_n/2$ about its points are pairwise disjoint and give a packing whose s-cost is

$$M_{q_n}(E) q_n^s > q_n^{s-t} \longrightarrow +\infty. \tag{3.7}$$

For every fixed $\delta > 0$, all sufficiently large n have $q_n/2 \le \delta$, and (3.7) is unbounded. Thus $P_\delta^s(E) = +\infty$ for every δ, and consequently $P_0^s(E) = +\infty$. □

Theorem — 3.2 (T-I.2.09: packing equals modified upper box dimension)

For every $E \subset \mathbb{R}^d$,

$$\dim_{\mathrm{P}} E = \overline{\dim}_{\mathrm{MB}} E. \tag{3.8}$$

Proof

The empty-set case follows from the definitions. Write $D = \overline{\dim}_{\mathrm{MB}} E$ and $p = \dim_{\mathrm{P}} E$.

Let $a > D$. By (3.1), there is a cover $E \subset \cup_j E_j$ by nonempty bounded sets satisfying

$$\sup_j \overline{\dim}_{\mathrm{B}} E_j < a.$$

Lemma 3.1 gives $P_0^a(E_j) = 0$ for every j. Equation (2.2) then gives $P^a(E) = 0$, so $p \le a$. Letting $a \downarrow D$ proves $p \le D$; if $D = +\infty$, this inequality is automatic.

For the reverse inequality, suppose $p < +\infty$ and take $s > p$. Theorem 2.4 gives $P^s(E) = 0$. By (2.2), there is a countable cover (A_j) of E such that

$$\sum_j \mathscr{P}_0^s(A_j) < 1. \tag{3.9}$$

Refine this cover to the nonempty sets

$$A_j \cap \overline{B}(0,k), \qquad j,k \geq 1.$$

They are bounded, they still cover E, and monotonicity gives each of them finite s-packing premeasure. The contrapositive of Lemma 3.1(2) therefore gives upper box dimension at most s for every refined member. Hence $D \leq s$. Letting $s \downarrow p$ proves $D \leq p$. If $p = +\infty$, that inequality is automatic. This proves (3.8). □

Corollary — 3.3 (C-I.2.10: Lipschitz behaviour of packing dimension)

Let $E \subset \mathbb{R}^m$ and let $f : E \to \mathbb{R}^n$ be Lipschitz. Then

$$\dim_{\mathrm{P}} f(E) \leq \dim_{\mathrm{P}} E. \tag{3.10}$$

Equality holds when f is bi-Lipschitz.

Proof

Let $E \subset \cup_j E_j$ be any cover by bounded sets as in (3.1), and put $A_j = E \cap E_j$. Discard empty A_j. Theorem 1.3 gives

$$\overline{\dim}_{\mathrm{B}} f(A_j) \leq \overline{\dim}_{\mathrm{B}} A_j \leq \overline{\dim}_{\mathrm{B}} E_j. \tag{3.11}$$

Each $f(A_j)$ is bounded: fix $x_j \in A_j$; if L is a Lipschitz constant, then $|f(x) - f(x_j)| \leq L|x - x_j|$ on A_j. The sets $f(A_j)$ cover $f(E)$. Taking the supremum in (3.11), then the infimum over the original covers, and using Theorem 3.2 proves (3.10). If f is bi-Lipschitz, apply (3.10) also to the inverse map on $f(E)$. □

Theorem — 3.4 (T-I.2.11: dimension hierarchy)

For every $E \subset \mathbb{R}^d$,

$$\dim_{\mathrm{H}} E \leq \dim_{\mathrm{P}} E. \tag{3.12}$$

If E is nonempty and bounded, then

$$\dim_{\mathrm{H}} E \leq \underline{\dim}_{\mathrm{B}} E \leq \overline{\dim}_{\mathrm{B}} E, \qquad \dim_{\mathrm{P}} E \leq \overline{\dim}_{\mathrm{B}} E. \tag{3.13}$$

Proof

First suppose E is nonempty and bounded. Put $\ell = \underline{\dim}_{\mathrm{B}} E$. If $\ell = +\infty$, the Hausdorff-to-lower-box inequality is automatic. Otherwise take $t > \ell$ and choose u with $\ell < u < t$. By the definition of the liminf, there are scales $r_k \downarrow 0$ such that

$$N_{r_k}(E) \leq r_k^{-u}. \tag{3.14}$$

The corresponding covering balls have diameter at most $2r_k$, so

$$\mathscr{H}^t_{2r_k}(E) \leq N_{r_k}(E)(2r_k)^t \leq 2^t r_k^{t-u} \longrightarrow 0. \tag{3.15}$$

For any fixed $\delta > 0$, arbitrarily large k satisfy $2r_k \le \delta$; the same covers are admissible at scale δ. Equation (3.15) therefore gives $\mathscr{H}^t_\delta(E) = 0$. Since δ was arbitrary, $\mathscr{H}^t(E) = 0$, and hence $\dim_{\mathrm{H}} E \le t$. Letting $t \downarrow \ell$ proves the first inequality in (3.13). The second is the liminf–limsup inequality. Theorem 3.2 and the one-set cover in (3.1) give

$$\dim_{\mathrm{P}} E = \overline{\dim}_{\mathrm{MB}} E \le \overline{\dim}_{\mathrm{B}} E,$$

which completes (3.13).

For arbitrary $E \subset \mathbb{R}^d$, take any cover $E \subset \cup_j E_j$ by bounded sets. Countable stability of Hausdorff dimension, monotonicity, and the bounded case give

$$\dim_{\mathrm{H}} E \le \sup_j \dim_{\mathrm{H}} E_j \le \sup_j \overline{\dim}_{\mathrm{B}} E_j. \tag{3.16}$$

Taking the infimum over covers and using Theorem 3.2 proves (3.12). The empty set has both dimensions zero. □

Proposition — 3.5 (P-I.2.12: lower box and packing are incomparable)

There are bounded subsets of the line for which lower box dimension is larger than packing dimension, and compact subsets for which it is smaller.

Proof

Let $D = Q \cap [0, 1]$. A singleton has packing dimension zero: for every $s > 0$, a centred packing of $\{x\}$ contains at most one ball, and

$$\mathscr{P}^s_\delta(\{x\}) \le (2\delta)^s,$$

so $P^s(\{x\}) = 0$. Countable stability gives $\dim_{\mathrm{P}} D = 0$, while Theorem 1.3 gives both box dimensions of D equal to one. Thus

$$\dim_{\mathrm{P}} D = 0 < 1 = \underline{\dim}_{\mathrm{B}} D. \tag{3.17}$$

For the reverse strict inequality, choose integers

$$1 = m_0 < m_1 < m_2 < \cdots, \qquad \frac{m_j}{m_{j+1}} \longrightarrow 0, \tag{3.18}$$

and set

$$S = \bigcup_{j \ge 0} \{m_{2j} + 1, \dots, m_{2j+1}\}. \tag{3.19}$$

For $s_n = \#(S \cap \{1, \dots, n\})$, the ends of the free and frozen blocks give

$$\frac{s_{m_{2j+1}}}{m_{2j+1}} \longrightarrow 1, \qquad \frac{s_{m_{2j+2}}}{m_{2j+2}} \longrightarrow 0. \tag{3.20}$$

Indeed, the first ratio is at least $1 - m_{2j}/m_{2j+1}$, and the second is at most m_{2j+1}/m_{2j+2}. Hence

$$\underline{d}(S) = 0, \qquad \overline{d}(S) = 1. \tag{3.21}$$

Consider the digit-restriction set

$$K_S = \left\{ \sum_{n=1}^{\infty} \frac{2\varepsilon_n}{3^n} \;:\; \varepsilon_n \in \{0,1\}\ (n \in S), \quad \varepsilon_n = 0\ (n \notin S) \right\}. \tag{3.22}$$

It is compact. Indeed, from any sequence of its points, successively select subsequences on which the first digit, then the second digit, and so on are constant. The diagonal subsequence converges because the tail after digit N is at most 3^{-N}, and its limit has the required fixed digits. Thus K_S is sequentially compact, hence compact as a subset of R.

At level n there are exactly 2^{S_n} permitted ternary cylinders, each of diameter at most 3^{-n}. Choosing one point in each cylinder gives a cover by 2^{S_n} balls of radius 3^{-n}, and hence

$$N_{3^{-n}}(K_S) \le 2^{S_n}. \tag{3.23}$$

The 2^{S_n} points obtained by setting every digit after n equal to zero belong to K_S. If two such points first differ at digit $k \le n$, their distance is at least

$$2\,3^{-k} - \sum_{j=k+1}^{n} 2\,3^{-j} > 3^{-k} \ge 3^{-n}.$$

They are therefore 3^{-n}-separated. Lemma 1.1 gives

$$2^{S_n} \le N_{3^{-n}/2}(K_S). \tag{3.24}$$

The discrete-scale lemma and (3.23)–(3.24) imply

$$\underline{\dim}_B K_S = \frac{\log 2}{\log 3}\underline{d}(S) = 0, \qquad \overline{\dim}_B K_S = \frac{\log 2}{\log 3}\overline{d}(S) = \frac{\log 2}{\log 3}. \tag{3.25}$$

It remains to compute packing dimension. Write $\beta = \log 2/\log 3$. Every nonempty relatively open subset U of K_S contains a permitted cylinder: for $x \in U$, choose a sufficiently long prefix of a ternary expansion representing x; all points of K_S sharing that prefix lie arbitrarily close to x. Such a cylinder is a translate of $3^{-N}K_{S_N}$, where

$$S_N = \{n - N \;:\; n \in S,\ n > N\}.$$

Every tail S_N still has upper density one, because the later free blocks in (3.19) dominate all preceding indices. The calculation (3.23)–(3.25), applied to S_N, shows that the cylinder has upper box dimension β. Consequently

$$\overline{\dim}_B U = \beta \tag{3.26}$$

for every nonempty relatively open $U \subset K_S$.

Let $K_S \subset \cup_j E_j$ be any countable cover by bounded sets, and let

$$F_j = \overline{E_j \cap K_S}^{K_S}.$$

These are closed subsets covering the complete metric space K_S. By the Baire Category Theorem, some F_j contains a nonempty relatively open set U. Closure invariance and monotonicity from Theorem 1.3 give

$$\overline{\dim}_{\mathrm{B}} E_j = \overline{\dim}_{\mathrm{B}} \overline{E_j} \geq \overline{\dim}_{\mathrm{B}} F_j \geq \overline{\dim}_{\mathrm{B}} U = \beta. \tag{3.27}$$

Thus every cover in (3.1) has supremum at least β, while the one-set cover and (3.25) give the reverse inequality. Theorem 3.2 now yields

$$\dim_{\mathrm{P}} K_S = \overline{\dim}_{\mathrm{MB}} K_S = \beta > 0 = \underline{\dim}_{\mathrm{B}} K_S. \tag{3.28}$$

This proves both directions of incomparability. □

I.2.4 Cartesian products

Give $X \times Y$ the max metric

$$d_\infty((x, y), (x', y')) = \max\{d_X(x, x'), d_Y(y, y')\}. \tag{4.1}$$

For nonempty totally bounded $E \subset X$ and $F \subset Y$, product covers and product separated sets give, for every $r > 0$,

$$N_r(E \times F) \leq N_r(E) N_r(F), \tag{4.2}$$

and

$$N_r(E \times F) \geq M_{2r}(E \times F) \geq M_{2r}(E) M_{2r}(F). \tag{4.3}$$

Indeed, a max ball is the Cartesian product of its factor balls, which proves (4.2). The product of a $2r$-separated subset of E with one of F is $2r$-separated in the max metric; Lemma 1.1 then proves (4.3).

Theorem — 4.1 (T-I.2.13: box dimensions of products)

One has

$$\begin{aligned} \underline{\dim}_{\mathrm{B}} E + \underline{\dim}_{\mathrm{B}} F &\leq \underline{\dim}_{\mathrm{B}} (E \times F) \\ &\leq \min\left\{\underline{\dim}_{\mathrm{B}} E + \overline{\dim}_{\mathrm{B}} F, \overline{\dim}_{\mathrm{B}} E + \underline{\dim}_{\mathrm{B}} F\right\}, \end{aligned} \tag{4.4}$$

and

$$\begin{aligned} \max\left\{\overline{\dim}_{\mathrm{B}} E + \underline{\dim}_{\mathrm{B}} F, \underline{\dim}_{\mathrm{B}} E + \overline{\dim}_{\mathrm{B}} F\right\} &\leq \overline{\dim}_{\mathrm{B}} (E \times F) \\ &\leq \overline{\dim}_{\mathrm{B}} E + \overline{\dim}_{\mathrm{B}} F. \end{aligned} \tag{4.5}$$

If both factors have box dimension, then the product does and

$$\dim_{\mathrm{B}} (E \times F) = \dim_{\mathrm{B}} E + \dim_{\mathrm{B}} F. \tag{4.6}$$

Proof

Divide the logarithms of (4.2) and (4.3) by $-\log r$. Lemma 1.2 says that the factor terms involving M_{2r} have the same lower and upper limits as the corresponding N_r terms.

For extended nonnegative sequences $(a_k), (b_k)$,

$$\liminf(a_k + b_k) \geq \liminf a_k + \liminf b_k, \tag{4.7}$$

$$\limsup(a_k + b_k) \leq \limsup a_k + \limsup b_k, \tag{4.8}$$

$$\limsup(a_k + b_k) \geq \max\{\limsup a_k + \liminf b_k, \liminf a_k + \limsup b_k\}, \tag{4.9}$$

and

$$\liminf(a_k + b_k) \leq \min\{\liminf a_k + \limsup b_k, \limsup a_k + \liminf b_k\}. \tag{4.10}$$

Equations (4.7)–(4.10) follow directly by bounding one sequence eventually above or below by its appropriate limiting envelope and then taking a subsequence realising the other limit. Applying (4.7) and (4.9) to the lower bound (4.3), and (4.8) and (4.10) to the upper bound (4.2), proves (4.4)–(4.5). If the lower and upper dimensions of both factors coincide, both bounds collapse to (4.6). □

We next prove the two upper product inequalities without using either of the deep lower product inequalities.

Lemma — 4.2 (bounded-box factor estimate)

Let $E \subset \mathbb{R}^m$ be arbitrary and let $F \subset \mathbb{R}^n$ be nonempty and bounded. Then

$$\dim_{\mathrm{H}}(E \times F) \leq \dim_{\mathrm{H}} E + \overline{\dim}_{\mathrm{B}} F. \tag{4.11}$$

Proof

The assertion is automatic if either term on the right is infinite. Choose

$$a > \dim_{\mathrm{H}} E, \qquad b > \overline{\dim}_{\mathrm{B}} F. \tag{4.12}$$

There is $r_0 > 0$ such that

$$N_r(F) \leq r^{-b} \qquad (0 < r \leq r_0). \tag{4.13}$$

We record explicitly the ball-cover consequence of $\mathscr{H}^a(E) = 0$. Given $\eta > 0$ and $\varepsilon > 0$, there is a countable cover of E by closed balls

$$(\overline{B}(x_i, r_i))_i, \qquad 0 < r_i \leq \min\{\eta, r_0\}, \qquad \sum_i r_i^a < \epsilon. \tag{4.14}$$

To obtain it, start with a countable cover (U_i) of sufficiently small diameter and with $\sum_i (\operatorname{diam} U_i)^a < \varepsilon/2$; discard sets missing E and choose $x_i \in U_i \cap E$. When $\operatorname{diam} U_i > 0$, use the ball centred at x_i of radius $\operatorname{diam} U_i$. Each zero-diameter nonempty member meets E in a singleton; give these countably many singletons positive radii whose a-powers sum to less than $\varepsilon/2$. The radii can individually be made at most $\min\{\eta, r_0\}$. This proves (4.14).

For each i, cover F by $N_{r_i}(F)$ closed radius-r_i balls. Their products with $\overline{B}(x_i, r_i)$ cover $E \times F$; in the max metric each product has diameter at most $2r_i$. Equations (4.13)–(4.14) give total $(a+b)$-cost at most

$$\sum_i N_{r_i}(F)(2r_i)^{a+b} \le 2^{a+b} \sum_i r_i^a < 2^{a+b}\epsilon. \tag{4.15}$$

The covering diameters are at most 2η. Letting $\varepsilon \downarrow 0$ shows $\mathscr{H}^{a+b}_{2\eta}(E \times F) = 0$; then letting $\eta \downarrow 0$ gives $\mathscr{H}^{a+b}(E \times F) = 0$. Hence $\dim_{\mathrm{H}}(E \times F) \le a + b$. Let a and b decrease to the values in (4.12) to obtain (4.11). □

Theorem — 4.3 (T-I.2.14: upper Hausdorff product inequality)

For arbitrary $E \subset \mathbb{R}^m$ and $F \subset \mathbb{R}^n$,

$$\dim_{\mathrm{H}}(E \times F) \le \dim_{\mathrm{H}} E + \dim_{\mathrm{P}} F. \tag{4.16}$$

The analogous inequality with the factors exchanged also holds.

Proof

Empty factors are immediate, and an infinite term on the right makes the claim automatic. Put $p = \dim_{\mathrm{P}} F < +\infty$ and take $b > p$. By Theorem 3.2 and the definition (3.1), there is a cover $F \subset \cup_j F_j$ by bounded sets with

$$\sup_j \overline{\dim}_{\mathrm{B}} F_j < b. \tag{4.17}$$

Since $E \times F \subset \cup_j (E \times F_j)$, countable stability of Hausdorff dimension and Lemma 4.2 give

$$\begin{aligned} \dim_{\mathrm{H}}(E \times F) &\le \sup_j \dim_{\mathrm{H}}(E \times F_j) \\ &\le \dim_{\mathrm{H}} E + \sup_j \overline{\dim}_{\mathrm{B}} F_j \\ &\le \dim_{\mathrm{H}} E + b. \end{aligned} \tag{4.18}$$

Letting $b \downarrow p$ proves (4.16). The coordinate swap is an isometry for the max metric, so applying (4.16) after swapping proves the symmetric version. □

Theorem — 4.4 (T-I.2.15: upper packing product inequality)

For arbitrary $E \subset \mathbb{R}^m$ and $F \subset \mathbb{R}^n$,

$$\dim_{\mathrm{P}}(E \times F) \le \dim_{\mathrm{P}} E + \dim_{\mathrm{P}} F. \tag{4.19}$$

Proof

The result is immediate for an empty factor or when the right side is infinite. Put $p = \dim_P E$, $q = \dim_P F$, and fix $\varepsilon > 0$. Theorem 3.2 and (3.1) give bounded covers

$$E \subset \bigcup_i E_i, \qquad F \subset \bigcup_j F_j, \tag{4.20}$$

such that

$$\sup_i \overline{\dim}_B E_i < p + \epsilon, \qquad \sup_j \overline{\dim}_B F_j < q + \epsilon. \tag{4.21}$$

The countable family $(E_i \times F_j)_{i,j}$ covers $E \times F$. The upper bound in Theorem 4.1 gives

$$\overline{\dim}_B(E_i \times F_j) \le \overline{\dim}_B E_i + \overline{\dim}_B F_j < p + q + 2\epsilon. \tag{4.22}$$

Taking this family in the modified upper box definition and applying Theorem 3.2 yields

$$\dim_P(E \times F) < p + q + 2\epsilon.$$

Let $\varepsilon \downarrow 0$. $\square$

I.2.5 Citation-owned lower bounds and sharpness

No assertion in this section is used in an F proof above. The statements are included to make the mathematical frontier of the chapter explicit.

Guided result — 5.1 (G-I.2.16: Howroyd product chain; no full proof here)

Let E, F be nonempty separable metric spaces, and equip their product with the metric d_2 from D-I.2.08. Then

$$\dim_H E + \dim_H F \le \dim_H(E \times F) \le \dim_H E + \dim_P F \le \dim_P(E \times F) \le \dim_P E + \dim_P F. \tag{5.1}$$

For Euclidean subsets, the second and fourth comparisons are Theorems 4.3 and 4.4. Their extensions to the metric spaces in the present statement, as well as the first and third comparisons, are cited without proof here.

Reference. The original product chain is due to Howroyd [How96, pp. 715–727]. Its exact form and metric-space extension are restated in Wei–Wen–Wen [WWW16, pp. 1–2, equations (1)–(2)].

Guided result — 5.2 (G-I.2.17: sharp Hausdorff products; no full proof here)

If $\alpha, \beta, \gamma, \lambda > 0$, $\beta \le \gamma$, and

$$\alpha + \beta \le \lambda \le \alpha + \gamma,$$

there are compact metric spaces E, F such that

$$\dim_{\mathrm{H}} E = \alpha, \quad \dim_{\mathrm{H}} F = \beta, \quad \dim_{\mathrm{P}} F = \gamma, \quad \dim_{\mathrm{H}}(E \times F) = \lambda. \tag{5.2}$$

Reference. Wei, Wen, and Wen [WWW16, Theorem 1, p. 2].

Guided result — 5.3 (G-I.2.18: sharp packing products; no full proof here)

If $\alpha, \beta, \gamma, \lambda > 0$, $\alpha \leq \gamma$, and

$$\alpha + \beta \leq \lambda \leq \gamma + \beta,$$

there are compact metric spaces E, F such that

$$\dim_{\mathrm{H}} E = \alpha, \quad \dim_{\mathrm{P}} F = \beta, \quad \dim_{\mathrm{P}} E = \gamma, \quad \dim_{\mathrm{P}}(E \times F) = \lambda. \tag{5.3}$$

Reference. Wei, Wen, and Wen [WWW16, Theorem 2, p. 2].

Guided result — 5.4 (G-I.2.19: sharp upper-box products; no full proof here)

If $\alpha, \beta, \gamma, \lambda > 0$, $\alpha \leq \gamma$, and

$$\alpha + \beta \leq \lambda \leq \gamma + \beta,$$

there are compact metric spaces E, F such that

$$\underline{\dim}_{\mathrm{B}} E = \alpha, \quad \overline{\dim}_{\mathrm{B}} F = \beta, \quad \overline{\dim}_{\mathrm{B}} E = \gamma, \quad \overline{\dim}_{\mathrm{B}}(E \times F) = \lambda. \tag{5.4}$$

Reference. Wei, Wen, and Wen [WWW16, Theorem 3, p. 2].

I.2.5.1 Construction map for Theorems 5.2–5.4 (not a proof)

For $S \subset \mathbb{N}$, the binary digit-restriction set is

$$E_S = \left\{ \sum_{n=1}^{\infty} \varepsilon_n 2^{-n} \, : \, \varepsilon_n \in \{0, 1\} \, (n \in S), \quad \varepsilon_n = 0 \, (n \notin S) \right\}.$$

Wei–Wen–Wen use these binary digit-restriction sets, encode their Hausdorff, packing, lower box, and upper box dimensions by lower and upper densities of index sets, and choose alternating blocks whose endpoints satisfy $k_j/k_{j+1} \to 0$. Synchronising two index sets makes the liminf or limsup of the sum of their densities attain the prescribed λ; finite Cartesian powers clear the restriction that the intermediate densities lie in $(0, 1)$. The quantitative density lemmas and the three parameter choices occupy pp. 3–12 of the cited paper. This paragraph is orientation only and does not replace those proofs.

I.2.6 Intuition and strategy

The following discussion explains the geometric ideas behind the proofs.

I.2.6.1 Equal-scale versus variable-scale geometry

Box dimension asks how many pieces are needed when every piece is examined at one common scale. This makes it computable, but also makes it sensitive to closure and destroys countable stability: a countable dense set looks as large as its closure at every positive resolution.

Packing measure allows every ball to choose its own small radius. A supremum over disjoint balls is the covering/packing dual of Hausdorff content. The raw supremum is still too rigid to be a measure. The second operation, covering the set by pieces and summing their zero-scale premeasures, is what restores countable subadditivity.

The notation must preserve this two-stage construction:

$$\mathscr{P}^s_\delta \longrightarrow \mathscr{P}^s_0 \longrightarrow \mathscr{P}^s.$$

Confusing the last two objects is a structural error, not a harmless abuse.

I.2.6.2 Why modified upper box dimension appears

If a bounded set has upper box dimension below s, every variable-radius packing can be sorted into dyadic radius classes. There are not enough balls in any class to keep the total s-cost positive. If the upper box dimension is above s, a sequence of especially efficient equal-radius packings has unbounded s-cost. Thus finite s-packing premeasure forces upper box dimension at most s.

Packing outer measure is obtained by covering with pieces before summing. Accordingly, its critical exponent is the best upper box exponent obtainable after a countable decomposition. That is exactly modified upper box dimension.

I.2.6.3 The hierarchy is not a total order

For bounded Euclidean sets,

$$\dim_{\mathrm{H}} E \le \min\{\underline{\dim}_{\mathrm{B}} E, \dim_{\mathrm{P}} E\} \le \max\{\underline{\dim}_{\mathrm{B}} E, \dim_{\mathrm{P}} E\} \le \overline{\dim}_{\mathrm{B}} E.$$

There is no fixed order between the two middle quantities. Dense countable sets make lower box dimension larger. Alternating free and frozen digit blocks make packing dimension larger. These two tests are useful whenever a new proposed dimension is introduced: test countable density and test oscillating digit frequencies.

I.2.6.4 Product synchronisation

At a fixed scale, a product cover is obtained by multiplying factor covers, and a product separated set is obtained by multiplying factor separated sets. The logarithm converts these products to sums. What prevents a universal exact formula is that liminf and limsup may be attained along different scale subsequences in the two factors.

The sharp examples of Wei–Wen–Wen deliberately synchronise or desynchronise long digit blocks. Their result says that the elementary product interval is not merely a proof artefact: every value allowed by the inequalities can occur.

I.2.6.5 Boundary between elementary and deep product estimates

The upper estimate

$$\dim_{\mathrm{H}}(E \times F) \leq \dim_{\mathrm{H}} E + \dim_{\mathrm{P}} F$$

is a covering argument. Decompose F into pieces with controlled upper box dimension, use a Hausdorff cover of E, and cover F at the radius assigned to each member of that cover.

The lower estimates require the existence or distribution of measures with controlled small-ball behaviour. Those ideas belong naturally with Frostman theory in I.3. In this chapter the deeper product estimates are explicitly cited without proof; they are not consequences of the elementary covering argument just described.

I.2.7 Exercises

Every exercise uses only D/R-I.1, the F portion of I.2, and A-RA.

I.2.7.1 Exercise 1 — Covers with ambient centres

Let E be nonempty and totally bounded. Define $N_r'(E)$ as in D-I.2.02, but allow the centres of the covering balls to lie anywhere in the ambient metric space. Prove

$$N_r'(E) \leq N_r(E) \leq N_{r/2}'(E),$$

and deduce that N_r' gives the same lower and upper box dimensions.

I.2.7.2 Exercise 2 — Geometric sampling of scales

Let E be a nonempty totally bounded metric set and fix $0 < c < 1$. Prove directly that both box dimensions may be computed from $r_k = c^k$. For $r_k = 2^{-k^2}$, explain why Lemma 1.2 does not apply directly, but show that sampling is nevertheless valid by comparing successive logarithmic scales.

I.2.7.3 Exercise 3 — A polynomial sequence

For $\alpha > 0$, let

$$E_\alpha = \{0\} \cup \{n^{-\alpha} : n \in \mathbb{N}\}.$$

Prove that the box dimension exists and

$$\dim_{\mathrm{B}} E_\alpha = \frac{1}{1+\alpha}.$$

I.2.7.4 Exercise 4 — The packing premeasure is not an outer measure

Let $D = Q \cap [0,1]$. Prove

$$\mathscr{P}_0^1(D) = 1, \qquad \mathscr{P}_0^1(\{x\}) = 0 \quad (x \in D).$$

Conclude that P_0^1 is not countably subadditive, whereas $P^1(D) = 0$.

I.2.7.5 Exercise 5 — General digit restrictions

For arbitrary $S \subset N$, prove

$$\underline{\dim}_{\mathrm{B}} K_S = \frac{\log 2}{\log 3}\underline{d}(S), \qquad \overline{\dim}_{\mathrm{B}} K_S = \frac{\log 2}{\log 3}\overline{d}(S),$$

and

$$\dim_{\mathrm{P}} K_S = \overline{\dim}_{\mathrm{B}} K_S.$$

I.2.7.6 Exercise 6 — Countable modifications

If $E \subset \mathbb{R}^d$ and $C \subset \mathbb{R}^d$ is countable, prove

$$\dim_{\mathrm{P}}(E \cup C) = \dim_{\mathrm{P}} E.$$

Give an example showing that the analogous assertion can fail for both box dimensions.

I.2.7.7 Exercise 7 — Hölder maps and box dimension

Let E be a nonempty totally bounded subset of a metric space (X, d_X), let (Y, d_Y) be a metric space, and let $0 \le L < \infty$. Suppose $f : E \to Y$ satisfies

$$d_Y(f(x), f(y)) \le L d_X(x, y)^\theta, \qquad 0 < \theta \le 1.$$

Prove

$$\underline{\dim}_{\mathrm{B}} f(E) \le \theta^{-1}\underline{\dim}_{\mathrm{B}} E, \qquad \overline{\dim}_{\mathrm{B}} f(E) \le \theta^{-1}\overline{\dim}_{\mathrm{B}} E.$$

I.2.7.8 Exercise 8 — Products with intervals

If a nonempty bounded set E has box dimension s, prove

$$\dim_{\mathrm{B}}(E \times [0,1]^k) = s + k$$

for every positive integer k, using the max metric.

I.2.7.9 Exercise 9 — Product oscillation bounds

Suppose

$$\underline{\dim}_{\mathrm{B}} E = a, \quad \overline{\dim}_{\mathrm{B}} E = A, \quad \underline{\dim}_{\mathrm{B}} F = b, \quad \overline{\dim}_{\mathrm{B}} F = B.$$

Assume these four numbers are finite.

Write the sharp intervals supplied by T-I.2.13 for the lower and upper box dimensions of $E \times F$. Determine when each interval collapses to one point.

I.2.7.10 Exercise 10 — A one-sided Hausdorff product estimate

Let $E \subset \mathbb{R}^m$ and let $F \subset \mathbb{R}^n$ be nonempty and bounded with $\overline{\dim}_{\mathrm{B}} F = 0$. Prove directly, without invoking T-I.2.14, that

$$\dim_{\mathrm{H}}(E \times F) \leq \dim_{\mathrm{H}} E.$$

I.2.7.11 Exercise 11 — Bi-Lipschitz product metrics

Prove

$$d_\infty \leq d_2 \leq \sqrt{2}\, d_\infty$$

on a product of two metric spaces. Deduce that all dimensions proved bi-Lipschitz invariant in I.1–I.2 have the same value for the two metrics in the Euclidean setting.

I.2.7.12 Exercise 12 — Audit the critical exponent

Let $p = \dim_{\mathrm{P}} E < +\infty$. Show that each of the three possibilities

$$\mathscr{P}^p(E) = 0, \qquad 0 < \mathscr{P}^p(E) < +\infty, \qquad \mathscr{P}^p(E) = +\infty$$

is logically compatible with the definition and the critical-jump theorem. Identify the exact step at which an attempted proof of a universal endpoint value would fail.

Chapter I.3

Frostman Measures, Energies, and Capacities

A lower bound for dimension can be encoded in a measure. A Frostman measure limits the mass in every small ball; an energy integral instead controls the average interaction between pairs of points. We develop the passage between these descriptions, formulate the associated capacity, and apply the energy method to orthogonal projections. Keep track of the strict inequalities between exponents: they are what make the relevant scale integrals converge.

The definitions below fix this chapter's conventions. Geometric intuition and proof strategy are discussed separately after the main development; exercises and their solutions provide a second route through the material.

Definitions and notation

All Euclidean norms are denoted by $|.|$. Hausdorff measures retain the diameter normalisation of I.1. Fourier transforms use the 2π convention fixed below.

Definition — I.3.01: Unrestricted Hausdorff content

For $E \subset \mathbb{R}^d$ and $s > 0$, define

$$\mathscr{H}^s_\infty(E) = \inf\left\{\sum_{j=1}^{\infty}(\operatorname{diam} U_j)^s \;:\; E \subset \bigcup_{j=1}^{\infty} U_j\right\}.$$

There is no diameter restriction on the covering sets. The empty cover has cost zero and covers only the empty set.

Definition — I.3.02: Frostman measure

Let $E \subset \mathbb{R}^d$ be Borel, $s > 0$, and $0 < \mu(\mathbb{R}^d) < +\infty$. A finite Borel measure μ carried by E is an s-**Frostman measure** if there is $C < +\infty$ such that

$$\mu(\overline{B}(x,r)) \le Cr^s \qquad (x \in \mathbb{R}^d,\ r > 0).$$

The least admissible C is the s-**Frostman constant** of μ, with value $+\infty$ if no finite constant exists. When $\mu(E) = 1$, it is an s-Frostman probability measure.

Definition — I.3.03: Riesz kernel, potential, and energy

Let $0 < s < d$. The s-**Riesz kernel** is

$$k_s(x) = |x|^{-s} \quad (x \neq 0), \qquad k_s(0) = +\infty.$$

For a finite Borel measure μ on $\mathbb{R}^d$, its s-**Riesz potential** is

$$U_s^\mu(x) = \int_{\mathbb{R}^d} |x-y|^{-s}\, d\mu(y) \in [0,+\infty],$$

and its s-**Riesz energy** is

$$I_s(\mu) = \iint_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^{-s}\, d\mu(x) d\mu(y) = \int U_s^\mu\, d\mu.$$

All integrals are nonnegative extended integrals.

Definition — I.3.04: Capacity and equilibrium measure

For a Borel set $E \subset \mathbb{R}^d$, let $\mathrm{Prob}(E)$ be the class of Borel probability measures carried by E. For nonempty compact E and $0 < s < d$, define

$$\mathrm{Cap}_s(E) = \left(\inf_{\mu\in\mathrm{Prob}(E)} I_s(\mu)\right)^{-1},$$

with $(+\infty)^{-1} = 0$ and $0^{-1} = +\infty$. Set $\mathrm{Cap}_s(\varnothing) = 0$. A measure $\mu_E \in \mathrm{Prob}(E)$ is an s-**equilibrium measure** if

$$I_s(\mu_E) = \inf_{\mu\in\mathrm{Prob}(E)} I_s(\mu).$$

Definition — I.3.05: Fourier and Gamma normalisation

For a finite Borel measure μ on $\mathbb{R}^d$, define

$$\hat{\mu}(\xi) = \int_{\mathbb{R}^d} e^{-2\pi i x\cdot\xi}\, d\mu(x), \qquad \xi \in \mathbb{R}^d.$$

For $a > 0$, define the Gamma function by

$$\Gamma(a) = \int_0^\infty u^{a-1} e^{-u}\, du.$$

For $0 < s < d$, set

$$c_{d,s} = \pi^{s-d/2} \frac{\Gamma((d-s)/2)}{\Gamma(s/2)}.$$

Definition — I.3.06: Orthogonal projections and direction measure

Let

$$S^1 = \{e \in \mathbb{R}^2 \,:\, |e| = 1\}.$$

For $e \in S^1$, the scalar orthogonal projection in direction e is

$$\pi_e : \mathbb{R}^2 \to \mathbb{R}, \qquad \pi_e(x) = x \cdot e.$$

For a finite Borel measure μ, write

$$\mu_e = (\pi_e)_\# \mu, \qquad \mu_e(A) = \mu(\pi_e^{-1}(A))$$

for its projected measure. The symbol σ denotes normalized arclength probability measure on S^1. A statement holding for **almost every direction** means outside a σ-null subset of S^1.

Throughout, Hausdorff measures use diameter normalisation. All measures are nonnegative Borel measures. The ambient Euclidean dimension is denoted by d unless the planar case is stated.

I.3.1 Hausdorff content and compactness of measures

For $E \subset \mathbb{R}^d$ and $s > 0$, define the unrestricted content

$$\mathscr{H}^s_\infty(E) = \inf\left\{ \sum_j (\operatorname{diam} U_j)^s \,:\, E \subset \bigcup_j U_j \right\}. \tag{1.1}$$

Unlike $\mathscr{H}^s_\delta$, the infimum in (1.1) imposes no upper bound on the covering diameters.

Lemma — 1.1 (L-I.3.01: content detects positivity)

For every $E \subset \mathbb{R}^d$ and $s > 0$,

$$\mathscr{H}^s_\infty(E) = 0 \iff \mathscr{H}^s(E) = 0. \tag{1.2}$$

Proof

Every δ-cover is an unrestricted cover, so

$$\mathcal{H}^s_\infty(E) \le \mathcal{H}^s_\delta(E) \le \mathcal{H}^s(E). \tag{1.3}$$

Thus $\mathcal{H}^s(E) = 0$ implies $\mathcal{H}^s_\infty(E) = 0$.

Conversely suppose $\mathcal{H}^s_\infty(E) = 0$. Fix $\delta > 0$ and $\varepsilon > 0$. Choose an unrestricted cover (U_j) with

$$\sum_j (\operatorname{diam} U_j)^s < \min\{\epsilon, \delta^s\}. \tag{1.4}$$

Every nonempty member of this cover has diameter strictly less than δ, so it is a δ-cover and

$$\mathcal{H}^s_\delta(E) < \epsilon.$$

Since ε was arbitrary, $\mathcal{H}^s_\delta(E) = 0$. This holds for every δ, hence $\mathcal{H}^s(E) = 0$. □

For a Borel set E, write $\operatorname{Prob}(E)$ for the Borel probability measures carried by E. When E is compact, a sequence (μ_j) in $\operatorname{Prob}(E)$ converges **weakly** to $\mu \in \operatorname{Prob}(E)$ if

$$\int_E f\, d\mu_j \longrightarrow \int_E f\, d\mu \qquad (f \in C(E)). \tag{1.5}$$

Lemma — 1.2 (L-I.3.02: compactness and lower semicontinuity)

Let $E \subset \mathbb{R}^d$ be compact.
(1) Every sequence in $\operatorname{Prob}(E)$ has a weakly convergent subsequence whose limit belongs to $\operatorname{Prob}(E)$.
(2) If $0 < s < d$ and μ_j converges weakly to μ, then

$$I_s(\mu) \le \liminf_{j\to\infty} I_s(\mu_j). \tag{1.6}$$

Proof

Because E is compact metric, $C(E)$ is separable in the uniform norm. Choose a dense sequence $(f_k)_{k\ge1}$. Given (μ_j), the sequence $(\int f_1 d\mu_j)_j$ is bounded; select a convergent subsequence. From it select a subsequence on which the integrals of f_2 converge, and continue. The diagonal subsequence, denoted again by (μ_j), has convergent integrals for every f_k.

For arbitrary $f \in C(E)$, uniform approximation by some f_k shows that $(\int f d\mu_j)_j$ is Cauchy. Define

$$L(f) = \lim_{j\to\infty} \int f\, d\mu_j.$$

Then L is a positive linear functional on $C(E)$, $L(1) = 1$, and $|L(f)| \le \|f\|_\infty$. The Riesz Representation Theorem supplies a unique Borel probability measure μ on E such that $L(f) = \int f d\mu$. This is (1.5) and proves the first assertion.

For the second, define on $E \times E$

$$K_N(x, y) = \min\{N, |x - y|^{-s}\}, \qquad K_N(x, x) = N. \tag{1.7}$$

The function K_N is bounded and continuous. Weak convergence implies

$$\iint K_N \, d\mu_j d\mu_j \longrightarrow \iint K_N \, d\mu d\mu. \tag{1.8}$$

To justify the product convergence in (1.8), first check it for functions $g(x)h(y)$, then for finite sums of such products, and use their uniform density in $C(E \times E)$. Finally, monotone convergence gives

$$I_s(\mu) = \sup_N \iint K_N \, d\mu d\mu \le \liminf_j I_s(\mu_j),$$

which is (1.6). □

I.3.2 A finite-tree proof of Frostman's lemma

Fix a half-open cube Q_0 of side length L whose interior contains the compact set E. Its 2^d half-open dyadic children partition it; iterating gives the level-k descendants $D_k(Q_0)$, each of side length $L2^{-k}$. The half-open convention assigns each point of Q_0 to exactly one cube at every level.

Lemma — 2.1 (L-I.3.03: finite dyadic Frostman flow)

Let $s > 0$ and

$$h = \mathscr{H}^s_\infty(E) > 0.$$

For every $n \ge 1$, there is a finitely supported $\nu_n \in \mathrm{Prob}(E)$ such that

$$\nu_n(Q) \le d^{s/2} h^{-1} \ell(Q)^s \tag{2.1}$$

for every $Q \in D_k(Q_0)$ and every $0 \le k \le n$.

Proof

On the finite tree of descendants through level n, define a number $w(Q)$ from the leaves upward. At level n, put

$$w(Q) = \begin{cases} \ell(Q)^s, & Q \cap E \ne \emptyset, \\ 0, & Q \cap E = \emptyset. \end{cases} \tag{2.2}$$

At levels $k < n$, put

$$w(Q) = \min\left\{\ell(Q)^s, \sum_{R \text{ child of } Q} w(R)\right\}. \tag{2.3}$$

Induction from the leaves shows that $w(Q)$ is the least sum of $\ell(R)^s$ over a family of descendants R of Q, at levels between k and n, whose union contains every occupied level-n descendant of Q. Indeed, a least cover either uses Q itself, at cost $\ell(Q)^s$, or does not use it and then splits into independent least covers inside its children.

Every such family for Q_0 covers E. Since a dyadic cube R has diameter $\sqrt{d}\ell(R)$, (1.1) gives

$$h \le d^{s/2} \sum_R \ell(R)^s.$$

Taking the least such sum yields

$$w(Q_0) \ge d^{-s/2}h. \tag{2.4}$$

We now send a mass $w(Q_0)$ down the tree. Give Q_0 incoming mass $m(Q_0) = w(Q_0)$. If Q is not a leaf and has incoming mass $m(Q) \le w(Q)$, distribute that mass among its children so that

$$0 \le m(R) \le w(R), \qquad \sum_{R \text{ child of } Q} m(R) = m(Q). \tag{2.5}$$

This is possible because (2.3) gives $m(Q) \le w(Q) \le \sum_R w(R)$; one may fill the children successively up to their capacities. Empty leaves receive zero mass. For every occupied leaf Q, choose $x_Q \in E \cap Q$ and place an atom of mass $m(Q)$ at x_Q. Let the resulting finite measure be λ_n. Conservation in (2.5) gives

$$\lambda_n(E) = w(Q_0), \qquad \lambda_n(Q) = m(Q) \le w(Q) \le \ell(Q)^s \tag{2.6}$$

for every node through level n. Normalize by $\nu_n = \lambda_n / w(Q_0)$. Combining (2.4) and (2.6) proves (2.1). □

An s-Frostman probability measure on E is a member $\mu \in \mathrm{Prob}(E)$ for which

$$\mu(\overline{B}(x,r)) \le Cr^s \qquad (x \in \mathbb{R}^d,\ r > 0) \tag{2.7}$$

with a constant $C < +\infty$.

Theorem — 2.2 (T-I.3.04: compact Euclidean Frostman lemma)

Let $E \subset \mathbb{R}^d$ be compact and $s > 0$. Then

$$\mathscr{H}^s(E) > 0 \iff E \text{ carries an } s\text{-Frostman probability measure.} \tag{2.8}$$

If $h = \mathscr{H}^s_\infty(E) > 0$, the measure can be chosen so that

$$\mu(\overline{B}(x,r)) \le C_{d,s} h^{-1} r^s \qquad (x \in \mathbb{R}^d,\ r > 0). \tag{2.9}$$

Conversely, a Frostman probability with constant C satisfies

$$\mathscr{H}^s(E) \geq C^{-1}. \tag{2.10}$$

Proof

Assume first that $\mathscr{H}^s(E) > 0$. Lemma 1.1 gives $h = \mathscr{H}^s_\infty(E) > 0$. Let (ν_n) be the measures from Lemma 2.1. By Lemma 1.2, a subsequence converges weakly to some $\mu \in \mathrm{Prob}(E)$.

Fix an open ball $B(x, r)$ with $0 < r \leq L$. Choose a dyadic level k such that

$$r \leq L2^{-k} < 2r. \tag{2.11}$$

The ball meets at most 4^d half-open cubes of that level. For every $n \geq k$, (2.1) gives

$$\nu_n(B(x, r)) \leq 4^d d^{s/2} h^{-1} (L2^{-k})^s \leq 4^d 2^s d^{s/2} h^{-1} r^s. \tag{2.12}$$

Weak convergence has the open-set inequality

$$\mu(G) \leq \liminf_n \nu_n(G) \tag{2.13}$$

for open G. For completeness, (2.13) follows by putting a compact subset of G under a continuous function supported in G, applying (1.5), and then using inner regularity. Equations (2.12)–(2.13) bound every open ball of radius at most L.

If $0 < r \leq L/2$, then

$$\mu(\overline{B}(x, r)) \leq \mu(B(x, 2r)) \leq 4^d 2^{2s} d^{s/2} h^{-1} r^s. \tag{2.14}$$

For $r > L/2$, use $\mu(\mathbb{R}^d) = 1$ and $h \leq \mathrm{diam}(E)^s \leq d^{s/2} L^s$ to enlarge the dimensional constant in (2.14). This proves (2.9) for all $r > 0$, and hence the forward implication in (2.8).

Conversely suppose μ is an s-Frostman probability on E with constant C. The mass-distribution principle T-I.1.08, applied with the power gauge r^s, gives

$$\mathscr{H}^s(E) \geq \frac{\mu(E)}{C} = C^{-1}.$$

This is (2.10) and completes the proof. □

I.3.3 Riesz potentials, energies, and capacity

For $0 < s < d$, define

$$U_s^\mu(x) = \int |x - y|^{-s} \, d\mu(y), \qquad I_s(\mu) = \int U_s^\mu(x) \, d\mu(x), \tag{3.1}$$

where the kernel equals $+\infty$ on the diagonal. These are nonnegative extended integrals.

Lemma — 3.1 (L-I.3.05: layer-cake identity)

For every finite Borel measure μ on $\mathbb{R}^d$ and every $0 < s < d$,

$$I_s(\mu) = s\int_0^\infty r^{-s-1}\left(\int \mu(\overline{B}(x,r))\,d\mu(x)\right)dr. \tag{3.2}$$

Proof

For $z \in \mathbb{R}^d$, with both sides interpreted as $+\infty$ when $z = 0$,

$$|z|^{-s} = s\int_{|z|}^\infty r^{-s-1}\,dr = s\int_0^\infty r^{-s-1}\mathbf{1}_{\{|z|\le r\}}\,dr. \tag{3.3}$$

Apply (3.3) with $z = x - y$ and use Tonelli's theorem twice. The inner integral in y is $\mu(\overline{B}(x,r))$, which gives (3.2). □

Corollary — 3.2 (C-I.3.06: Frostman growth gives finite lower energy)

Let $t > 0$ and $0 < s < \min\{t,d\}$, and let μ be a t-Frostman probability measure on $\mathbb{R}^d$ with constant C. Then

$$I_s(\mu) \le 1 + \frac{sC}{t-s} < +\infty. \tag{3.4}$$

Proof

The Frostman estimate and the trivial probability bound give

$$\int \mu(\overline{B}(x,r))\,d\mu(x) \le \begin{cases} Cr^t, & 0 < r \le 1, \\ 1, & r > 1. \end{cases} \tag{3.5}$$

Insert (3.5) into (3.2):

$$\begin{aligned} I_s(\mu) &\le sC\int_0^1 r^{t-s-1}\,dr + s\int_1^\infty r^{-s-1}\,dr \\ &= \frac{sC}{t-s} + 1. \end{aligned}$$

□

Corollary — 3.3 (C-I.3.07: finite energy gives positive measure)

Let $0 < s < d$, let $E \subset \mathbb{R}^d$ be Borel, and let $\mu \in \mathrm{Prob}(E)$. If $I_s(\mu) < +\infty$, then

$$\mathscr{H}^s(E) > 0, \qquad \dim_{\mathrm{H}} E \ge s. \tag{3.6}$$

More precisely, there is a Borel set $F \subset E$ with $\mu(F) > 0$ such that $\nu = \mu|F$ obeys

$$\nu(\overline{B}(x,r)) \le Mr^s \qquad (x \in F,\ r > 0) \tag{3.7}$$

for some finite M.

Proof

Since

$$\int U_s^\mu \, d\mu = I_s(\mu) < +\infty,$$

the potential is finite at μ-almost every point. It is Borel measurable as the monotone limit of the integrals of the bounded continuous kernels from (1.7). Hence for some $M < +\infty$, the Borel set

$$F = E \cap \{x : U_s^\mu(x) \le M\} \tag{3.8}$$

has positive μ-measure. If $x \in F$ and $r > 0$, then

$$\begin{aligned} M &\ge U_s^\mu(x) \\ &\ge \int_{\overline{B}(x,r)} |x-y|^{-s} \, d\mu(y) \\ &\ge r^{-s}\mu(\overline{B}(x,r)). \end{aligned} \tag{3.9}$$

Since $\nu \le \mu$, equation (3.9) gives (3.7). The mass-distribution principle T-I.1.08 applied to ν, carried by F, yields

$$\mathscr{H}^s(F) \ge \frac{\nu(F)}{M} > 0.$$

Monotonicity gives $\mathscr{H}^s(E) > 0$; the critical-exponent characterization in I.1 then gives $\dim_{\mathrm{H}} E \ge s$. □

For a nonempty compact $E \subset \mathbb{R}^d$ and $0 < s < d$, define

$$\mathrm{Cap}_s(E) = \left(\inf_{\mu \in \mathrm{Prob}(E)} I_s(\mu) \right)^{-1}, \tag{3.10}$$

with the reciprocal conventions in D-I.3.04, and set the capacity of the empty set equal to zero.

Theorem — 3.4 (T-I.3.08: energy and capacity characterize dimension)

If $E \subset \mathbb{R}^d$ is nonempty and compact, then

$$\dim_{\mathrm{H}} E = \sup\left(\{s \in (0,d) : \exists \mu \in \mathrm{Prob}(E),\ I_s(\mu) < +\infty\} \cup \{0\}\right) \tag{3.11}$$

and

$$\dim_{\mathrm{H}} E = \sup\left(\{s \in (0,d) : \mathrm{Cap}_s(E) > 0\} \cup \{0\}\right). \tag{3.12}$$

Proof

Let $\alpha = \dim_{\mathrm{H}} E$. If $0 < s < \alpha$, choose t with

$$s < t < \alpha.$$

The Hausdorff critical-exponent theorem gives $\mathscr{H}^t(E) = +\infty$. Frostman's lemma supplies a t-Frostman probability measure μ on E, and Corollary 3.2 gives $I_s(\mu) < +\infty$.

Conversely, if some $\mu \in \mathrm{Prob}(E)$ has finite s-energy, Corollary 3.3 gives $\alpha \geq s$. Therefore the first supremum is exactly α, including $\alpha = 0$ and $\alpha = d$.

By (3.10), $\mathrm{Cap}_s(E) > 0$ exactly when the infimum of the energies is finite, which is equivalent to the existence of one probability measure of finite s-energy. Thus (3.12) is the same formula as (3.11). Neither argument determines the capacity at $s = \alpha$. □

Theorem — 3.5 (T-I.3.09: capacity calculus and equilibrium measures)

Let $0 < s < d$ and let $E, F \subset \mathbb{R}^d$ be compact.

(1) If $E \subset F$, then

$$\mathrm{Cap}_s(E) \leq \mathrm{Cap}_s(F). \tag{3.13}$$

(2) If a is a nonzero real scalar and $b \in \mathbb{R}^d$, then

$$\mathrm{Cap}_s(aE + b) = |a|^s \,\mathrm{Cap}_s(E). \tag{3.14}$$

(3) Every nonempty compact E has an s-equilibrium measure.

Proof

If $E \subset F$, then $\mathrm{Prob}(E) \subset \mathrm{Prob}(F)$. The infimum in (3.10) for F is therefore no larger than the one for E; taking reciprocals proves (3.13).

Let $T(x) = ax + b$. Push-forward by T is a bijection from $\mathrm{Prob}(E)$ onto $\mathrm{Prob}(aE + b)$. For $\mu \in \mathrm{Prob}(E)$,

$$\begin{aligned} I_s(T_\# \mu) &= \iint |T(x) - T(y)|^{-s} \, d\mu(x) d\mu(y) \\ &= |a|^{-s} I_s(\mu). \end{aligned} \tag{3.15}$$

Taking infima and reciprocals gives (3.14), including the endpoint reciprocal conventions.

Finally, put

$$m = \inf_{\mu \in \mathrm{Prob}(E)} I_s(\mu).$$

If $m = +\infty$, every member of $\mathrm{Prob}(E)$ is a minimizer. If m is finite, choose a minimizing sequence (μ_j). Lemma 1.2 gives a weakly convergent subsequence with limit $\mu \in \mathrm{Prob}(E)$, and lower semicontinuity gives

$$I_s(\mu) \leq \liminf_j I_s(\mu_j) = m.$$

The reverse inequality follows from the definition of m, so equality holds. Thus μ is an equilibrium measure. When the capacity is positive, the minimum is finite and equals $\mathrm{Cap}_s(E)^{-1}$. □

I.3.4 Fourier representation of Riesz energy

We first prove the only Fourier calculation needed in this chapter. It is included inside the proof so that both the phase convention and the numerical constant are auditable.

Theorem — 4.1 (T-I.3.10: exact Fourier–energy identity)

Let $0 < s < d$ and let μ be a finite compactly supported Borel measure on $\mathbb{R}^d$. Then

$$I_s(\mu) = c_{d,s} \int_{\mathbb{R}^d} |\xi|^{s-d} |\hat{\mu}(\xi)|^2 \, d\xi, \qquad c_{d,s} = \pi^{s-d/2} \frac{\Gamma((d-s)/2)}{\Gamma(s/2)}. \tag{4.1}$$

The equality is an equality of nonnegative extended real numbers.

Proof

We begin with the one-dimensional Gaussian. Put

$$G(\xi) = \int_{\mathbb{R}} e^{-\pi x^2} e^{-2\pi i x \xi} \, dx.$$

The derivative of the integrand is dominated by the integrable function $2\pi|x| \exp(-\pi x^2)$, so differentiation under the integral sign is valid. Integration by parts, with vanishing boundary term, gives

$$\begin{aligned} G'(\xi) &= -2\pi i \int_{\mathbb{R}} x e^{-\pi x^2} e^{-2\pi i x \xi} \, dx \\ &= -2\pi \xi G(\xi). \end{aligned} \tag{4.2}$$

Moreover, Tonelli's theorem and polar coordinates give

$$\left(\int_{\mathbb{R}} e^{-\pi x^2} \, dx \right)^2 = \int_{\mathbb{R}^2} e^{-\pi |x|^2} \, dx = 2\pi \int_0^\infty r e^{-\pi r^2} \, dr = 1.$$

The one-dimensional integral is positive, so $G(0) = 1$. Solving (4.2) yields

$$G(\xi) = e^{-\pi \xi^2}. \tag{4.3}$$

Taking products in d coordinates and then scaling gives, for $t > 0$ and $z \in \mathbb{R}^d$,

$$e^{-\pi t |z|^2} = t^{-d/2} \int_{\mathbb{R}^d} e^{-\pi |\xi|^2 / t} e^{2\pi i z \cdot \xi} \, d\xi. \tag{4.4}$$

This is a direct consequence of (4.3), not an appeal to Fourier inversion.
For $z \neq 0$, the substitution $u = \pi t |z|^2$ in the defining Gamma integral gives

$$|z|^{-s} = \frac{\pi^{s/2}}{\Gamma(s/2)} \int_0^\infty t^{s/2-1} e^{-\pi t |z|^2} \, dt. \tag{4.5}$$

At $z = 0$, both sides of (4.5) are $+\infty$, so it remains valid as an extended identity. Integrating (4.5) with $z = x - y$ against $\mu \times \mu$ and using Tonelli gives

$$I_s(\mu) = \frac{\pi^{s/2}}{\Gamma(s/2)} \int_0^\infty t^{s/2-1} \iint e^{-\pi t|x-y|^2}\, d\mu(x)d\mu(y)\, dt. \tag{4.6}$$

For each fixed $t > 0$, insert (4.4). Absolute integrability follows from

$$t^{-d/2}\mu(\mathbb{R}^d)^2 \int_{\mathbb{R}^d} e^{-\pi|\xi|^2/t}\, d\xi = \mu(\mathbb{R}^d)^2 < +\infty.$$

Consequently Fubini's theorem is applicable, and the two measure integrals factor to give

$$\iint e^{-\pi t|x-y|^2}\, d\mu(x)d\mu(y) = t^{-d/2} \int_{\mathbb{R}^d} e^{-\pi|\xi|^2/t}|\hat{\mu}(\xi)|^2\, d\xi. \tag{4.7}$$

All quantities in (4.6)–(4.7) are now nonnegative. A second use of Tonelli therefore gives

$$I_s(\mu) = \frac{\pi^{s/2}}{\Gamma(s/2)} \int_{\mathbb{R}^d} |\hat{\mu}(\xi)|^2 \left(\int_0^\infty t^{(s-d)/2-1} e^{-\pi|\xi|^2/t}\, dt \right) d\xi. \tag{4.8}$$

For $\xi \neq 0$, set $u = \pi|\xi|^2/t$. Since $(d-s)/2 > 0$,

$$\begin{aligned} \int_0^\infty t^{(s-d)/2-1} e^{-\pi|\xi|^2/t}\, dt &= (\pi|\xi|^2)^{(s-d)/2}\Gamma((d-s)/2) \\ &= \pi^{(s-d)/2}\Gamma((d-s)/2)|\xi|^{s-d}. \end{aligned} \tag{4.9}$$

The exceptional point $\xi = 0$ has Lebesgue measure zero; the weight $|\xi|^{s-d}$ is locally integrable because $s > 0$. Substituting (4.9) into (4.8) produces (4.1), since

$$\frac{\pi^{s/2}}{\Gamma(s/2)} \pi^{(s-d)/2}\Gamma((d-s)/2) = c_{d,s}.$$

Every interchange was either absolutely integrable or nonnegative, so the argument also covers the value $+\infty$. □

I.3.5 Averaged energies and the first projection theorem

Lemma — 5.1 (L-I.3.11: average projected energy)

Let $0 < t < 1$ and let μ be a finite compactly supported Borel measure on $\mathbb{R}^2$. Then

$$\int_{S^1} I_t(\mu_e)\, d\sigma(e) = A_t I_t(\mu), \tag{5.1}$$

where

$$A_t = \int_{S^1} |e \cdot (1,0)|^{-t}\, d\sigma(e) < +\infty. \tag{5.2}$$

Both sides of (5.1) are allowed to be infinite.

Proof

The definition of push-forward and Tonelli's theorem give, for each e,

$$I_t(\mu_e) = \iint |(x-y)\cdot e|^{-t}\, d\mu(x)d\mu(y). \tag{5.3}$$

If $z \neq 0$, rotation invariance of normalized arclength implies

$$\int_{S^1} |z\cdot e|^{-t}\, d\sigma(e) = |z|^{-t} A_t. \tag{5.4}$$

The equality also holds at $z = 0$ in the extended sense. To check the finiteness asserted in (5.2), parametrize $e = (\cos\theta, \sin\theta)$. The only singularities of $|\cos\theta|^{-t}$ in one period occur at two simple zeros of cosine. Near each zero it is bounded by a constant times $|u|^{-t}$, whose integral is finite exactly when $t < 1$.

Finally, integrate (5.3) over e, use Tonelli to change the order of the three nonnegative integrations, and apply (5.4). This gives

$$\int_{S^1} I_t(\mu_e)\, d\sigma(e) = A_t \iint |x-y|^{-t}\, d\mu(x)d\mu(y),$$

which is (5.1). □

Theorem — 5.2 (T-I.3.12: first orthogonal projection theorem)

Let $E \subset \mathbb{R}^2$ be nonempty and compact. Then, for σ-almost every $e \in S^1$,

$$\dim_{\mathrm{H}} \pi_e(E) = \min\{\dim_{\mathrm{H}} E, 1\}. \tag{5.5}$$

Proof

Put

$$\alpha = \min\{\dim_{\mathrm{H}} E, 1\}.$$

Every π_e is 1-Lipschitz, so Hausdorff-dimension monotonicity gives $\dim_{\mathrm{H}} \pi_e(E) \le \dim_{\mathrm{H}} E$. Also $\pi_e(E) \subset R$, and hence $\dim_{\mathrm{H}} \pi_e(E) \le 1$. Thus the upper bound in (5.5) holds for every e.

If $\alpha = 0$, this upper bound and nonnegativity already give (5.5). Suppose $\alpha > 0$, and choose a strictly increasing sequence (t_j) in $(0, \alpha)$ with t_j tending to α. The energy characterization in Theorem 3.4 supplies $\mu_j \in \mathrm{Prob}(E)$ such that

$$I_{t_j}(\mu_j) < +\infty. \tag{5.6}$$

Because $t_j < \alpha \le 1$, Lemma 5.1 and (5.6) imply

$$\int_{S^1} I_{t_j}((\mu_j)_e)\, d\sigma(e) < +\infty.$$

Therefore $I_{t_j}((\mu_j)_e) < +\infty$ for almost every e. Intersecting these countably many full-measure sets of directions yields a full-measure set $\Omega \subset S^1$ on which all these energy inequalities hold. If $e \in \Omega$, then $(\mu_j)_e$ is a probability measure carried by the compact set $\pi_e(E)$. Corollary 3.3 therefore gives

$$\dim_{\mathrm{H}} \pi_e(E) \geq t_j \qquad (j \geq 1).$$

Letting j tend to infinity yields $\dim_{\mathrm{H}} \pi_e(E) \geq \alpha$. Together with the pointwise upper bound, this proves (5.5). □

I.3.6 Intuition and strategy

The following discussion explains the geometric ideas behind the proofs.

I.3.6.1 Frostman measures are distributed witnesses

A Hausdorff cover proves that a set is small by paying for the diameters of sets that cover it. A Frostman measure proves that a set is large by showing that no small ball can capture too much of a unit mass:

$$\mu(B(x,r)) \lesssim r^s.$$

The two statements are dual. If the mass is spread this thinly, any cover must have positive total s-cost. Conversely, positive Hausdorff content can be routed down a finite tree of dyadic cubes without overloading a cube. A compactness argument turns those finite routings into one limiting measure.

The construction in the proof is deliberately finite before taking a limit. This prevents the usual informal phrase “distribute mass according to content” from hiding a countable-consistency problem.

I.3.6.2 Energy measures crowding at every scale

The layer-cake formula rewrites

$$I_s(\mu) = \iint |x-y|^{-s}\, d\mu(x) d\mu(y)$$

as an integral of average ball masses over all radii. Finite energy therefore means that pairs of mass points do not crowd too strongly at small scales. A t-Frostman estimate has more decay than an s-energy requires when $t > s$; the spare exponent makes the scale integral converge.

The converse is not a pointwise statement on all of the measure. Finite average potential first gives a positive-mass level set on which the potential is bounded. Only after restricting to that set does one obtain uniform ball growth and invoke the mass-distribution principle.

I.3.6.3 Capacity packages the best energy

Capacity asks for the least possible energy among probability measures on a compact set and then takes its reciprocal. A set has positive s-capacity exactly when it can carry one finite-s-energy probability measure. The Hausdorff dimension is consequently the largest exponent below which this is possible.

The critical exponent remains delicate. Knowing

$$\dim_{\mathrm{H}} E = s$$

does not determine $\mathrm{Cap}_s(E)$. Every safe formulation in this chapter uses a strictly smaller energy exponent or a supremum over such exponents.

I.3.6.4 Why the Fourier weight has exponent $s - d$

The kernel $|x|^{-s}$ is homogeneous of degree $-s$. Fourier transformation reverses homogeneity relative to dimension, so its transform has degree $s - d$. The identity

$$I_s(\mu) = c_{d,s} \int |\xi|^{s-d} |\hat{\mu}(\xi)|^2 \, d\xi$$

turns geometric crowding into frequency concentration. It is the bridge to the Fourier-decay volumes of the series.

The constant is not decorative: it depends on the Fourier phase and its 2π normalization. The formal proof derives it from a Gaussian and a Gamma integral, so no convention is imported silently.

I.3.6.5 Projection by averaging singularity

Projecting a difference vector z in direction e replaces its length by $|z \cdot e|$. Averaging over directions gives

$$\int_{S^1} |z \cdot e|^{-t} \, d\sigma(e) = A_t |z|^{-t}.$$

The constant is finite precisely for $t < 1$: near a perpendicular direction, the dot product has a simple zero and the integrand behaves like $|u|^{-t}$. Thus a finite planar t-energy produces finite projected t-energy for almost every direction. Taking a countable sequence of exponents tending upward to $\min(\dim_{\mathrm{H}} E, 1)$ proves the dimension formula.

This argument deliberately stops at dimension. At $t = 1$ the directional average diverges, so positive Lebesgue measure of projections requires a stronger argument and is not smuggled into the baseline theorem.

I.3.7 Exercises

Every exercise uses only D/R-I.1–I.3 and A-RA. No later projection or Fourier-decay theorem is permitted.

I.3.7.1 Exercise 1 — Atoms have infinite energy

Let $0 < s < d$ and let μ be a finite Borel measure on $\mathbb{R}^d$. Prove that if $\mu(\{x_0\}) > 0$ for some x_0, then $I_s(\mu) = +\infty$. Conclude that every finite-energy measure is nonatomic.

I.3.7.2 Exercise 2 — Euclidean similarities

Fix $0 < s < d$. Let $T(x) = aOx + b$, where $a \neq 0$, O is an orthogonal transformation of $\mathbb{R}^d$, and $b \in \mathbb{R}^d$. Prove

$$I_s(T_\# \mu) = |a|^{-s} I_s(\mu)$$

for every finite Borel measure μ. Deduce

$$\operatorname{Cap}_s(T(E)) = |a|^s \operatorname{Cap}_s(E)$$

for every compact E.

I.3.7.3 Exercise 3 — Monotonicity in the energy exponent

Let μ be a probability measure carried by a bounded set of diameter $D > 0$. If $0 < s < t < d$, prove

$$I_s(\mu) \leq D^{t-s} I_t(\mu).$$

Explain why this is the local reason that the set of finite-energy exponents is an interval.

I.3.7.4 Exercise 4 — Diameter and countable sets

Fix $0 < s < d$. Let $E \subset \mathbb{R}^d$ be nonempty and compact with diameter D.

1. Prove $\operatorname{Cap}_s(E) \leq D^s$, with the convention that the right side is zero when $D = 0$.
2. Prove that if E is countable, then $\operatorname{Cap}_s(E) = 0$.

I.3.7.5 Exercise 5 — Exact energy of an interval

Let λ be Lebesgue measure restricted to $[0, 1]$. For $0 < s < 1$, prove

$$I_s(\lambda) = \frac{2}{(1-s)(2-s)}.$$

Use this to recover $\dim_{\mathrm{H}}[0, 1] = 1$ from the energy criterion.

I.3.7.6 Exercise 6 — The middle-third Cantor measure

The **middle-third Cantor set** is

$$C = \left\{ \sum_{k=1}^{\infty} 2a_k 3^{-k} \ : \ a_k \in \{0, 1\} \right\}.$$

For $a_1, \dots, a_n \in \{0, 1\}$, its level-$n$ **basic interval** is

$$\left[\sum_{k=1}^{n} 2a_k 3^{-k},\ \sum_{k=1}^{n} 2a_k 3^{-k} + 3^{-n}\right].$$

Let

$$q = \frac{\log 2}{\log 3}.$$

To define the **Cantor measure** without an existence assumption, put

$$b_k(u) = \lfloor 2^k u \rfloor - 2\lfloor 2^{k-1} u \rfloor, \qquad T(u) = \sum_{k=1}^{\infty} 2b_k(u)3^{-k} \quad (0 \le u < 1), \qquad \mu = T_\#(\mathscr{L}^1|_{[0,1)}).$$

Verify that this probability measure assigns mass 2^{-n} to every level-n basic interval. Prove that

$$\mu(\overline{B}(x,r)) \le 8r^q \qquad (x \in \mathbb{R}, r > 0).$$

Deduce directly that $\dim_{\mathrm{H}} C = q$.

I.3.7.7 Exercise 7 — Fourier audit for the interval

For the measure λ in Exercise 5, prove

$$\hat{\lambda}(\xi) = e^{-\pi i \xi} \frac{\sin(\pi\xi)}{\pi\xi},$$

where the quotient is interpreted as one at $\xi = 0$. Check directly that

$$\int_{\mathbb{R}} |\xi|^{s-1} |\hat{\lambda}(\xi)|^2 \, d\xi < +\infty$$

for $0 < s < 1$, and use Theorem 4.1 and Exercise 5 to identify the integral exactly in terms of $c_{1,s}$.

I.3.7.8 Exercise 8 — Compactly supported L^2 densities

Let $f \ge 0$ be compactly supported, $f \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, and $\int f = 1$. Let $d\mu = f dx$. Using Plancherel's theorem from A-RA and Theorem 4.1, prove

$$I_s(\mu) < +\infty \qquad (0 < s < d).$$

I.3.7.9 Exercise 9 — A quantitative energy-to-measure bound

Let $E \subset \mathbb{R}^d$ be Borel, $0 < s < d$, and let μ be a probability measure carried by E. Suppose $0 < I_s(\mu) < +\infty$. Prove

$$\mathscr{H}^s(E) \ge \frac{1}{4I_s(\mu)}.$$

Your proof should make the level set of the potential and every constant explicit.

I.3.7.10 Exercise 10 — A quantitative exceptional-direction estimate

Let $0 < t < 1$, let μ be a finite compactly supported Borel measure on $\mathbb{R}^2$, and let $L > 0$. Prove

$$\sigma\{e \in S^1 : I_t(\mu_e) > L\} \leq \frac{A_t I_t(\mu)}{L}.$$

What does the estimate say when $I_t(\mu) < +\infty$ and L tends to infinity?

I.3.7.11 Exercise 11 — Oriented and unoriented directions

Let $0 < t < 1$, let μ be a finite compactly supported Borel measure on $\mathbb{R}^2$, and let $E \subset \mathbb{R}^2$ be nonempty and compact. An **unoriented line direction** is a one-dimensional linear subspace of $\mathbb{R}^2$. Write

$$\mathcal{G} = \{\mathbb{R}e : e \in S^1\}, \qquad q : S^1 \to \mathcal{G}, \quad q(e) = \mathbb{R}e.$$

Equip $\mathcal{G}$ with the quotient topology under the identification e with $-e$ and the direction probability measure $q_\#\sigma$ on its Borel sets.

Prove

$$I_t(\mu_{-e}) = I_t(\mu_e), \qquad \dim_{\mathrm{H}} \pi_{-e}(E) = \dim_{\mathrm{H}} \pi_e(E).$$

Explain why stating Theorem 5.2 on S^1 or on the space of unoriented lines produces the same exceptional-direction conclusion.

I.3.7.12 Exercise 12 — Audit the endpoint $t = 1$

Prove

$$A_1 = \int_{S^1} |e \cdot (1,0)|^{-1} \, d\sigma(e) = +\infty.$$

Identify the exact step in Lemma 5.1 and Theorem 5.2 that cannot be run with $t = 1$. Explain why proving $\dim_{\mathrm{H}} \pi_e(E) = 1$ does not by itself prove that the projection has positive one-dimensional Lebesgue measure.

Chapter I.4

Dimensions of Measures and Exact Dimensionality

The size of a measure near a typical point need not describe every part of its support. Local dimensions quantify this pointwise behavior, while dimensions of a measure ask how small a set can carry positive or full mass. We relate these notions and introduce exact dimensionality. Entropy offers a third, averaged description at a fixed scale. The examples are designed to show which implications hold and which plausible equalities fail.

The definitions below fix this chapter's conventions. Geometric intuition and proof strategy are discussed separately after the main development; exercises and their solutions provide a second route through the material.

Definitions and notation

Throughout, μ denotes a Borel probability measure on $\mathbb{R}^d$. Closed balls are used in local-dimension limits. Changing to open balls does not alter the limits, but that equivalence is not used as an unstated convention.

Definition — I.4.01: Lower and upper local dimensions

For $x \in \mathbb{R}^d$, define

$$\underline{d}_\mu(x) = \liminf_{r\downarrow 0} \frac{\log \mu(\overline{B}(x,r))}{\log r}, \qquad \overline{d}_\mu(x) = \limsup_{r\downarrow 0} \frac{\log \mu(\overline{B}(x,r))}{\log r}.$$

Only radii $0 < r < 1$ matter, so the denominator is negative. We set $\log 0 = -\infty$; its quotient by $\log r < 0$ is $+\infty$. If the two values agree, their common value is the **local dimension** $d_\mu(x)$.

Definition — I.4.02: Essential extrema

For a Borel function $f : \mathbb{R}^d \to [0, +\infty]$, define

$$\operatorname*{ess\,inf}_{\mu} f = \sup\{a \in \mathbb{R} : f(x) \geq a \text{ for } \mu\text{-a.e. } x\},$$

and

$$\operatorname*{ess\,sup}_{\mu} f = \inf\{a \in \mathbb{R} : f(x) \leq a \text{ for } \mu\text{-a.e. } x\},$$

with the usual extended-real endpoint conventions.

Definition — I.4.03: Four set-carried dimensions of a measure

Define the lower and upper Hausdorff dimensions of μ by

$$\underline{\dim}_{\mathrm{H}}\mu = \inf\{\dim_{\mathrm{H}} E : E \text{ Borel}, \ \mu(E) > 0\},$$

$$\overline{\dim}_{\mathrm{H}}\mu = \inf\{\dim_{\mathrm{H}} E : E \text{ Borel}, \ \mu(E) = 1\}.$$

Define the lower and upper packing dimensions by

$$\underline{\dim}_{\mathrm{P}}\mu = \inf\{\dim_{\mathrm{P}} E : E \text{ Borel}, \ \mu(E) > 0\},$$

$$\overline{\dim}_{\mathrm{P}}\mu = \inf\{\dim_{\mathrm{P}} E : E \text{ Borel}, \ \mu(E) = 1\}.$$

"Lower" refers to positive-mass carriers and "upper" to full-mass carriers; it does not refer to lower or upper box dimension.

Definition — I.4.04: Exact dimensionality

The probability measure μ is **exact dimensional of dimension** α if

$$\underline{d}_{\mu}(x) = \overline{d}_{\mu}(x) = \alpha \qquad \text{for } \mu\text{-a.e. } x.$$

The value α is then denoted $\dim \mu$. The phrase is not used when the almost-everywhere local limit fails to exist or is not almost surely constant.

Definition — I.4.05: Adic partitions and Shannon entropy

For an integer $b \geq 2$ and $n \geq 0$, let

$$\mathscr{D}_n^{(b)} = \left\{ \prod_{j=1}^{d} [k_j b^{-n}, (k_j + 1) b^{-n}) : k \in \mathbb{Z}^d \right\}.$$

These half-open cubes form the level-n b-**adic partition**. For a finite or countable measurable partition P whose entropy is well defined, set

$$H(\mu, \mathscr{P}) = -\sum_{P \in \mathscr{P}} \mu(P) \log \mu(P),$$

with $0 \log 0 = 0$. Natural logarithms are used. If μ is compactly supported, only finitely many atoms of D_n^b have positive mass.

Definition — I.4.06: Entropy dimensions

For a compactly supported probability measure μ, define

$$\underline{\dim}_{\mathrm{e}}^{(b)} \mu = \liminf_{n\to\infty} \frac{H(\mu, \mathscr{D}_n^{(b)})}{n \log b},$$

$$\overline{\dim}_{\mathrm{e}}^{(b)} \mu = \limsup_{n\to\infty} \frac{H(\mu, \mathscr{D}_n^{(b)})}{n \log b}.$$

L-I.4.10 proves that these values do not depend on the integer base. After that proof the superscript is dropped. If the lower and upper values agree, their common value is the **entropy dimension** $\dim_{\mathrm{e}} \mu$.

Definition — I.4.07: Cantor coding and inhomogeneous Bernoulli products

Let $\Sigma = \{0, 1\}^{\mathbb{N}}$, where $\mathbb{N} = \{1, 2, \dots\}$. If $\omega \neq \tau$, let

$$N(\omega, \tau) = \min\{k \geq 1 \,:\, \omega_k \neq \tau_k\}, \qquad d_3(\omega, \tau) = 3^{1-N(\omega,\tau)},$$

and set $d_3(\omega, \omega) = 0$. The length-n cylinder at ω is

$$[\omega|n] = \{\tau \,:\, \tau_k = \omega_k \text{ for } 1 \leq k \leq n\}.$$

The Cantor coding map is

$$\pi(\omega) = \sum_{k=1}^{\infty} 2\omega_k 3^{-k}.$$

Given numbers $p_k \in (0, 1)$, the **inhomogeneous Bernoulli product** ν_p is the probability measure on Σ determined by

$$\nu_p([\omega|n]) = \prod_{k=1}^{n} p_k^{1-\omega_k}(1 - p_k)^{\omega_k}.$$

Its Cantor realization is $\mu_p = \pi_\# \nu_p$. The binary entropy function is

$$h(p) = -p \log p - (1 - p) \log(1 - p).$$

All measures are Borel probability measures on Euclidean space unless another compact metric space is explicitly named. Hausdorff and packing dimensions retain the conventions of I.1–I.2.

I.4.1 Local dimensions and the Hausdorff dictionary

For $x \in \mathbb{R}^d$, put

$$\underline{d}_\mu(x) = \liminf_{r\downarrow 0} \frac{\log \mu(\overline{B}(x,r))}{\log r}, \qquad \overline{d}_\mu(x) = \limsup_{r\downarrow 0} \frac{\log \mu(\overline{B}(x,r))}{\log r}, \tag{1.1}$$

with the conventions in D-I.4.01.

Lemma — 1.1 (L-I.4.01: geometric sampling and measurability)

Fix $0 < c < 1$. Both limits in (1.1) are unchanged if r is restricted to the sequence c^n. Both local-dimension functions are Borel measurable.

Proof

It is convenient to write

$$a_x(r) = \frac{-\log \mu(\overline{B}(x,r))}{-\log r} \in [0,+\infty].$$

If $c^{n+1} < r \le c^n$, monotonicity of balls gives

$$-\log \mu(\overline{B}(x,c^n)) \le -\log \mu(\overline{B}(x,r)) \le -\log \mu(\overline{B}(x,c^{n+1})), \tag{1.2}$$

while

$$n|\log c| \le -\log r < (n+1)|\log c|. \tag{1.3}$$

Dividing the extreme terms in (1.2) by the extreme denominators in (1.3) shows, with factors $n/(n+1)$ that tend to one, that the full liminf and limsup equal their values on c^n.

For fixed $r > 0$, the map

$$x \longmapsto \mu(\overline{B}(x,r))$$

is upper semicontinuous. Indeed, if x_j tends to x, then for every $\eta > 0$ and all large j,

$$\overline{B}(x_j,r) \subset \overline{B}(x,r+\eta).$$

Hence

$$\limsup_j \mu(\overline{B}(x_j,r)) \le \mu(\overline{B}(x,r+\eta)).$$

Letting η decrease to zero and using continuity from above of the finite measure proves the claim. Thus $a_x(c^n)$ is an extended Borel function of x. Countable liminf and limsup preserve Borel measurability, and the first part of the proof identifies them with (1.1). ☐

We next isolate the covering selection used in the converse local estimate.

Lemma — 1.2 (L-I.4.02: Euclidean $5r$ selection)

Let V be a family of positive-radius closed balls in $\mathbb{R}^d$, and suppose their radii are bounded above. There is a finite or countable pairwise disjoint subfamily $W \subset V$ such that

$$\bigcup_{B\in\mathscr{V}} B \subset \bigcup_{B\in\mathscr{W}} 5B, \tag{1.4}$$

where $5B$ is the concentric ball with five times the radius.

Proof

Choose $R > 0$ larger than every radius and divide V into classes

$$\mathscr{V}_k = \{B \in \mathscr{V} \ : \ 2^{-k-1}R < r(B) \le 2^{-k}R\}, \qquad k \ge 0.$$

Proceed through the classes in increasing k. In class k, after removing balls meeting a ball already selected in an earlier class, choose a maximal pairwise disjoint subfamily of the remaining balls. To justify maximality, order the pairwise disjoint subfamilies by inclusion. The union of every chain is again pairwise disjoint and is an upper bound for that chain, so Zorn's lemma supplies a maximal member. The union W of all selected families is pairwise disjoint. It is countable because the interiors of its members are disjoint nonempty open sets and each contains a distinct point of the countable dense set Q^d.

Let $B \in V_k$. If it was removed by an earlier selected ball B', then $r(B) \le r(B')$; since the balls meet, $B \subset 3B'$. Otherwise maximality in class k gives a selected $B' \in V_k$ meeting B. Now $r(B) < 2r(B')$, so every $y \in B$ satisfies

$$|y - c(B')| \le 2r(B) + r(B') < 5r(B').$$

Thus $B \subset 5B'$ in either case, proving (1.4). □

Lemma — 1.3 (L-I.4.03: local Hausdorff comparison sets)

Let $s > 0$.

(1) If E is Borel, $\mu(E) > 0$, and $\underline{d}_\mu(x) > s$ for μ-almost every $x \in E$, then $\dim_{\mathrm{H}} E \ge s$.

(2) The Borel set

$$A_s = \{x \ : \ \underline{d}_\mu(x) < s\}$$

satisfies $\dim_{\mathrm{H}} A_s \le s$.

Proof

Discarding a null subset in the first assertion, assume the strict local inequality at every point of E. For $m \ge 1$, let

$$E_m = \{x \in E \,:\, \mu(\overline{B}(x,r)) \le r^s \text{ whenever } 0 < r < 1/m\}. \tag{1.5}$$

The condition may be tested first at rational radii and then passed to every radius by continuity from above, so E_m is Borel. The strict liminf inequality implies $E = \cup_m E_m$. Some E_m has positive measure. For $\nu = \mu|E_m$, (1.5) gives the s-growth bound at centres in E_m for small radii; for $r \ge 1/m$,

$$\nu(\overline{B}(x,r)) \le 1 \le m^s r^s.$$

The mass-distribution principle applied to ν gives $\mathscr{H}^s(E_m) > 0$, and hence $\dim_{\mathrm{H}} E \ge s$.

For the second assertion, fix an integer $j \ge 1$ and a scale $\delta > 0$. At every $x \in A_s \cap B(0,j)$ there are arbitrarily small radii r for which

$$\frac{\log \mu(\overline{B}(x,r))}{\log r} < s,$$

or, because $\log r < 0$,

$$\mu(\overline{B}(x,r)) > r^s. \tag{1.6}$$

Take all such balls with $r < \delta/10$. Lemma 1.2 selects disjoint balls $B_i = \overline{B}(x_i, r_i)$ whose fivefold dilates cover $A_s \cap B(0,j)$. By (1.6) and disjointness,

$$\sum_i (\operatorname{diam} 5B_i)^s = 10^s \sum_i r_i^s < 10^s \sum_i \mu(B_i) \le 10^s. \tag{1.7}$$

Thus $\mathscr{H}^s_\delta(A_s \cap B(0,j)) \le 10^s$ for every δ, so its s-dimensional Hausdorff measure is finite and its dimension is at most s. Countable stability over j proves $\dim_{\mathrm{H}} A_s \le s$. ☐

Theorem — 1.4 (T-I.4.04: Hausdorff local-to-global dictionary)

For every Borel probability measure μ on $\mathbb{R}^d$,

$$\underline{\dim}_{\mathrm{H}}\mu = \operatorname*{ess\,inf}_{\mu} \underline{d}_\mu, \qquad \overline{\dim}_{\mathrm{H}}\mu = \operatorname*{ess\,sup}_{\mu} \underline{d}_\mu. \tag{1.8}$$

Proof

Put $a = \text{ess inf}_\mu\, \underline{d}_\mu$. If $s < a$, the set $\{\underline{d}_\mu > s\}$ has full measure. Every positive-mass Borel E therefore meets it in positive measure, and Lemma 1.3 applied to that intersection gives $\dim_{\mathrm{H}} E \ge s$. Hence $\underline{\dim}_{\mathrm{H}}\mu \ge s$.

If $s > a$, the set $\{\underline{d}_\mu < s\}$ has positive measure by the definition of essential infimum. Lemma 1.3 gives this set Hausdorff dimension at most s, so $\underline{\dim}_{\mathrm{H}}\mu \le s$. Letting s approach a from both sides proves the first equality, including the endpoint cases.

Put $b = \text{ess sup}_\mu\, \underline{d}_\mu$. If $s > b$, the set $\{\underline{d}_\mu < s\}$ has full measure and dimension at most s, so $\overline{\dim}_{\mathrm{H}}\mu \le s$. If $s < b$, then $\{\underline{d}_\mu > s\}$ has positive measure. Every full-measure Borel set intersects it in positive measure and Lemma 1.3 forces that full-measure set to have dimension at least s. Thus $\overline{\dim}_{\mathrm{H}}\mu \ge s$. Letting s tend to b proves the second equality. ☐

I.4.2 Upper local dimension and the packing dictionary

Lemma — 2.1 (L-I.4.05: reverse ball bounds and upper box dimension)

Suppose $F \subset \mathbb{R}^d$ is bounded and, for some $q \geq 0$ and $r_0 > 0$,

$$\mu(\overline{B}(x,r)) \geq r^q \qquad (x \in F,\ 0 < r < r_0). \tag{2.1}$$

Then $\overline{\dim}_{\mathrm{B}} F \leq q$. Consequently, for every $s > 0$,

$$\dim_{\mathrm{P}}\{x : \overline{d}_\mu(x) < s\} \leq s. \tag{2.2}$$

Proof

For $0 < r < 2r_0$, take a maximal r-separated subset $\{x_1, ..., x_N\}$ of F. The balls $\overline{B}(x_i, r/2)$ are pairwise disjoint, while maximality makes the radius-r balls cover F. By (2.1),

$$1 \geq \sum_{i=1}^{N} \mu(\overline{B}(x_i, r/2)) \geq N(r/2)^q.$$

Thus

$$N_r(F) \leq N \leq 2^q r^{-q},$$

which proves the upper-box bound.

Let $A = \{x : \overline{d}_\mu(x) < s\}$. For each $x \in A$, choose a rational $q < s$ strictly between $\overline{d}_\mu(x)$ and s. For all sufficiently small r,

$$\frac{\log \mu(\overline{B}(x,r))}{\log r} < q,$$

so $\mu(\overline{B}(x,r)) > r^q$. Decompose A according to the rational q, a uniform small-scale threshold, and intersection with $B(0,j)$. Each piece has upper box dimension at most its corresponding q by the first part. The modified-upper-box characterization and countable stability of packing dimension give (2.2). □

Lemma — 2.2 (L-I.4.06: low upper box is invisible at high local dimension)

If $F \subset \mathbb{R}^d$ is bounded and $\overline{\dim}_{\mathrm{B}} F < s$, then

$$\mu^*(F \cap \{x : \overline{d}_\mu(x) > s\}) = 0. \tag{2.3}$$

Here $\mu^*(A) = \inf\{\mu(B) : A \subset B,\ B \text{ Borel}\}$ is the outer measure induced by μ. This formulation does not require F to be measurable; for Borel F, it is the ordinary measure assertion.

Consequently, if a Borel set E satisfies

$$\mu(E \cap \{x : \overline{d}_\mu(x) > s\}) > 0, \tag{2.4}$$

then $\dim_{\mathrm{P}} E \geq s$.

Proof

Replace F by its closure. Closure preserves upper box dimension by T-I.2.03, and an outer-measure-zero conclusion for the closure implies the same conclusion for F. We may therefore assume that F is compact. In particular, the sets A_n below are Borel by Lemma 1.1's ball-mass measurability argument.

Choose u with $\overline{\dim}_{\mathrm{B}} F < u < s$ and put $r_n = 2^{-n}$. For all large n,

$$N_{r_n}(F) \leq r_n^{-u}. \tag{2.5}$$

Define

$$A_n = \{x \in F \,:\, \mu(\overline{B}(x, 4r_n)) \leq (4r_n)^s\}. \tag{2.6}$$

Cover F by the number of radius-r_n balls in (2.5), with centres in F. For each covering ball that meets A_n, choose a point x_i in the intersection. The covering ball is contained in $\overline{B}(x_i, 2r_n)$ and hence in $\overline{B}(x_i, 4r_n)$. Therefore

$$\mu(A_n) \leq N_{r_n}(F)(4r_n)^s \leq 4^s r_n^{s-u}. \tag{2.7}$$

The right side is summable in n. Consequently

$$\mu(\limsup_n A_n) \leq \lim_{N\to\infty} \sum_{n\geq N} \mu(A_n) = 0. \tag{2.8}$$

If $x \in F$ and $\overline{d}_\mu(x) > s$, choose t strictly between these two numbers. Lemma 1.1, applied to the geometric radii $4r_n$, shows that for infinitely many large n,

$$\mu(\overline{B}(x, 4r_n)) < (4r_n)^t \leq (4r_n)^s.$$

Thus $x \in \limsup A_n$, and (2.3) follows from (2.8).

Suppose (2.4) holds but $\dim_{\mathrm{P}} E < s$. Some bounded part of E still has positive measure in the set from (2.4). The identity between packing and modified upper box dimension supplies a countable cover of that bounded part by sets F_j with $\overline{\dim}_{\mathrm{B}} F_j < s$. Equation (2.3) makes every F_j-intersection with $\{\overline{d}_\mu > s\}$ have outer measure zero. Countable subadditivity of outer measure contradicts (2.4). Hence $\dim_{\mathrm{P}} E \geq s$. □

Theorem — 2.3 (T-I.4.07: packing local-to-global dictionary)

For every Borel probability measure μ on $\mathbb{R}^d$,

$$\underline{\dim}_{\mathrm{P}} \mu = \operatorname*{ess\,inf}_\mu \overline{d}_\mu, \qquad \overline{\dim}_{\mathrm{P}} \mu = \operatorname*{ess\,sup}_\mu \overline{d}_\mu. \tag{2.9}$$

Proof

Let $a = \operatorname{ess\,inf}_\mu \overline{d}_\mu$. If $s < a$, then $\{\overline{d}_\mu > s\}$ has full measure. Every positive-mass Borel set satisfies (2.4), so Lemma 2.2 gives packing dimension at least s. Hence $\underline{\dim}_{\mathrm{P}}\mu \geq s$. If $s > a$, the set $\{\overline{d}_\mu < s\}$ has positive measure and packing dimension at most s by Lemma 2.1. Hence $\underline{\dim}_{\mathrm{P}}\mu \leq s$. Letting s tend to a proves the first equality.

Let $b = \operatorname{ess\,sup}_\mu \overline{d}_\mu$. If $s > b$, then $\{\overline{d}_\mu < s\}$ is a full-measure set of packing dimension at most s, so $\overline{\dim}_{\mathrm{P}}\mu \leq s$. If $s < b$, then $\{\overline{d}_\mu > s\}$ has positive measure. Every full-measure Borel set satisfies (2.4), so its packing dimension is at least s. Thus $\overline{\dim}_{\mathrm{P}}\mu \geq s$. Taking limits proves the second equality. □

I.4.3 The dimension lattice and maps

Corollary — 3.1 (C-I.4.08: dimension lattice and exactness)

Every probability measure satisfies

$$\underline{\dim}_{\mathrm{H}}\mu \leq \min\{\overline{\dim}_{\mathrm{H}}\mu, \underline{\dim}_{\mathrm{P}}\mu\} \leq \max\{\overline{\dim}_{\mathrm{H}}\mu, \underline{\dim}_{\mathrm{P}}\mu\} \leq \overline{\dim}_{\mathrm{P}}\mu. \tag{3.1}$$

It is exact dimensional of dimension α if and only if all four measure dimensions in (3.1) equal α.

Proof

The essential infimum of a function does not exceed its essential supremum, and $\underline{d}_\mu \leq \overline{d}_\mu$ pointwise. Applying these facts to the four formulas (1.8) and (2.9) gives every inequality in (3.1).

If μ is exact dimensional of dimension α, both local functions equal α almost everywhere, so all four essential extrema equal α. Conversely, equality of the lower and upper Hausdorff dimensions at α gives

$$\operatorname*{ess\,inf}_{\mu} \underline{d}_\mu = \operatorname*{ess\,sup}_{\mu} \underline{d}_\mu = \alpha,$$

and hence $\underline{d}_\mu = \alpha$ almost everywhere. The two packing-dimension identities similarly give $\overline{d}_\mu = \alpha$ almost everywhere. Thus the local limit exists and equals α almost everywhere. □

Theorem — 3.2 (T-I.4.09: Lipschitz and bi-Lipschitz covariance)

Let $f : \mathbb{R}^d \to \mathbb{R}^m$ be L-Lipschitz and put $\nu = f_\# \mu$. For every x,

$$\underline{d}_\nu(f(x)) \leq \underline{d}_\mu(x), \qquad \overline{d}_\nu(f(x)) \leq \overline{d}_\mu(x). \tag{3.2}$$

Each of the four global measure dimensions of ν is at most the corresponding dimension of μ. If f is a bi-Lipschitz embedding, all pointwise and global inequalities are equalities, and exact dimensionality with its dimension is preserved.

Proof

If $L = 0$, then ν is a point mass and the assertions follow immediately. Assume $L > 0$. The Lipschitz inequality gives

$$\mu(\overline{B}(x,r)) \le \nu(\overline{B}(f(x), Lr)). \tag{3.3}$$

Take negative logarithms in (3.3), divide by $-\log(Lr)$, and note that

$$\frac{-\log r}{-\log(Lr)} \longrightarrow 1.$$

The liminf and limsup give (3.2).

For a global dimension defined using a positive-mass carrier, choose a positive-mass Borel carrier and then, by inner regularity, a positive-mass compact subset K. The compact set $f(K)$ has positive ν-measure and

$$\dim_{\mathrm{H}} f(K) \le \dim_{\mathrm{H}} K, \qquad \dim_{\mathrm{P}} f(K) \le \dim_{\mathrm{P}} K.$$

Taking infima proves both lower-dimension inequalities. For a full-mass carrier E, choose compact $K_j \subset E$ with $\mu(K_j) > 1 - 1/j$. The Borel set $\cup_j f(K_j)$ has full ν-measure. Countable stability and Lipschitz monotonicity give its Hausdorff or packing dimension at most that of E, proving both upper-dimension inequalities.

If f is bi-Lipschitz onto its image, apply the same arguments to its Lipschitz inverse. This reverses (3.2) and all four global inequalities. The exact-dimensional assertion follows from equality of the two local dimensions almost everywhere. □

I.4.4 Entropy dimension

For a compactly supported μ and integer $b \ge 2$, abbreviate

$$H_b(n) = H(\mu, \mathscr{D}_n^{(b)}).$$

Lemma — 4.1 (L-I.4.10: entropy calculus and base independence)

The following statements hold.

(1) If a finite partition Q refines a finite partition P, then $H(\mu, P) \le H(\mu, Q)$.

(2) If every atom of P meets at most M atoms of Q on the support of μ, then

$$H(\mu, \mathscr{Q}) \le H(\mu, \mathscr{P}) + \log M. \tag{4.1}$$

(3) If $\mu = p\nu + (1-p)\eta$, where $0 \le p \le 1$, then

$$pH(\nu, \mathscr{P}) + (1-p)H(\eta, \mathscr{P}) \le H(\mu, \mathscr{P}) \tag{4.2}$$

and

$$H(\mu, \mathscr{P}) \le h(p) + pH(\nu, \mathscr{P}) + (1-p)H(\eta, \mathscr{P}). \tag{4.3}$$

(4) The lower and upper entropy dimensions in D-I.4.06 are independent of the integer base.

Proof

Let $\mathscr{P} \vee \mathscr{Q} = \{P \cap Q : P \in \mathscr{P}, Q \in \mathscr{Q},\ P \cap Q \neq \emptyset\}$ be the common refinement. For an atom P with positive mass, define its conditional probability vector by $q_{Q|P} = \mu(P \cap Q)/\mu(P)$, indexed by $Q \in \mathscr{Q}$. Its entropy means $H(q_{Q|P}) = -\sum_{Q\in\mathscr{Q}} q_{Q|P} \log q_{Q|P}$. Atoms P of mass zero contribute zero to the sum below and need no conditional vector. Direct expansion of $-u \log u$ gives the chain identity

$$H(\mu, \mathscr{P} \vee \mathscr{Q}) = H(\mu, \mathscr{P}) + \sum_{P\in\mathscr{P}} \mu(P)H(q_{Q|P}). \tag{4.4}$$

Every entropy on the right is nonnegative, proving refinement monotonicity. If at most M conditional weights are nonzero, their entropy is at most $\log M$. Indeed, if the positive weights are $q_1, ..., q_k$, weighted Jensen for the concave logarithm gives

$$\sum_{i=1}^{k} q_i \log(1/q_i) \le \log\left(\sum_{i=1}^{k} q_i(1/q_i)\right) = \log k \le \log M.$$

Equation (4.4), followed by refinement monotonicity from Q to $\mathscr{P} \vee \mathscr{Q}$, proves (4.1).

For (4.2), apply concavity of $-u \log u$ to every atom mass and sum. For the upper bound, split each atom of mass $p\nu(P) + (1-p)\eta(P)$ into the two masses $p\nu(P)$ and $(1-p)\eta(P)$. Splitting can only increase entropy, and the entropy of the split vector is

$$h(p) + pH(\nu, \mathscr{P}) + (1-p)H(\eta, \mathscr{P}),$$

which proves (4.3).

Finally fix integer bases $b, c \ge 2$. For each n, choose $m = m(n)$ so that

$$m \log c = n \log b + O(1). \tag{4.5}$$

The side lengths b^{-n} and c^{-m} then differ by a factor bounded only by b, c. A cube from either grid meets at most $M_{b,c,d}$ cubes of the other grid. Apply (4.1) in both directions to their common refinement to obtain

$$|H_b(n) - H_c(m(n))| \le C_{b,c,d}. \tag{4.6}$$

Divide by $n \log b$; equation (4.5) makes the alternative denominator differ by $O(1)$. The sequence $m(n)$ misses at most a bounded number of consecutive integers, and entropy is nondecreasing under refinement with increment at most $d \log c$ per level. Therefore its sampled liminf and limsup equal the full ones. Equations (4.5)–(4.6) prove base independence. □

Theorem — 4.2 (T-I.4.11: Hausdorff–entropy–packing sandwich)

If μ is compactly supported, then

$$\underline{\dim}_{\mathrm{H}}\mu \leq \underline{\dim}_{\mathrm{e}}\mu \leq \overline{\dim}_{\mathrm{e}}\mu \leq \overline{\dim}_{\mathrm{P}}\mu. \tag{4.7}$$

Proof

Fix a base b. Write

$$e = \liminf_{n} \frac{H_b(n)}{n \log b}.$$

Choose u, q, t with $e < u < q < t$. Along an infinite sequence (n_k),

$$H_b(n_k) \leq u n_k \log b. \tag{4.8}$$

Let A_k be the union of those level-n_k cubes Q for which

$$\mu(Q) \geq b^{-q n_k}. \tag{4.9}$$

For a random point with law μ, the quantity $-\log \mu(Q)$ at its level cube has expectation $H_b(n_k)$. Outside A_k it exceeds $q n_k \log b$. Hence (4.8) gives

$$\mu(A_k) \geq 1 - u/q > 0. \tag{4.10}$$

There are at most $b^{q n_k}$ cubes satisfying (4.9). Put $A = \limsup_k A_k$. Since

$$\mu(A) = \lim_{K\to\infty} \mu\left(\bigcup_{k\geq K} A_k\right) \geq 1 - u/q, \tag{4.11}$$

the set A has positive measure. For every K, the cubes defining A_k, $k \geq K$, cover A, and their total t-cost is at most

$$d^{t/2} \sum_{k\geq K} b^{-(t-q)n_k}. \tag{4.12}$$

The tail tends to zero because $n_k \geq k$ after reindexing. Thus $\mathscr{H}^t(A) = 0$, so $\underline{\dim}_{\mathrm{H}}\mu \leq t$. Letting t, q, u decrease to e proves the first inequality in (4.7). The middle inequality is immediate.

Let $p = \overline{\dim}_{\mathrm{P}}\mu$ and fix $s > p$. There is a full-measure Borel set E with $\dim_{\mathrm{P}} E < s$. By the modified-upper-box formula, cover E by bounded sets F_j whose upper box dimensions are all strictly below s. Intersect with the compact support and take closures; this preserves upper box dimension and makes the pieces compact. For any $\delta > 0$, a finite union $F = F_1 \cup \ldots \cup F_J$ may be chosen with

$$\mu(F) > 1 - \delta. \tag{4.13}$$

This finite union has upper box dimension $u < s$ for some u. Therefore the number M_n of b-adic level-n cubes meeting F satisfies

$$\log M_n \leq un \log b + o(n). \tag{4.14}$$

The compact support meets at most Cb^{dn} such cubes. Decompose μ as the mixture of its normalized restrictions to F and F^c. The entropy mixture bound (4.3), together with the maximum entropy on a finite alphabet, gives

$$H_b(n) \leq \log 2 + (1-\delta)\log M_n + \delta(dn \log b + O(1)). \tag{4.15}$$

If the complement in (4.13) has mass smaller than δ, enlarge its coefficient to δ; the same upper bound remains valid. Divide (4.15) by $n \log b$ and take limsup:

$$\overline{\dim}_{\mathrm{e}}\mu \leq (1-\delta)u + \delta d \leq s + \delta d.$$

First let δ decrease to zero and then let s decrease to p. This proves the last inequality in (4.7). ☐

Corollary — 4.3 (C-I.4.12: exact measures have entropy dimension)

If a compactly supported probability measure is exact dimensional of dimension α, then

$$\dim_{\mathrm{e}} \mu = \alpha. \tag{4.16}$$

Proof

Corollary 3.1 makes all four set-carried dimensions equal to α. The two middle quantities in (4.7) are squeezed between α and α; hence both equal α. ☐

I.4.5 Mixtures and inhomogeneous Bernoulli products

Theorem — 5.1 (T-I.4.13: separated mixtures)

Let μ_0, μ_1 be compactly supported exact-dimensional probabilities of dimensions α_0, α_1, and suppose their supports have positive distance. For $0 < p < 1$, put

$$\mu = p\mu_0 + (1-p)\mu_1.$$

Then

$$\underline{\dim}_{\mathrm{H}}\mu = \underline{\dim}_{\mathrm{P}}\mu = \min\{\alpha_0, \alpha_1\}, \tag{5.1}$$

$$\overline{\dim}_{\mathrm{H}}\mu = \overline{\dim}_{\mathrm{P}}\mu = \max\{\alpha_0, \alpha_1\}, \tag{5.2}$$

and

$$\dim_{\mathrm{e}} \mu = p\alpha_0 + (1-p)\alpha_1. \tag{5.3}$$

Proof

At a point in spt μ_0, every sufficiently small ball misses spt μ_1, and hence

$$\mu(\overline{B}(x,r)) = p\mu_0(\overline{B}(x,r)).$$

The constant factor p disappears after division by $\log r$. Thus both local dimensions of μ equal α_0 at μ_0-almost every point. The parallel statement with α_1 holds on the other support. The four dictionary formulas immediately give (5.1)–(5.2).

Corollary 4.3 gives $\dim_e \mu_i = \alpha_i$. For every adic partition, the mixture inequalities (4.2)–(4.3) place $H(\mu, D_n)$ between

$$pH(\mu_0, \mathscr{D}_n) + (1-p)H(\mu_1, \mathscr{D}_n)$$

and that quantity plus $h(p)$. Divide by the scale denominator and take the limit to obtain (5.3). ☐

Lemma — 5.2 (L-I.4.14: bounded independent strong law)

Let (X_k) be independent real random variables satisfying $EX_k = 0$ and $|X_k| \le M$ for every k. Then

$$\frac{1}{n}\sum_{k=1}^{n} X_k \longrightarrow 0 \quad \text{almost surely.} \tag{5.4}$$

Proof

Put $S_n = \sum_{k\le n} X_k$. Independence and centering give

$$\mathrm{Var}(S_n) = \sum_{k=1}^{n} \mathrm{Var}(X_k) \le M^2 n. \tag{5.5}$$

Chebyshev's inequality therefore yields

$$\mathbb{P}(|S_{m^2}| > \varepsilon m^2) \le \frac{M^2}{\varepsilon^2 m^2}. \tag{5.6}$$

The sum of (5.6) over m is finite. The elementary summable-tail argument used in (2.8) shows that only finitely many of these events occur. Applying this for rational positive ε gives

$$S_{m^2}/m^2 \longrightarrow 0 \quad \text{almost surely.} \tag{5.7}$$

If $m^2 \le n < (m+1)^2$, then

$$|S_n - S_{m^2}| \le M(2m+1).$$

After division by n, this error tends to zero, and (5.7) proves (5.4). ☐

Theorem — 5.3 (T-I.4.15: inhomogeneous Bernoulli formula)

Let $\varepsilon \in (0, 1/2]$ and suppose

$$\varepsilon \le p_k \le 1 - \varepsilon \qquad (k \ge 1). \tag{5.8}$$

For the Cantor realization μ_p of D-I.4.07, define

$$a = \liminf_{n\to\infty} \frac{\sum_{k=1}^{n} h(p_k)}{n \log 3}, \qquad b = \limsup_{n\to\infty} \frac{\sum_{k=1}^{n} h(p_k)}{n \log 3}. \tag{5.9}$$

Then for μ_p-almost every x,

$$\underline{d}_{\mu_p}(x) = a, \qquad \overline{d}_{\mu_p}(x) = b. \tag{5.10}$$

Consequently

$$\underline{\dim}_{\mathrm{H}}\mu_p = \overline{\dim}_{\mathrm{H}}\mu_p = a, \qquad \underline{\dim}_{\mathrm{P}}\mu_p = \overline{\dim}_{\mathrm{P}}\mu_p = b, \tag{5.11}$$

and

$$\underline{\dim}_{\mathrm{e}}\mu_p = a, \qquad \overline{\dim}_{\mathrm{e}}\mu_p = b. \tag{5.12}$$

Proof

First observe that the coding map is bi-Lipschitz. If ω, τ first differ at coordinate k, the leading difference in their coded values has absolute value $2 \cdot 3^{-k}$, while the entire tail has absolute value at most

$$\sum_{j>k} 2 \cdot 3^{-j} = 3^{-k}.$$

Since $d_3(\omega, \tau) = 3^{1-k}$, it follows that

$$\frac{1}{3} d_3(\omega, \tau) \le |\pi(\omega) - \pi(\tau)| \le d_3(\omega, \tau). \tag{5.13}$$

The closed d_3-ball of radius 3^{-n} at ω is exactly the cylinder $[\omega|n]$.
On the product probability space define

$$Y_k(\omega) = -(1 - \omega_k) \log p_k - \omega_k \log(1 - p_k). \tag{5.14}$$

The variables Y_k are independent,

$$\mathbb{E}Y_k = h(p_k), \tag{5.15}$$

and (5.8) bounds $Y_k - EY_k$ uniformly. Lemma 5.2 gives

$$\frac{1}{n}\left(-\log \nu_p([\omega|n]) - \sum_{k=1}^{n} h(p_k)\right) \longrightarrow 0 \tag{5.16}$$

for ν_p-almost every ω. Equations (5.9), (5.16), and Lemma 1.1 compute the lower and upper local dimensions in the symbolic metric as a and b. The ball inclusions from the bi-Lipschitz estimate (5.13) preserve these limits under π, proving (5.10). The two dictionaries give (5.11).

It remains to audit entropy. A single sequence has mass at most $(1-\varepsilon)^n$ in every length-n cylinder, hence has zero mass. The countable collection of adic boundary points is therefore null. Up to this null set, the occupied atoms of the level-n triadic partition are precisely the images of length-n cylinders. Product expansion of entropy gives

$$H(\mu_p, \mathscr{D}_n^{(3)}) = \sum_{k=1}^{n} h(p_k). \tag{5.17}$$

For completeness, (5.17) follows inductively: splitting every cylinder of mass w into masses wp_{n+1} and $w(1-p_{n+1})$ increases its entropy by $wh(p_{n+1})$; summing w gives one. Divide (5.17) by $n \log 3$, use (5.9), and invoke base independence from Lemma 4.1. This proves (5.12). □

Corollary — 5.4 (C-I.4.16: explicit separation examples)

Let $0 < p < q \le 1/2$. There is a compactly supported probability μ such that

$$\underline{\dim}_{\mathrm{H}}\mu = \overline{\dim}_{\mathrm{H}}\mu = \frac{h(p)}{\log 3} < \frac{h(q)}{\log 3} = \underline{\dim}_{\mathrm{P}}\mu = \overline{\dim}_{\mathrm{P}}\mu. \tag{5.18}$$

Moreover, there is a compactly supported probability for which all four measure dimensions are strictly different in the order displayed in (3.1).

Proof

Choose consecutive blocks of positive integer lengths L_j such that

$$L_j \ge j(L_1 + \cdots + L_{j-1}) \qquad (j \ge 2). \tag{5.19}$$

Set $p_k = p$ on odd blocks and $p_k = q$ on even blocks. At the end of block j, the proportion of all earlier coordinates is at most $1/(j+1)$. Hence the average of $h(p_k)$ tends along odd block endpoints to $h(p)$ and along even endpoints to $h(q)$. Every summand lies between these two values, so

$$\liminf_n \frac{1}{n}\sum_{k=1}^{n} h(p_k) = h(p), \qquad \limsup_n \frac{1}{n}\sum_{k=1}^{n} h(p_k) = h(q). \tag{5.20}$$

Theorem 5.3 gives (5.18). The inequality is strict because

$$h'(u) = \log\frac{1-u}{u} > 0 \qquad (0 < u < 1/2).$$

For the fourfold separation, choose

$$0 < p_1 < p_2 < q_1 < q_2 \le 1/2.$$

Using the same blocks, form a measure μ_i with alternating probabilities p_i, q_i, for $i = 1, 2$. Translate μ_2 by two units and put

$$\rho = \theta\mu_1 + (1-\theta)(T_{\#}\mu_2), \qquad 0 < \theta < 1, \quad T(x) = x + 2.$$

The supports are positively separated. As in the local calculation in Theorem 5.1, multiplication by a mixture weight does not change either local dimension on its component. Theorem 5.3 therefore says that the essential values of the lower local dimension are

$$a_i = \frac{h(p_i)}{\log 3},$$

and those of the upper local dimension are

$$b_i = \frac{h(q_i)}{\log 3}.$$

The dictionaries yield

$$\underline{\dim}_{\mathrm{H}}\rho = a_1, \quad \overline{\dim}_{\mathrm{H}}\rho = a_2, \quad \underline{\dim}_{\mathrm{P}}\rho = b_1, \quad \overline{\dim}_{\mathrm{P}}\rho = b_2.$$

Strict monotonicity of h gives $a_1 < a_2 < b_1 < b_2$, completing the example. □

I.4.6 Intuition and strategy

The following discussion explains the geometric ideas behind the proofs.

I.4.6.1 Four global dimensions answer four questions

There are two independent choices. One may measure carriers by Hausdorff or packing dimension, and one may ask for a carrier of positive mass or of full mass. This produces four dimensions:

$$\underline{\dim}_{\mathrm{H}}\mu, \quad \overline{\dim}_{\mathrm{H}}\mu, \quad \underline{\dim}_{\mathrm{P}}\mu, \quad \overline{\dim}_{\mathrm{P}}\mu.$$

The lower value detects the smallest-dimensional component with positive mass. The upper value must accommodate every component except a null set. A mixture of two exact measures therefore turns minimum and maximum into different answers.

I.4.6.2 Why lower local dimension belongs to Hausdorff dimension

If the lower local dimension exceeds s, then eventually

$$\mu(B(x,r)) \le r^s.$$

After passing to a positive-mass subset with a uniform scale threshold, this is a Frostman bound and forces Hausdorff dimension at least s.

In the reverse direction, a point whose lower local dimension is below s has arbitrarily small balls with mass larger than r^s. A disjoint-ball selection converts those heavy balls into a Hausdorff cover with bounded s-cost. The alternation between “eventually” and “infinitely often” is the liminf mechanism.

I.4.6.3 Why upper local dimension belongs to packing dimension

When the upper local dimension is below s, every sufficiently small ball at the point is heavy. A separated set cannot contain many such points because the corresponding half-radius balls are disjoint and their masses sum to at most one. This produces an upper-box bound on countably many pieces and hence a packing bound.

For the converse, a bounded set of upper box dimension below s has too few balls at dyadic scales to contain positive measure of points whose balls are exceptionally light infinitely often. The estimates form a summable series. This is the technical reason the packing dictionary is not merely the Hausdorff proof with names exchanged.

I.4.6.4 Entropy is an average-scale dimension

The quantity

$$H(\mu, \mathscr{D}_n) = \int -\log \mu(\mathscr{D}_n(x))\, d\mu(x)$$

is the expected number of natural-log units needed to identify the scale-n cell of a random point. Normalizing by $n \log b$ converts this information to a dimension.

Low entropy along a subsequence places a fixed positive mass in relatively few cubes; a limsup of those cubes is a positive-mass set of small Hausdorff dimension. Conversely, a full-mass set covered by pieces of low upper box dimension forces most entropy to live on few cubes; the small remaining mass can contribute at most its proportion times the ambient dimension. This is the mechanism behind

$$\underline{\dim}_{\mathrm{H}}\mu \le \underline{\dim}_{\mathrm{e}}\mu \le \overline{\dim}_{\mathrm{e}}\mu \le \overline{\dim}_{\mathrm{P}}\mu.$$

Comparable grids have only bounded conditional ambiguity, so the integer base changes entropy by $O(1)$ rather than by a linear amount.

I.4.6.5 Exactness, mixtures, and oscillation

An exact-dimensional measure has one scaling exponent at almost every point. All four global dimensions and both entropy dimensions are then forced to that exponent.

A separated mixture retains the local exponent of each component. Its lower dimensions take the minimum, its upper dimensions take the maximum, while entropy averages the component dimensions with the mixture weights.

An inhomogeneous Bernoulli product illustrates a different failure. Long blocks with two weight distributions cause the typical logarithmic cylinder mass to oscillate between two slopes. The Hausdorff dimensions record the liminf slope and the packing dimensions the limsup slope. Mixing two such measures with separated supports makes all four global dimensions distinct.

I.4.7 Exercises

Every exercise uses only D/R-I.1–I.4 and A-RA.

I.4.7.1 Exercise 1 — Atoms and points outside the support

Let μ be a probability measure on $\mathbb{R}^d$.

1. Prove that if $\mu(\{x\}) > 0$, then $\underline{d}_\mu(x) = \overline{d}_\mu(x) = 0$.
2. Prove that if x lies outside $\operatorname{spt}\mu$, then both local dimensions are $+\infty$.

I.4.7.2 Exercise 2 — Equivalent metrics

Let μ be a Borel probability measure on a space X with two metrics d_1, d_2 satisfying

$$c\, d_1(x, y) \le d_2(x, y) \le C\, d_1(x, y)$$

for all $x, y \in X$, with $0 < c \le C < +\infty$. These metrics give the same topology and therefore the same Borel sets. Define local dimensions by D-I.4.01's formula, using closed balls for the respective metric. Prove that every lower and upper local dimension is the same for the two metrics.

I.4.7.3 Exercise 3 — Countably supported measures

Let μ be a probability carried by a countable subset of $\mathbb{R}^d$. Prove that all four measure dimensions are zero. If the support is compact, prove that its entropy dimension also exists and equals zero.

I.4.7.4 Exercise 4 — Ahlfors-regular measures

Assume μ is compactly supported and there are $\alpha \ge 0$, $r_0 > 0$, and $0 < c \le C < +\infty$ such that

$$cr^\alpha \le \mu(\overline{B}(x, r)) \le Cr^\alpha$$

for $x \in \operatorname{spt}\mu$ and $0 < r < r_0$. This two-sided bound is called **α-Ahlfors regularity** of μ at small scales. Prove that μ is exact dimensional of dimension α and identify all six global dimensions in I.4.

I.4.7.5 Exercise 5 — Strict dimension loss under a Lipschitz map

Let μ be normalized two-dimensional Lebesgue measure on $[0, 1]^2$, and let $f(x, y) = x$. Prove that μ is exact dimensional of dimension two while $f_\#\mu$ is exact dimensional of dimension one. Thus every inequality in the Lipschitz part of T-I.4.09 can be strict.

I.4.7.6 Exercise 6 — Entropy on a finite alphabet

Let $(p_1, \dots, p_N)$ be a probability vector. Prove

$$-\sum_{i=1}^{N} p_i \log p_i \le \log N,$$

with equality exactly when $p_i = 1/N$ for every $1 \le i \le N$. Use this to rederive the bounded-intersection inequality (4.1).

I.4.7.7 Exercise 7 — Finite separated mixtures

Let μ_i, $1 \le i \le N$, be compactly supported exact-dimensional probability measures with pairwise positively separated supports and dimensions α_i. For positive weights p_i summing to one, compute all four set-carried dimensions and the entropy dimension of $\sum_i p_i \mu_i$.

I.4.7.8 Exercise 8 — An atom plus an interval

For $0 < p < 1$, define

$$\mu = p\delta_0 + (1-p)\lambda|_{[1,2]}.$$

Compute all four set-carried dimensions and the entropy dimension. Explain why this one example distinguishes minimum, maximum, and average behavior.

I.4.7.9 Exercise 9 — Translated adic grids

Translate every b-adic grid by a fixed vector $v \in \mathbb{R}^d$. Prove that the lower and upper entropy dimensions computed from the translated grids agree with those in D-I.4.06.

I.4.7.10 Exercise 10 — Periodic inhomogeneous Bernoulli products

Assume $p_{k+m} = p_k$ for some positive integer m, with all weights bounded away from zero and one. Prove that μ_p is exact dimensional and

$$\dim \mu_p = \dim_{\mathrm{e}} \mu_p = \frac{h(p_1) + \cdots + h(p_m)}{m \log 3}.$$

I.4.7.11 Exercise 11 — A concrete oscillating product

Use the dominating-block construction with probabilities $1/4$ and $1/2$. Compute the two almost-sure local dimensions, all four set-carried dimensions, and the lower and upper entropy dimensions. Prove that the measure is not exact dimensional.

I.4.7.12 Exercise 12 — The middle dimensions are incomparable

Give one compactly supported probability with

$$\overline{\dim}_{\mathrm{H}}\mu > \underline{\dim}_{\mathrm{P}}\mu$$

and another with the reverse strict inequality. Your examples must follow from proved results in this chapter, not from an external construction.

Chapter I.5

Assouad Dimensions, Tangents, and Microsets

Hausdorff and box dimension describe global size, but a small region can be much more crowded than the set looks globally. Assouad dimension measures the largest local covering growth between two scales; lower dimension measures the smallest. Dimension spectra tie those scales together and reveal information lost by either endpoint. We first develop the covering estimates and examples, then study what survives when a set is magnified and a limiting picture is taken. The distinction between similarities and the more restrictive homotheties matters for the tangent and microset conventions below.

The definitions below fix this chapter's conventions. Geometric intuition and proof strategy are discussed separately after the main development; exercises and their solutions provide a second route through the material.

Definitions and notation

Throughout, $d \geq 1$, Euclidean balls are closed, and N_r is the centred covering number from D-I.2.02. The observation window is

$$W = [0,1]^d, \qquad W^\circ = (0,1)^d.$$

Definition — I.5.01: Local covering convention

For a nonempty bounded set $F \subset \mathbb{R}^d$, $x \in F$, and $0 < r < R$, the **local covering number at** $(x; R, r)$ is

$$\mathcal{N}_F(x; R, r) = N_r(F \cap \overline{B}(x, R)).$$

Write $|F| = \operatorname{diam} F$. A set is called **nondegenerate** here if it contains at least two points. If F is nondegenerate, every Assouad or lower-dimension quantifier below uses

$$0 < r < R \le |F|.$$

The empty set and every singleton are assigned all dimensions in D-I.5.02–D-I.5.04 equal to zero. Thus a zero diameter never occurs in a scale denominator.

Definition — I.5.02: Assouad and lower dimensions

For a nondegenerate bounded F, its **Assouad dimension** is

$$\dim_{\mathrm{A}} F = \inf\left\{ s \ge 0 \ : \ \begin{array}{l} \text{there is } C \ge 1 \text{ such that for every } x \in F \text{ and} \\ 0 < r < R \le |F|, \quad \mathcal{N}_F(x; R, r) \le C(R/r)^s \end{array} \right\}.$$

Its **lower dimension** is

$$\dim_{\mathrm{L}} F = \sup\left\{ s \ge 0 \ : \ \begin{array}{l} \text{there is } c > 0 \text{ such that for every } x \in F \text{ and} \\ 0 < r < R \le |F|, \quad \mathcal{N}_F(x; R, r) \ge c(R/r)^s \end{array} \right\}.$$

The constants are uniform in the point and both scales.

Definition — I.5.03: Modified lower dimension

For nonempty bounded F, define its **modified lower dimension** by

$$\dim_{\mathrm{ML}} F = \sup\{\dim_{\mathrm{L}} E \ : \ \varnothing \ne E \subset F\}.$$

Set $\dim_{\mathrm{ML}} \varnothing = 0$, consistently with D-I.5.01. Every nonempty subset of a bounded Euclidean set is totally bounded, so its lower dimension is defined by D-I.5.02. This is the standard monotone modification; no closure is inserted into the supremum.

Definition — I.5.04: Assouad/lower spectra and scale profile

Let F be nondegenerate and bounded and let $0 < \theta < 1$. Put

$$r = R^{1/\theta} \qquad (0 < R < \min\{1, |F|\}),$$

so $0 < r < R$. The **Assouad spectrum** is

$$\dim_{\mathrm{A}}^{\theta} F = \inf\left\{ s \ge 0 \ : \ \begin{array}{l} \text{there is } C \ge 1 \text{ such that for every } x \in F \text{ and} \\ 0 < R < \min\{1, |F|\}, \quad N_{R^{1/\theta}}(F \cap \overline{B}(x, R)) \le C(R/R^{1/\theta})^s \end{array} \right\}.$$

The **lower spectrum** is

$$\dim_{\mathrm{L}}^{\theta} F = \sup\left\{ s \ge 0 \ : \ \begin{array}{l} \text{there is } c > 0 \text{ such that for every } x \in F \text{ and} \\ 0 < R < \min\{1, |F|\}, \quad N_{R^{1/\theta}}(F \cap \overline{B}(x, R)) \ge c(R/R^{1/\theta})^s \end{array} \right\}.$$

The constant is uniform in the point and scale. The **Assouad scale profile** of F is the function

$$\theta \longmapsto \left(\dim_{\mathrm{L}}^{\theta} F, \dim_{\mathrm{A}}^{\theta} F\right), \qquad 0 < \theta < 1.$$

This phrase does not denote a Falconer–Howroyd capacity or projection dimension profile.

Definition — I.5.05: Hausdorff metric and compact hyperspace

For a metric space (X, d_X) and nonempty compact $A, B \subset X$, define

$$d_{\mathrm{H}}(A, B) = \max\left\{\sup_{a\in A} \operatorname{dist}(a, B), \sup_{b\in B} \operatorname{dist}(b, A)\right\}.$$

Here distances are computed with d_X. The **compact hyperspace** $K(X)$ is the set of all nonempty compact subsets of X, equipped with d_H. The displayed function is a metric: symmetry is immediate; zero distance forces equality because the two sets are closed; the triangle inequality follows by approximating each point through the intermediate set and then taking suprema. It is finite because compact sets are bounded. In particular, $K(W)$ uses Euclidean distance. Convergence in a compact hyperspace always means Hausdorff metric convergence. The name does not assert that $K(X)$ itself is compact for arbitrary X; compactness for the observation window is proved in L-I.5.09.

Definition — I.5.06: Similarities, homotheties, and weak tangents

A **similarity** of $\mathbb{R}^d$ is a map

$$T(x) = cOx + t,$$

where $c > 0$, O is orthogonal, and $t \in \mathbb{R}^d$. Its **similarity ratio** is c. It is a **homothety** when O is the identity.
Let $F \subset \mathbb{R}^d$ be nonempty and compact. A set $E \in K(W)$ is a **windowed weak tangent** of F if $E \cap W^{\circ}$ is nonempty and there are similarities T_k with ratios $c_k \to \infty$ such that every $T_k(F) \cap W$ is nonempty and

$$d_{\mathrm{H}}\big(E, T_k(F) \cap W\big) \longrightarrow 0.$$

The terms **weak tangent** and **tangent** mean this stricter windowed notion inside the series. A source allowing bounded ratios, zooming out, or limits confined to the boundary uses a broader convention and must be labelled as such.

Definition — I.5.07: Strict minisets, microsets, and star dimension

For nonempty compact $F \subset \mathbb{R}^d$, a **strict miniset** is a nonempty set

$$(\lambda F + t) \cap W, \qquad \lambda \geq 1, \quad t \in \mathbb{R}^d.$$

A set $E \in K(W)$ is a **strict microset** of F if $E \cap W^\circ$ is nonempty and there are $\lambda_k \to \infty$ and $t_k \in \mathbb{R}^d$ such that

$$(\lambda_k F + t_k) \cap W \longrightarrow E \quad \text{in } d_{\mathrm{H}}.$$

The **star dimension** used in this series is

$$\dim_* F = \sup\{\dim_{\mathrm{H}} E \,:\, E \text{ is a strict microset of } F\}.$$

Every strict microset is a windowed weak tangent obtained by homotheties.

Definition — I.5.08: Furstenberg minisets, galleries, and grid exponent

Let $F \in K(W)$. A **Furstenberg miniset** of F is any $A \in K(W)$ for which

$$A \subset cF + t$$

for some $c \geq 1$ and $t \in \mathbb{R}^d$. A **Furstenberg microset** is a Hausdorff limit of Furstenberg minisets. This subset convention is broader than the strict-intersection convention in D-I.5.07.
For $F \in K(W)$, its **Furstenberg microset family** is

$$\mathscr{G}_F = \left\{ A \in \mathscr{K}(W) \,:\, \begin{array}{l} \text{there are } A_j \in \mathscr{K}(W),\ c_j \geq 1,\ t_j \in \mathbb{R}^d \text{ such that} \\ A_j \subset c_j F + t_j \text{ and } d_{\mathrm{H}}(A_j, A) \to 0 \end{array} \right\}.$$

Thus G_F is precisely the family of all Furstenberg microsets of F. The fact that it is a gallery is part of the cited result G-I.5.14, not part of this definition.
A nonempty family $G \subset K(W)$ is a **gallery** if it is closed in $K(W)$ and contains every Furstenberg miniset of each of its members.
Let D_n be the half-open dyadic cubes

$$2^{-n}(k + [0,1)^d), \qquad k \in \mathbb{Z}^d,$$

and put

$$Q_n(A) = \#\{Q \in \mathscr{D}_n \,:\, Q \cap A \neq \varnothing\}.$$

The **dyadic gallery exponent** is

$$\Delta(\mathscr{G}) = \limsup_{n\to\infty} \frac{1}{n} \log_2 \left(\sup_{A \in \mathscr{G}} Q_n(A) \right).$$

Throughout, F, E are nonempty compact Euclidean sets unless another hypothesis is stated. Balls are closed, N_r and M_r are the centred covering and separated-set numbers of D-I.2.02, and

$$W = [0,1]^d.$$

I.5.1 Local covering calculus

Lemma — L-I.5.01 — Local covering calculus and dyadic equivalence [F]

For nonempty totally bounded metric sets A, B and positive radii (with one common ambient metric for the union):

$$r \le R \quad \Longrightarrow \quad N_R(A) \le N_r(A), \tag{1.1}$$

$$N_r(A \cup B) \le N_r(A) + N_r(B), \tag{1.2}$$

and, with the max product metric,

$$N_r(A \times B) \le N_r(A) N_r(B). \tag{1.3}$$

If A, B are nonempty compact subsets of one metric space and $d_H(A,B) < \varepsilon$, then

$$N_{r+2\varepsilon}(A) \le N_r(B), \qquad N_{r+2\varepsilon}(B) \le N_r(A). \tag{1.4}$$

Two further comparisons will keep track of the centres. For nonempty totally bounded sets $A \subset B$ in one metric space,

$$N_{2r}(A) \le N_r(B). \tag{1.4a}$$

For a nonempty bounded $A \subset \mathbb{R}^d$, $r > 0$, and $\lambda \ge 1$, put $D_d(\lambda) = (\lceil 2\lambda\sqrt{d}\rceil + 1)^d$. Then

$$N_r(A) \le D_d(\lambda) N_{\lambda r}(A). \tag{1.4b}$$

In particular, if $A \subset B \subset \mathbb{R}^d$ are nonempty and bounded, then $N_r(A) \le D_d(2) N_r(B)$. Unlike (1.1), this is not exact monotonicity: a fixed recentering factor is necessary. For example, in the real line $A = \{-1, 1\}$ and $B = \{-1, 0, 1\}$ satisfy $N_1(A) = 2$ but $N_1(B) = 1$.
Here is a precise dyadic version of the local dimension estimates. Let $F \subset \mathbb{R}^d$ be compact with $D = \operatorname{diam} F > 0$, and fix $s \ge 0$. For every integer m, let $\mathscr{D}_m$ be the half-open cubes $2^{-m}(k + [0,1)^d)$, $k \in \mathbb{Z}^d$. If $Q \in D_m$ and $n > m$, put

$$M_n(Q; F) = \#\{Q' \in \mathscr{D}_n : Q' \subset Q, Q' \cap F \ne \varnothing\}, \qquad n > m, \tag{1.5}$$

The exponent-s Assouad upper estimate in D-I.5.02 is equivalent to the existence of $C > 0$ such that

$$M_n(Q; F) \le C 2^{(n-m)s} \quad (m \in \mathbb{Z},\ 2^{-m} \le D,\ n > m,\ Q \in \mathscr{D}_m). \tag{1.6}$$

For the lower estimate, define the finite family of adjacent parents

$$\mathscr{A}_m(x) = \{Q \in \mathscr{D}_m : Q \cap \overline{B}(x, 2^{-m}) \neq \varnothing\}. \tag{1.6a}$$

It has at most 4^d members. The exponent-s lower estimate in D-I.5.02 is equivalent to the existence of $c > 0$ such that

$$\max_{Q \in \mathscr{A}_m(x)} M_n(Q; F) \geq c2^{(n-m)s} \quad (x \in F,\ m \in \mathbb{Z},\ 2^{-m} \leq D,\ n > m). \tag{1.6b}$$

The parent attaining the maximum may depend on x, m, n. No lower bound in every individual occupied parent is asserted. The constants may change between the ball and dyadic formulations; their changes depend only on d, s and the original constant. The distinction is essential even for $F = [0, 1]$: the parent $[1, 2)$ meets F only at 1, so it has just one occupied descendant at every level, although the interval has lower dimension one. An adjacent parent captures the interval's local complexity.

Proof

Equations (1.1) and (1.2) follow by retaining a finer cover and by joining two covers. If A and B are covered by r-balls with centres a_i and b_j, then the products of those balls are r-balls for the max metric and cover $A \times B$; this proves (1.3).

Assume $d_H(A, B) < \varepsilon$ and let

$$B \subset \bigcup_{j=1}^{N} B(b_j, r), \qquad b_j \in B, \qquad N = N_r(B).$$

Choose $a_j \in A$ with $d(a_j, b_j) < \varepsilon$. If $a \in A$, choose $b \in B$ with $d(a, b) < \varepsilon$, then choose j with $d(b, b_j) \leq r$. The triangle inequality gives

$$d(a, a_j) < r + 2\varepsilon.$$

Thus the balls $B(a_j, r + 2\varepsilon)$ cover A, proving the first inequality in (1.4); symmetry proves the second.

For (1.4a), start with a centred r-cover of B. Discard balls missing A, and from each remaining intersection choose a point of A. Its $2r$-ball covers that intersection by the triangle inequality. To prove (1.4b), enclose each ball of a centred λr-cover of A in an axis-parallel cube of side $2\lambda r$. Partition the cube into at most $D_d(\lambda)$ cells of diameter at most r; assign boundary points to one cell. For each cell meeting the covered part of A, choose a centre in that intersection. The resulting r-balls cover A. Combining (1.4a), with r replaced by $r/2$, and (1.4b) for B gives the stated fixed-factor subset comparison.

A ball of radius ℓ meets at most 4^d half-open grid cubes of side ℓ. A cube of side ℓ has diameter at most $\sqrt{d}\ell$. Both facts follow coordinate by coordinate and include points on grid boundaries.

Suppose first that the exponent-s ball upper estimate holds. Write $\ell = 2^{-m}$, $h = 2^{-n}$, and choose $x \in Q \cap F$ if that set is nonempty. It is contained in the local piece with outer radius $\min\{\sqrt{d}\ell, D\}$. Cover that piece by centred h-balls. Each such ball meets at most 4^d level-n cubes, so

$$M_n(Q;F) \le 4^d C(\sqrt{d}\,\ell/h)^s.$$

Here $h < \ell \le D$, so the inner radius is admissible. Empty parents have count zero. This proves (1.6).

Conversely suppose (1.6). For $0 < r < R \le D/2$, choose dyadic lengths

$$R \le \ell < 2R, \qquad \frac{r}{2\sqrt{d}} < h \le \frac{r}{\sqrt{d}}. \tag{1.7}$$

Their levels satisfy $n > m$. At most 4^d level-m parents meet $B(x,R)$. In each of their occupied level-n cubes meeting the target local piece, choose a point of that piece. Its r-ball covers the intersection because the cube diameter is at most r. Consequently

$$N_r(F \cap B(x,R)) \le 4^d C(\ell/h)^s \le 4^d C(4\sqrt{d})^s (R/r)^s.$$

For $D/2 < R \le D$, cover F by at most $D_d(4)$ centred $D/4$-balls, using (1.4b) and $N_D(F) = 1$. If $r < D/4$, cover each of those local pieces at radius $r/2$ by the estimate just proved, and recenter retained balls in $F \cap B(x,R)$. This gives the same exponent with a constant depending only on d, s, C. If $r \ge D/4$, (1.4b) gives a uniform bound directly for the target piece. Thus all outer scales are covered.

Now assume the exponent-s ball lower estimate. Fix x, m, n as in (1.6b), and again put $\ell = 2^{-m}$, $h = 2^{-n}$. When $\sqrt{d}h < \ell$, the occupied level-n cubes meeting $F \cap B(x,\ell)$ give a centred $\sqrt{d}h$-cover of this set. Hence their number is at least $c(\ell/(\sqrt{d}h))^s$. They lie in at most 4^d parents in $A_m(x)$, so one parent has at least a 4^{-d} fraction of this count. When $\sqrt{d}h \ge \ell$, there is at least one occupied fine cube, while $(\ell/h)^s \le d^{s/2}$. Reducing the lower constant handles this case too. This proves (1.6b), uniformly in all its variables.

Finally suppose (1.6b). Set $a = 1 + \sqrt{d}$. Given $x \in F$ and $0 < r < R \le D$, choose dyadic lengths

$$\frac{R}{2a} < \ell \le \frac{R}{a}, \qquad r \le h < 2r.$$

Every parent in $A_m(x)$ lies in $B(x, a\ell)$, hence in $B(x,R)$. If $h < \ell$, use (1.6b) at these levels. Each counted fine cube contains a point of the target local piece, and an r-ball meets at most 4^d fine cubes because $h \ge r$. Any centred r-cover therefore has at least

$$4^{-d} c(\ell/h)^s \ge 4^{-d} c(4a)^{-s} (R/r)^s$$

members. If $h \ge \ell$, including the equal-level case $n = m$, then $R/r < 4a$; the estimate follows instead from $N_r \ge 1$. This proves the lower equivalence without any assertion about a fixed individual parent. □

Lemma — L-I.5.02 — Local Frostman tree from lower homogeneity [F]

Assume there are $c > 0$ and $t > 0$ such that

$$N_r(F \cap B(x,R)) \ge c(R/r)^t \tag{1.8}$$

whenever $x \in F$ and $0 < r < R \le |F|$. If $0 \le s < t$, then there is a constant

$C = C(c, s, t, d)$ such that, for every $x \in F$ and every $0 < R \leq |F|$, there is a probability measure $\mu_{x,R}$ carried by $F \cap B(x, R)$ satisfying

$$\mu_{x,R}(B(y,r)) \leq C(r/R)^s \qquad (y \in \mathbb{R}^d, 0 < r < R). \tag{1.9}$$

Proof

For $s = 0$, take any point mass in $F \cap B(x, R)$ and $C = 1$. Assume $0 < s < t$. Choose an integer M so large that, with

$$q = (32M)^{-1}, \qquad b = \left\lfloor c(4M)^t \right\rfloor, \tag{1.10}$$

we have

$$b \geq 2, \qquad b \geq q^{-s}. \tag{1.11}$$

This is possible because $t > s$.

The construction is indexed by finite words over $\{1, \dots, b\}$. A word of length n is an ordered n-tuple of these symbols; the unique length-zero word is denoted by $\emptyset$. Appending j to u gives uj. An infinite sequence of symbols is a branch, and its length-n cylinder is the set of all branches having those first n symbols.

Put $\rho_0 = R/4$ and choose the root centre $x_\emptyset = x$. Suppose a centre $x_u \in F$ and a radius $\rho_n = q^n \rho_0$ have been chosen. Apply (1.8) inside $B(x_u, \rho_n/2)$ at scale

$$a_n = \frac{\rho_n}{8M} = 4\rho_{n+1}.$$

Using $N_{a_n} \leq M_{a_n}$ from L-I.2.01 gives an a_n-separated subset of cardinality at least

$$c\left(\frac{\rho_n/2}{a_n}\right)^t = c(4M)^t \geq b.$$

Choose exactly b of its points and call them $x_{u1}, \dots, x_{ub}$. The balls $B(x_{uj}, \rho_{n+1})$ lie in $B(x_u, \rho_n)$, and sibling centres are more than $4\rho_{n+1}$ apart. Iteration produces a rooted b-ary family of nested balls. If two level-n branches first split at level $j + 1$, each later centre moves from its level-$j + 1$ centre by at most

$$\frac{1}{2} \sum_{i=j+1}^{n-1} \rho_i \leq \frac{\rho_{j+1}}{2(1-q)}.$$

Since $q \leq 1/32$, their level-n centres are separated by more than $2\rho_{j+1} \geq 2\rho_n$. Thus the open construction balls at a fixed level have disjoint interiors. If a ball of radius at most ρ_n meets a level-n construction ball, its centre lies in a ball of radius $2\rho_n$. The disjoint balls of radius $\rho_n/2$ about all such centres lie in a ball of radius $5\rho_n/2$. Comparing Euclidean volumes bounds their number by $A_d = 5^d$, independently of n.

Give each length-n cylinder mass b^{-n}. For an explicit realization, take Lebesgue probability on $[0, 1)$ and use the symbols

$$j_k(u) = 1 + \lfloor b^k u \rfloor - b\lfloor b^{k-1}u \rfloor \quad (k \geq 1).$$

Each specified prefix of length n corresponds to one half-open interval of length b^{-n}. Send u to the unique point in the nested balls along this branch, and call the push-forward probability μ. The map is Borel: the level-n centre is a finite-valued Borel function of u, and these centres converge pointwise. The radii tend to zero and all centres lie in the closed set F, so the limit set is contained in $F \cap B(x, R)$.

For $R/4 < r < R$, the probability bound gives $\mu(B(y, r)) \leq 1 \leq 4^s(r/R)^s$, which is covered by the constant below. For $0 < r \leq R/4$, choose an integer $n \geq 0$ so that

$$\rho_{n+1} < r \leq \rho_n.$$

The packing observation gives

$$\mu(B(y, r)) \leq A_d b^{-n} \leq A_d q^{ns}.$$

Since $\rho_n = q^n R/4$ and $r > q\rho_n$,

$$q^{ns} = \left(\frac{4\rho_n}{R}\right)^s \leq 4^s q^{-s}(r/R)^s.$$

Thus (1.9) holds with $C = A_d 4^s q^{-s}$. □

Theorem — T-I.5.03 — Dimension hierarchy [F]

For every nonempty compact $F \subset \mathbb{R}^d$,

$$0 \leq \dim_L F \leq \dim_H F \leq \dim_P F \leq \overline{\dim}_B F \leq \dim_A F \leq d, \tag{1.12}$$

and

$$\dim_L F \leq \underline{\dim}_B F \leq \overline{\dim}_B F. \tag{1.13}$$

Proof

Only $\dim_L F \leq \dim_H F$, the comparisons with Assouad dimension, and the ambient upper bound are new. The first inequality is trivial when $\dim_L F = 0$. Otherwise let $0 \leq s < \dim_L F$ and choose $s < t < \dim_L F$. The defining lower estimate holds with exponent t, so L-I.5.02 supplies an s-Frostman probability carried by F. The mass-distribution principle T-I.1.08 yields $\dim_H F \geq s$. Letting s increase to $\dim_L F$ proves the first new inequality.

The inequalities

$$\dim_H F \leq \dim_P F \leq \overline{\dim}_B F$$

are T-I.2.11. If $D = |F| > 0$, an Assouad estimate at the fixed outer scale $R = D$ gives

$$N_r(F) \leq C(D/r)^u$$

for every $u > \dim_A F$; hence $\overline{\dim}_B F \leq \dim_A F$. For the local piece $A = F \cap B(x, R)$, one has $N_R(A) = 1$, with centre x. Equation (1.4b), with $\lambda = R/r$, gives $N_r(A) \leq D_d(R/r) \leq C_d(R/r)^d$. All centres are in A, and this proves $\dim_A F \leq d$.

Finally, the lower estimate at $R = D$ gives

$$N_r(F) \geq c(D/r)^s$$

for every $s < \dim_L F$. Taking the lower logarithmic limit proves $\dim_L F \leq \underline{\dim}_B F$. The remaining inequality in (1.13) is definitional. Degenerate compact sets satisfy all claims by convention. □

Theorem — T-I.5.04 — Stability, nonmonotonicity, and products [F]

Assouad dimension is monotone, finitely stable, closure-stable, and bi-Lipschitz invariant on nonempty bounded Euclidean sets. Lower dimension is bi-Lipschitz invariant but is not monotone. Modified lower dimension is monotone and bi-Lipschitz invariant, with

$$\dim_L F \leq \dim_{ML} F \leq \dim_A F. \tag{1.14}$$

For the max product metric,

$$\dim_A(E \times F) \leq \dim_A E + \dim_A F, \tag{1.15}$$

$$\dim_L(E \times F) \geq \dim_L E + \dim_L F. \tag{1.16}$$

Proof

If $E \subset F$, cover $F \cap B(x, R)$ by $r/2$-balls. Retain only balls meeting $E \cap B(x, R)$ and recenter each retained ball at such a point of E; the resulting r-balls are a centred cover of the smaller local piece. This fixed change of inner scale proves Assouad monotonicity. For a finite union, if a local ball centred in $E \cup F$ meets both sets, choose one point of each nonempty intersection. Each part lies in a ball of radius $2R$ centred in its own set. Covering the two parts separately, making the same harmless recentering, and using (1.2) gives

$$\dim_A(E \cup F) = \max\{\dim_A E, \dim_A F\}.$$

Scales larger than the diameter of one component are absorbed into the constant.

For closure stability, monotonicity gives $\dim_A A \leq \dim_A \mathrm{cl}(A)$. Conversely, fix $x \in \mathrm{cl}(A)$ and $0 < r < R$. Choose $x_0 \in A$ with $|x - x_0| < R/4$. Every point of $\mathrm{cl}(A) \cap B(x, R)$ is within $r/4$ of a point of $A \cap B(x_0, 2R)$. An $r/4$-cover of the latter set, followed by this approximation and recentering at points of the target local piece, gives

$$N_r\big(\overline{A} \cap B(x, R)\big) \leq N_{r/4}\big(A \cap B(x_0, 2R)\big).$$

An Assouad estimate for A bounds the right side by a constant times $(R/r)^u$; outer scales above $|A|$ are absorbed into the constant. Thus $\dim_A \mathrm{cl}(A) \leq \dim_A A$.

Let T be bi-Lipschitz and choose $K \geq 1$ so that $K^{-1}|x-y| \leq |Tx - Ty| \leq K|x-y|$ on its domain. The inclusions

$$T(F \cap B(x, R/K)) \subset T(F) \cap B(Tx, R) \subset T(F \cap B(x, KR)) \tag{1.17}$$

give the transfer, with the following explicit recentering costs. A centred $r/(2K)$-cover of the right-hand preimage piece maps to balls of radius $r/2$; recenter retained balls in the middle set to obtain a centred r-cover. Thus an Assouad upper estimate transfers with scale ratio at most $2K^2R/r$. If $KR > \operatorname{diam} F$, replace that outer radius by $\operatorname{diam} F$, which covers all of F and only decreases the ratio.

Conversely, pull any centred r-cover of the middle set back to its preimage in F, and recenter on $F \cap B(x, R/K)$. This gives a centred $2Kr$-cover of the latter set, hence

$$N_r(T(F) \cap B(Tx, R)) \geq N_{2Kr}(F \cap B(x, R/K)).$$

When $2Kr < R/K$, the defining lower estimate applies, with ratio $R/(2K^2r)$; here $R \leq \operatorname{diam} T(F) \leq K \operatorname{diam} F$ ensures the outer radius is admissible. When $2Kr \geq R/K$, the ratio R/r is at most $2K^2$, so $N_r \geq 1$ gives the desired estimate after reducing its constant. Apply the same argument to T^{-1} for equality of both dimensions.

The set $[0, 1]$ has lower dimension one by elementary interval covering. The compact superset $[0, 1] \cup \{2\}$ has lower dimension zero: take the centre 2, any $R < 1/2$, and r/R arbitrarily small; the local covering number remains one. Thus lower dimension is not monotone.

Modified lower dimension is monotone directly from its supremum over subsets and is bi-Lipschitz invariant by the preceding paragraph. Its lower bound in (1.14) follows by choosing the subset F. For every nonempty $A \subset F$, one has $\dim_{\mathrm{L}} A \leq \dim_{\mathrm{A}} A$: otherwise choose exponents $v < u$ strictly between the two dimensions and combine the uniform upper and lower estimates to obtain $c(R/r)^u \leq C(R/r)^v$ for arbitrarily large R/r, a contradiction. Assouad monotonicity therefore gives

$$\dim_{\mathrm{L}} A \leq \dim_{\mathrm{A}} A \leq \dim_{\mathrm{A}} F;$$

taking the supremum proves the upper bound.

For (1.15), a max-metric local product ball is exactly the product of its two coordinate local balls. Apply (1.3) to Assouad covers with exponents above the two dimensions and multiply the bounds.

For (1.16), choose a nonnegative admissible lower exponent a for E and b for F. Every nonnegative exponent strictly below a positive lower dimension is admissible, and exponent zero is always admissible. We may use their local estimates up to the product diameter $D = \max\{\operatorname{diam} E, \operatorname{diam} F\}$: above a positive factor diameter D_E, the local piece is all of that factor; for $r < D_E$ use its estimate at D_E and the inequality $(D_E/r)^a \geq (D_E/D)^a(R/r)^a$. For $r \geq D_E$, the ratio R/r is bounded by D/D_E, and $N_r \geq 1$ suffices. A singleton factor has exponent zero and covering number one throughout. The upper estimates used in (1.15) extend to these outer scales by the same argument, or simply by increasing the numerator to R.

When $0 < r < R/2$, in the two coordinate local balls choose maximal $2r$-separated sets. By L-I.2.01 their cardinalities are at least the corresponding N_{2r} values, hence at least constant multiples of $(R/r)^a$ and $(R/r)^b$. Their Cartesian product is $2r$-separated in the max metric. L-I.2.01 then gives

$$N_r\big((E \times F) \cap B((x,y),R)\big) \geq c(R/r)^{a+b}.$$

When $R/2 \leq r < R$, the ratio R/r is bounded and the same estimate follows after reducing c. Let positive-dimensional factors' exponents increase to their lower dimensions, keeping exponent zero for a zero-dimensional factor. This proves (1.16). ☐

Proposition — P-I.5.05 — Polynomial sequences separate the dimensions [F]

For $p > 0$, set

$$F_p = \{0\} \cup \{n^{-p} : n \geq 1\}.$$

Then

$$\dim_{\mathrm{H}} F_p = \dim_{\mathrm{P}} F_p = \dim_{\mathrm{L}} F_p = 0, \tag{1.18}$$

$$\underline{\dim}_{\mathrm{B}} F_p = \overline{\dim}_{\mathrm{B}} F_p = \frac{1}{p+1}, \qquad \dim_{\mathrm{A}} F_p = 1. \tag{1.19}$$

Proof

The set is countable, so its Hausdorff dimension is zero. Packing dimension is countably stable and every singleton has packing dimension zero, proving the packing assertion. Every point n^{-p} is isolated. Centring a ball small enough to contain only that point and then taking r/R arbitrarily small shows that no positive lower exponent is possible; hence $\dim_{\mathrm{L}} F_p = 0$.

The mean-value theorem gives constants $0 < c_p < C_p$ such that

$$c_p n^{-(p+1)} \leq n^{-p} - (n+1)^{-p} \leq C_p n^{-(p+1)} \tag{1.20}$$

for all n. Given small r, choose n comparable to $r^{-1/(p+1)}$. The first n points need at most n balls. The tail lies in $[0, n^{-p}]$ and needs at most

$$1 + Cn^{-p}/r \leq C' r^{-1/(p+1)}$$

balls. This proves the upper box bound. Conversely, by the lower estimate in (1.20), choose, for sufficiently small r,

$$n = \left\lfloor \left(\frac{c_p}{2^{p+3} r}\right)^{1/(p+1)} \right\rfloor \geq 1.$$

Consecutive points in the block $n \leq m \leq 2n$ have gap at least $c_p(2n)^{-(p+1)} \geq 4r$. The block is therefore $2r$-separated and has $n+1$ points, so L-I.2.01 gives the matching lower bound for $N_r(F_p)$. Thus both box dimensions equal $1/(p+1)$.

The ambient bound gives $\dim_{\mathrm{A}} F_p \leq 1$. To prove the reverse inequality, let

$$R_N = N^{-p}, \qquad r_N = \frac{c_p}{4}(2N)^{-(p+1)}.$$

The $N+1$ points n^{-p}, $N \le n \le 2N$, lie in $B(0, R_N)$ and are $2r_N$-separated by (1.20). Hence

$$N_{r_N}(F_p \cap B(0, R_N)) \ge N + 1.$$

Since R_N/r_N is comparable to N, no exponent below one can satisfy a uniform Assouad upper estimate. Therefore $\dim_A F_p = 1$. □

I.5.2 Dimension spectra

It is convenient in this section to use the equivalent smaller-scale variable ρ. Thus the two scales are ρ and ρ^θ, with $0 < \rho < 1$.

Theorem — T-I.5.06 — Spectrum hierarchy [F]

For every $0 < \theta < 1$,

$$\dim_L F \le \dim_L^\theta F \le \underline{\dim}_B F \le \overline{\dim}_B F \le \dim_A^\theta F \le \dim_A F. \tag{2.1}$$

Proof

For a singleton all quantities are zero by convention. Assume that $|F| > 0$. The first and last inequalities follow by restricting the pairs of scales in D-I.5.02 to $R = \rho^\theta$ and $r = \rho$. The middle box inequality is definitional. We prove the other two by iteration.

Let $a > \dim_A^\theta F$. For all sufficiently small u and every $x \in F$,

$$N_u(F \cap B(x, u^\theta)) \le C(u^\theta/u)^a. \tag{2.2}$$

Fix small r and form the increasing sequence

$$r_0 = r, \qquad r_j = r^{\theta^j}.$$

Stop at the first n for which $r_n \ge r_*$, where $0 < r_* < \min\{1, |F|\}$ is fixed. Then r_{n-1} is bounded below by the positive constant $r_*^{1/\theta}$, so $N_{r_{n-1}}(F) \le C_*$ uniformly. Starting with an r_{n-1}-cover of F, cover the intersection of F with each r_j-ball by r_{j-1}-balls. This is exactly (2.2), since $r_j = r_{j-1}^\theta$. Induction gives

$$N_r(F) \le C_* C^n (r_{n-1}/r)^a. \tag{2.3}$$

Since $n = O(\log\log(1/r))$,

$$\frac{n}{-\log r} \longrightarrow 0.$$

Taking the upper logarithmic limit in (2.3) gives $\overline{\dim}_{\mathrm{B}} F \leq a$, and then a decreases to $\dim_{\mathrm{A}}^{\theta} F$.

If the lower spectrum is zero, its comparison with lower box dimension is immediate. Otherwise let $0 < s < \dim_{\mathrm{L}}^{\theta} F$. There is $c > 0$ such that

$$N_{R^{1/\theta}}(F \cap B(x,R)) \geq c(R/R^{1/\theta})^s \tag{2.4}$$

for all sufficiently small R. We give the separated-tree iteration explicitly. Put $\delta_0 = 2r$ and define upward

$$\delta_j = 4\delta_{j-1}^{\theta} \qquad (j \geq 1). \tag{2.5}$$

Stop at the first n for which δ_n reaches a fixed sufficiently small reference scale δ_*. The recurrence has the exact logarithmic form

$$-\log \delta_j = \theta^j \left(-\log(2r) + \frac{\log 4}{1-\theta} \right) - \frac{\log 4}{1-\theta}.$$

It follows that $n = O(\log\log(1/r))$ and $\delta_* \leq \delta_n < 4\delta_*^{\theta}$. Choose δ_* so small that all parent radii used below are within the range of (2.4). Begin with one parent centre in F. If the level-j parent centres are δ_j-separated, apply (2.4) in $B(x, \delta_j/4)$. Its prescribed inner scale is

$$(\delta_j/4)^{1/\theta} = \delta_{j-1}.$$

By $N_\delta \leq M_\delta$, choose at least

$$c\left(\frac{\delta_j}{4\delta_{j-1}}\right)^s$$

children which are δ_{j-1}-separated and lie in $B(x, \delta_j/4)$. Children of distinct parents are also δ_{j-1}-separated because their parents are δ_j-separated and, for all scales used, $\delta_{j-1} < \delta_j/2$. Iterating from level n down to level zero produces a δ_0-separated subset of F with at least

$$c^n 4^{-sn} (\delta_n/\delta_0)^s \tag{2.6}$$

points. Since $\delta_0 = 2r$, L-I.2.01 gives

$$N_r(F) \geq M_{2r}(F) \geq c^n 4^{-sn} (\delta_n/(2r))^s. \tag{2.7}$$

The factor $c^n 4^{-sn}$ has logarithm $O(n) = o(-\log r)$, while δ_n is bounded above and below by positive constants fixed independently of r. Taking the lower logarithmic limit in (2.7) gives

$$\underline{\dim}_{\mathrm{B}} F \geq s.$$

Letting s increase to the lower spectrum completes (2.1). $\square$

Theorem — T-I.5.07 — Two-parameter spectrum comparison [F]

Let $0 < \theta_1 < \theta_2 < 1$. Then

$$\frac{1-\theta_2}{1-\theta_1} \dim_{\mathrm{A}}^{\theta_2} F \le \dim_{\mathrm{A}}^{\theta_1} F \tag{2.8}$$

and

$$\dim_{\mathrm{A}}^{\theta_1} F \le \frac{1-\theta_2}{1-\theta_1} \dim_{\mathrm{A}}^{\theta_2} F + \frac{\theta_2-\theta_1}{1-\theta_1} \dim_{\mathrm{A}}^{\theta_1/\theta_2} F. \tag{2.9}$$

For the lower spectra,

$$\frac{1-\theta_2}{1-\theta_1} \dim_{\mathrm{L}}^{\theta_2} F \le \dim_{\mathrm{L}}^{\theta_1} F \tag{2.10}$$

and

$$\dim_{\mathrm{L}}^{\theta_1} F \le \frac{\theta_1(1-\theta_2)}{\theta_2(1-\theta_1)} \dim_{\mathrm{L}}^{\theta_2} F + \frac{\theta_2-\theta_1}{\theta_2(1-\theta_1)} \dim_{\mathrm{A}}^{\theta_1/\theta_2} F. \tag{2.11}$$

Proof

For a singleton the four inequalities are immediate. Assume that $|F| > 0$. Put $H = D_d(2)$, as in L-I.5.01. We will use the precise subset comparison $N_r(A) \le HN_r(B)$ for $A \subset B$, rather than exact monotonicity. We also use the following two-stage version. If a target piece is covered by M balls centred in F, and each corresponding local piece of F has a centred r-cover with at most P balls, the combined balls cover the target but may have centres outside it. Discard balls missing the target and recenter the others, giving a centred $2r$-cover. Equation (1.4b) then bounds its centred r-covering number by HMP.

Choose

$$s < \dim_{\mathrm{A}}^{\theta_2} F, \qquad t > \dim_{\mathrm{A}}^{\theta_2} F, \qquad u > \dim_{\mathrm{A}}^{\theta_1/\theta_2} F.$$

For arbitrarily small ρ, the failure of an exponent-s upper bound gives

$$\sup_{x\in F} N_\rho(B(x,\rho^{\theta_2}) \cap F) \ge (\rho^{\theta_2}/\rho)^s.$$

Because $\rho^{\theta_2} < \rho^{\theta_1}$, the subset comparison gives at least $\mathscr{H}^{-1}$ times this lower bound for the larger outer ball ρ^{θ_1}. Since

$$(\rho^{\theta_2}/\rho)^s = (\rho^{\theta_1}/\rho)^{s(1-\theta_2)/(1-\theta_1)},$$

this fixed factor does not affect the exponent, and we obtain (2.8) after sending s to the spectrum. If that spectrum is zero, (2.8) follows directly from nonnegativity.

For every small ρ, first cover $B(x,\rho^{\theta_1}) \cap F$ by balls of radius ρ^{θ_2} and then cover each resulting piece at radius ρ. The first stage uses parameter θ_1/θ_2, while the second uses parameter θ_2. Apply the two-stage centred-cover comparison above; its factor H is included in C. Therefore

$$\begin{aligned}\sup_x N_\rho(B(x,\rho^{\theta_1})\cap F) &\le C(\rho^{\theta_1}/\rho^{\theta_2})^u(\rho^{\theta_2}/\rho)^t \\ &= C(\rho^{\theta_1}/\rho)^{u(\theta_2-\theta_1)/(1-\theta_1)+t(1-\theta_2)/(1-\theta_1)}.\end{aligned}$$

Letting t, u decrease to their dimensions proves (2.9).

For (2.10), take $0 \le s < \dim_{\mathrm{L}}^{\theta_2} F$ when this dimension is positive. The lower bound at the smaller outer ball ρ^{θ_2} gives a lower bound divided by H inside the larger outer ball ρ^{θ_1}. The same exponent conversion gives (2.10). A zero lower spectrum makes the claimed inequality immediate.

Finally choose

$$t > \dim_{\mathrm{L}}^{\theta_2} F, \qquad u > \dim_{\mathrm{A}}^{\theta_1/\theta_2} F.$$

By failure of the lower-spectrum bound at exponent t with proposed constant one, there are arbitrarily small ρ and a centre x for which the number of ρ^{θ_1/θ_2}-balls needed to cover $B(x,\rho^{\theta_1})\cap F$ is at most

$$(\rho^{\theta_1}/\rho^{\theta_1/\theta_2})^t.$$

Cover each intermediate ball by ρ-balls using the Assouad spectrum at parameter θ_1/θ_2. Recenter the resulting cover and apply (1.4b) as above, absorbing H into C. This gives

$$\begin{aligned}N_\rho(B(x,\rho^{\theta_1})\cap F) &\le C(\rho^{\theta_1}/\rho^{\theta_1/\theta_2})^t(\rho^{\theta_1/\theta_2}/\rho)^u \\ &= C(\rho^{\theta_1}/\rho)^v,\end{aligned}$$

where

$$v = \frac{\theta_1(1-\theta_2)}{\theta_2(1-\theta_1)}t + \frac{\theta_2-\theta_1}{\theta_2(1-\theta_1)}u.$$

Such small scales obstruct every lower-spectrum exponent greater than v. Let t, u decrease to the indicated dimensions to obtain (2.11). ☐

Corollary — C-I.5.08 — Continuity of the Assouad scale profile [F]

The functions

$$\theta \mapsto \dim_{\mathrm{A}}^{\theta} F, \qquad \theta \mapsto \dim_{\mathrm{L}}^{\theta} F$$

are continuous on $(0,1)$.

Proof

Fix a closed interval $I \subset (0,1)$. By T-I.5.03 and T-I.5.06 every dimension appearing in T-I.5.07 lies in $[0,d]$. For $\theta_1 < \theta_2$ in I, (2.8)–(2.9) imply

$$|\dim_{\mathrm{A}}^{\theta_1} F - \dim_{\mathrm{A}}^{\theta_2} F| \le C_{I,d}|\theta_2-\theta_1|.$$

Indeed, subtract the term with coefficient $(1-\theta_2)/(1-\theta_1)$ and bound the remaining spectrum by d. Equations (2.10)–(2.11) give the same estimate for the lower spectrum, with a possibly different $C_{I,d}$. Thus both functions are Lipschitz on every closed interval compactly contained in $(0,1)$, hence continuous there. Since every point of $(0,1)$ lies in such an interval, the proof is complete. □

I.5.3 Hausdorff hyperspaces and weak tangents

Lemma — L-I.5.09 — Compactness of the Hausdorff hyperspace [F]

The metric space $(K(W), d_H)$ is compact.

Proof

First prove total boundedness. Given $\varepsilon > 0$, choose a finite set $Z = \{z_1, \dots, z_N\}$ in W such that every point of W has distance less than $\varepsilon/2$ from some member of Z. For $A \in K(W)$, let

$$Z_A = \{z_j : \operatorname{dist}(z_j, A) < \varepsilon\}.$$

This set is nonempty and $d_H(A, Z_A) < 2\varepsilon$. There are only finitely many nonempty subsets of Z, so these subsets form a finite 2ε-net in $K(W)$.

It remains to prove completeness. Let (A_n) be d_H-Cauchy. Select a subsequence (A_{n_j}) such that

$$d_H(A_{n_j}, A_{n_{j+1}}) < 2^{-j}. \tag{3.1}$$

Choose $a_1 \in A_{n_1}$ and inductively $a_{j+1} \in A_{n_{j+1}}$ with $|a_{j+1} - a_j| < 2^{-j}$. Then (a_j) converges in compact W, so the set

$$A = \left\{ a \in W \;:\; \begin{array}{l} \text{there are } j_k \to \infty \text{ and } x_k \in A_{n_{j_k}} \\ \text{with } x_k \to a \end{array} \right\} \tag{3.2}$$

is nonempty. It is closed in W, since it can equivalently be written as

$$A = \bigcap_{j \ge 1} \overline{\bigcup_{k \ge j} A_{n_k}}.$$

Indeed, a subsequential limit lies in every displayed closure; conversely, membership in every closure lets us choose strictly increasing indices and points approaching the given point. Hence A is compact.

For any fixed j, repeated use of (3.1) shows that each point of A_{n_j} lies within $\sum_{i \ge j} 2^{-i}$ of a point of A; the same chaining argument starting from far later sets shows that every point of A lies within that distance of A_{n_j}. Consequently

$$d_H(A_{n_j}, A) \le \sum_{i \ge j} 2^{-i} \longrightarrow 0.$$

The original Cauchy sequence has the same limit. Thus $K(W)$ is complete and totally bounded, hence compact. □

Theorem — T-I.5.10 — Weak tangents do not raise Assouad dimension [F]

If E is a windowed weak tangent of F, then

$$\dim_{\mathrm{A}} E \leq \dim_{\mathrm{A}} F. \tag{3.3}$$

Proof

If F is a singleton, every nonempty window intersection is a singleton. Its Hausdorff limit is also a singleton, so both Assouad dimensions are zero. Assume now that $|F| > 0$ and let $s > \dim_{\mathrm{A}} F$. Since the similarity ratios tend to infinity, discard finitely many indices so that every transformed diameter exceeds $2\,\mathrm{diam}(W)$. A similarity preserves ratios of distances, so one constant C then satisfies

$$N_r(T_k F \cap B(x, R)) \leq C(R/r)^s \tag{3.4}$$

for every k, every $x \in T_k F$, and all admissible scales; scales exceeding the transformed diameter are absorbed into C.

If E is a singleton, its Assouad dimension is zero and there is nothing more to prove. Otherwise fix $y \in E$ and $0 < r < R \leq \mathrm{diam}\, E$. Choose k so large that

$$d_H(E, T_k F \cap W) < r/8.$$

There is $x \in T_k F \cap W$ with $|x - y| < r/8$. Cover $T_k F \cap B(x, 2R)$ by $r/8$-balls with centres a_i in that piece. For every $p \in E \cap B(y, R)$, Hausdorff approximation gives $z \in T_k F \cap W$ with $|p - z| < r/8$. Then $|z - x| < R + r/4 < 2R$, so z lies in one of the covering balls and $|p - a_i| < r/4$. Retain precisely those indices whose $r/4$-ball meets $E \cap B(y, R)$, and choose a new centre e_i in that intersection. The triangle inequality gives $|p - e_i| < r/2$ whenever $|p - a_i| < r/4$. Thus the resulting r-balls are centred in the target local piece, as required by D-I.2.02. Since $2R \leq 2\,\mathrm{diam}(W)$ is below the transformed diameter, (3.4) applies and gives

$$N_r(E \cap B(y, R)) \leq N_{r/8}(T_k F \cap B(x, 2R)) \leq C16^s(R/r)^s.$$

The constant is independent of y, r, R. Let s decrease to $\dim_{\mathrm{A}} F$. □

Theorem — T-I.5.11 — Interior weak tangents dominate lower dimension [F]

If E is a windowed weak tangent of F, then

$$\dim_{\mathrm{L}} F \leq \dim_{\mathrm{H}} E. \tag{3.5}$$

Proof

If $\dim_{\mathrm{L}} F = 0$, the conclusion follows from the nonnegativity of Hausdorff dimension. Assume $\dim_{\mathrm{L}} F > 0$. Choose $z \in E \cap W^\circ$ and $R_0 > 0$ so small that

$$B(z, 4R_0) \subset W^\circ. \tag{3.6}$$

Let $0 \leq s < t < \dim_{\mathrm{L}} F$. For all large k, choose $z_k \in T_k F \cap W$ with $|z_k - z| < R_0$. Then $B(z_k, 2R_0) \subset W$, and the lower estimate with exponent t and its constant are preserved by T_k.

Apply L-I.5.02 to $T_k F$ at the root ball $B(z_k, R_0)$. We obtain a probability μ_k carried by

$$T_k F \cap B(z_k, R_0) \subset T_k F \cap W$$

and a constant C, independent of k, such that

$$\mu_k(B(y, r)) \leq C(r/R_0)^s \qquad (0 < r < R_0). \tag{3.7}$$

By L-I.3.02, pass to a weakly convergent subsequence $\mu_k \to \mu$ on W. To check its carrier explicitly, put $\eta_k = d_H(T_k F \cap W, E) \to 0$. For every $\varepsilon > 0$, the continuous function $f_\varepsilon(w) = \min\{1, \mathrm{dist}(w, E)/\varepsilon\}$ satisfies

$$\int f_\varepsilon \, d\mu = \lim_k \int f_\varepsilon \, d\mu_k \leq \lim_k \eta_k/\varepsilon = 0.$$

It equals one when $\mathrm{dist}(w, E) \geq \varepsilon$. Taking $\varepsilon = 1/j$, $j \geq 1$, shows that $\mu(W \setminus E) = 0$.

For an open ball, the Portmanteau inequality and (3.7) give

$$\mu(B(y, r)) \leq \liminf_k \mu_k(B(y, r)) \leq C(r/R_0)^s.$$

The same estimate for closed balls follows by decreasing slightly larger open radii. The mass-distribution principle yields $\dim_{\mathrm{H}} E \geq s$. Let s increase to $\dim_{\mathrm{L}} F$. The use of (3.6) is exactly where the interior-window hypothesis is needed. □

I.5.4 Realization and strict star dimension

Theorem — T-I.5.12 — Tangent realization at Assouad dimension [F]

Let $\dim_{\mathrm{A}} F = s$. There is a strict microset $E \subset W$, obtained by homotheties with ratios tending to infinity, such that

$$\mathscr{H}^s(E) > 0, \qquad \dim_{\mathrm{H}} E = \dim_{\mathrm{A}} E = s. \tag{4.1}$$

Proof

Assume first that $s > 0$. Let $Q(m)$ be the half-open dyadic cubes of side 2^{-m}; these form a genuine partition, and closures are taken only when a compact geometric cube is needed. By the dyadic characterization in L-I.5.01, for every k there are integers $m_k < n_k$ and $Q_k \in Q(m_k)$ such that

$$M_{n_k}(Q_k; F) \geq 2^{(n_k - m_k)(s - 1/k)}, \qquad n_k - m_k \longrightarrow \infty. \tag{4.2}$$

Indeed, otherwise an exponent below s would have a uniform dyadic upper bound. Choose one point of F from every occupied level-n_k subcube of Q_k, and give these points equal mass. The resulting probability μ_k satisfies

$$\mu_k(Q') = M_{n_k}(Q_k; F)^{-1} \tag{4.3}$$

for each occupied level-n_k subcube $Q' \subset Q_k$.

We need a finite-depth regularity claim. If $0 < t < s$ and ℓ is a positive integer, then for some sufficiently large k there is a dyadic cube $Q = Q(\ell, t) \subset Q_k$, of level $L \geq m_k + \ell$, with positive mass, such that

$$\frac{\mu_k(Q')}{\mu_k(Q)} \leq \left(\frac{|Q'|}{|Q|}\right)^t \tag{4.4}$$

for every dyadic $Q' \subset Q$ whose level lies between L and $L + \ell$. Here $|Q|$ denotes side length.

Suppose the claim failed for fixed t, ℓ. Take k large enough that $n_k - m_k > 2\ell$. Among the level-$m_k + \ell$ descendants of Q_k, choose one, call it Q^0, with

$$\mu_k(Q^0) \geq 2^{-\ell d}. \tag{4.5}$$

Failure of (4.4) gives a strict descendant Q^1 at most ℓ levels below Q^0 for which the reverse strict inequality holds. Continue choosing bad descendants, and let Q^q be the last chosen cube whose level does not exceed n_k; the next bad descendant crosses level n_k. Successive level jumps are at most ℓ, so its level l_q satisfies

$$n_k - \ell \leq l_q \leq n_k. \tag{4.6}$$

Multiplying the bad ratios and using (4.5) gives

$$\mu_k(Q^q) > 2^{-(n_k - m_k)t} 2^{-\ell(d-t)}. \tag{4.7}$$

Choose $t < t' < s$. For large k, (4.2) implies

$$M_{n_k}(Q_k; F) \geq 2^{(n_k - m_k)t'}.$$

The cube Q^q contains at most 2^{dell} level-n_k cubes by (4.6), so (4.3) gives

$$\mu_k(Q^q) \leq 2^{d\ell} 2^{-(n_k - m_k)t'}. \tag{4.8}$$

Equations (4.7)–(4.8) contradict $n_k - m_k \to \infty$. The claim is proved.

Choose positive $t_\ell < s$ with $t_\ell \to s$, and apply the claim. The admissible outer-scale cutoff in D-I.5.01 makes the levels m_k bounded below; since $L \geq m_k + \ell$, the selected levels tend to infinity. Let T_ℓ be the homothety carrying the closure of $Q(\ell, t_\ell)$ onto the fixed interior dyadic cube

$$J = [1/4, 1/2]^d \subset W^\circ.$$

Since the level L tends to infinity, the ratios of T_ℓ tend to infinity. By L-I.5.09, pass to a subsequence such that

$$T_\ell(F) \cap W \longrightarrow E \tag{4.9}$$

in d_H. Normalize μ_k on $Q(\ell, t_\ell)$ and push it forward by T_ℓ; call the resulting probability ν_ℓ. It is carried by $T_\ell(F) \cap J$. Pass again to a subsequence so that ν_ℓ converges weakly to a probability ν, using L-I.3.02. The continuous-distance argument in T-I.5.11 shows that ν is carried by E; since every ν_ℓ is carried by the fixed closed cube J, the same argument gives full mass to J. Thus ν is carried by $E \cap J$, so $E \cap W^\circ$ is nonempty and (4.9) makes E a strict microset.

Fix a sufficiently small radius r and choose $n \geq 2$ with

$$2^{-n-1} < r \leq 2^{-n}.$$

A ball of radius r meets at most C_d level-n dyadic cubes. The map T_ℓ sends a level-L cube onto the level-2 cube J; hence, for fixed n and all $\ell \geq n-2$, the relevant inverse dyadic cubes have level $L+n-2$, between L and $L+\ell$. Equation (4.4) therefore gives

$$\nu_\ell(B(y,r)) \leq C_d 4^{t_\ell} 2^{-nt_\ell}. \tag{4.10}$$

Use a slightly larger open ball and the Portmanteau inequality, then let $\ell \to \infty$. Since $t_\ell \to s$, (4.10) yields

$$\nu(B(y,r)) \leq C'_d r^s. \tag{4.11}$$

The mass-distribution principle gives $\mathscr{H}^s(E) > 0$ and $\dim_{\mathrm{H}} E \geq s$. T-I.5.10 gives $\dim_{\mathrm{A}} E \leq \dim_{\mathrm{A}} F = s$, while T-I.5.03 gives $\dim_{\mathrm{H}} E \leq \dim_{\mathrm{A}} E$. All three dimensions are therefore equal to s.

If $s = 0$, choose $x \in F$, a point $z \in W^\circ$, and homotheties

$$T_k(y) = k(y-x) + z.$$

Every $T_k F \cap W$ contains z; L-I.5.09 gives a Hausdorff-convergent subsequence with nonempty limit E meeting the interior. Theorem I.5.10 gives $\dim_{\mathrm{A}} E = 0$, and $\mathscr{H}^0(E) > 0$ because E is nonempty. This completes all cases. □

Corollary — C-I.5.13 — Star dimension equals Assouad dimension [F]

For every nonempty compact $F \subset W$,

$$\dim_* F = \dim_{\mathrm{A}} F, \tag{4.12}$$

and a strict microset attains the supremum.

Proof

Every strict microset is a windowed weak tangent by D-I.5.07, so T-I.5.10 and T-I.5.03 give

$$\dim_{\mathrm{H}} E \leq \dim_{\mathrm{A}} E \leq \dim_{\mathrm{A}} F$$

for every strict microset E. Hence $\dim_* F \le \dim_{\mathrm{A}} F$. The strict microset constructed in T-I.5.12 has Hausdorff dimension exactly $\dim_{\mathrm{A}} F$, proving the reverse inequality and attainment. □

I.5.5 The quarantined gallery theorem

Guided result — G-I.5.14 — Furstenberg gallery characterization [G]

For every nonempty compact $F \subset W$, the family G_F from D-I.5.08 is a gallery. More generally, let $G \subset K(W)$ be a gallery. Then there exists $A \in G$ such that

$$\dim_{\mathrm{H}} A = \Delta(\mathscr{G}). \tag{5.1}$$

Furthermore,

$$\sup_{X\in\mathscr{G}} \dim_{\mathrm{H}} X = \sup_{X\in\mathscr{G}} \dim_{\mathrm{A}} X = \Delta(\mathscr{G}), \tag{5.2}$$

and both suprema are attained. If G_F is the Furstenberg microset gallery of a nonempty compact $F \subset W$, then

$$\dim_{\mathrm{A}} F = \sup_{X\in\mathscr{G}_F} \dim_{\mathrm{H}} X = \Delta(\mathscr{G}_F). \tag{5.3}$$

References. See Furstenberg's original gallery theorem [Fur08, Definitions 5.1–5.2, Proposition 5.2, and Theorem 5.1, pp. 419–421] and Fraser's modern formulation [Fra20, Theorem 2.3.1, pp. 16–17].

Guided architecture and exact omitted block

Put

$$a_n = \sup_{X\in\mathscr{G}} Q_n(X). \tag{5.4}$$

If a level-m dyadic cube meets $X \in G$, magnify the compact closure of that piece to W. The result is a Furstenberg miniset of X, hence belongs to G. Counting its level-n cubes shows, with at most a fixed boundary factor C_d, that

$$a_{m+n} \le C_d a_m a_n. \tag{5.5}$$

For completeness, put $b_n = \log_2(C_d a_n)$ for $n \ge 1$ and $b_0 = 0$. Then (5.5) gives $b_{m+n} \le b_m + b_n$. For fixed $k \ge 1$, write $n = qk + r$, $0 \le r < k$. Iteration gives $b_n \le qb_k + b_r$, so $\limsup_n b_n/n \le b_k/k$. Taking the infimum over k, and using $b_n/n \ge \inf_k b_k/k$, proves existence of the limit. Since $\log_2 C_d/n$ tends to zero, the limsup in D-I.5.08 is this exponential growth rate of a_n.

The same magnification argument applied to a fixed $X \in G$ and a level-m cube Q, with $m \ge 0$, gives

$$M_{m+n}(Q;X) \le C_d a_n. \tag{5.6}$$

By L-I.5.01 and the definition of the limsup, for every $\varepsilon > 0$ the right side is at most $C_\varepsilon 2^{n(\Delta(G)+\varepsilon)}$. Hence

$$\dim_{\mathrm{A}} X \le \Delta(\mathscr{G}) \qquad (X \in \mathscr{G}). \tag{5.7}$$

The finitely many possible negative parent levels in L-I.5.01 cause no problem: $\operatorname{diam} X \le \sqrt{d}$ bounds them below, and for positive fine levels their descendants are bounded by the global count a_n. At nonpositive fine levels there are only boundedly many descendants. These give the same exponent with an enlarged constant. Singletons satisfy (5.7) directly.

The measure-selection input uses the source's endpoint convention, which we specify before stating it. For $n \ge 1$ and $0 \le j < 2^n$, let

$$J_{n,j} = \begin{cases} [j2^{-n}, (j+1)2^{-n}), & 0 \le j < 2^n - 1, \\ [1-2^{-n}, 1], & j = 2^n - 1. \end{cases}$$

Let $\mathscr{P}_n$ be the partition of W into products of d such intervals, and let $P_n(x)$ be its unique cell containing x. Write

$$\widetilde{Q}_n(X) = \#\{P \in \mathscr{P}_n \,:\, P \cap X \ne \varnothing\}.$$

Each source cell is a union of at most 2^d nonempty intersections of global half-open grid cells with W; each global cell meeting W belongs to exactly one source cell. Consequently

$$\widetilde{Q}_n(X) \le Q_n(X) \le 2^d \widetilde{Q}_n(X) \qquad (X \subset W).$$

The two gallery suprema therefore have the same logarithmic growth exponent. In particular, Furstenberg's Δ with base $p = 2$ is exactly the value in D-I.5.08, although the finite partitions are not identical.

The omitted measure-selection block is the following existence assertion.

FUR08 gallery-measure selection block. For every gallery G, there are $A \in G$ and a Borel probability μ carried by A such that, for μ-almost every x,

$$\lim_{n\to\infty} -\frac{\log_2 \mu(P_n(x))}{n} = \Delta(\mathscr{G}). \tag{5.8}$$

The exact source is Furstenberg, Theorem 5.1 on printed p. 421, together with the cell-mass limit definition on printed p. 411, using $p = 2$. The source proves the assertion via Theorem 4.2 and Proposition 5.2. Its technical selection argument is not reproduced here.

Orientation, not a proof of the selection block. Giving equal mass to the occupied cells at one fine scale captures their counting exponent. Furstenberg averages successive magnifications of these measures and selects a limiting law that retains this information at arbitrarily fine scales. The difficult step is the selection leading to the almost-sure limit (5.8); that entire step is the explicitly imported assertion above.

We prove every deduction surrounding that block. If $\Delta(G) = 0$, (5.7) and nonnegativity give $\dim_{\mathrm{H}} A = \dim_{\mathrm{A}} A = 0$ for any $A \in G$, proving (5.1)–(5.2) and attainment directly. Suppose henceforth that $\Delta(G) > 0$. Fix $0 < q < \Delta(G)$ and set

$$E_N = \{x \in A : \mu(P_n(x)) \le 2^{-nq} \text{ for every } n \ge N\}. \tag{5.9}$$

Equation (5.8) implies $\mu(\cup_N E_N) = 1$, so some E_N has positive measure. Normalize the restriction of μ to this E_N. A ball of radius r, where $2^{-n-1} < r \le 2^{-n}$ and $n \ge N$, meets at most 4^d cells of P_n. Each cell meeting E_N has μ-mass at most 2^{-nq}, by choosing a point of that intersection in (5.9). The sets E_N are Borel, since each cell-mass function is constant on a finite Borel partition. Therefore the normalized restriction ν satisfies

$$\nu(B(y, r)) \le \frac{4^d}{\mu(E_N)} 2^{-nq} \le \frac{2^q 4^d}{\mu(E_N)} r^q. \tag{5.10}$$

Increase the constant to handle larger radii, using $\nu(B) \le 1$. The mass-distribution principle gives $\dim_{\mathrm{H}} A \ge q$. Letting q increase to $\Delta(G)$ and combining with (5.7) yields

$$\dim_{\mathrm{H}} A = \dim_{\mathrm{A}} A = \Delta(\mathscr{G}).$$

This proves (5.1), (5.2), and attainment once the selection block is granted.

For a compact F, Furstenberg's Definitions 1.2–1.5 and the following remark state that the family G_F is a gallery. This is the separate externally cited gallery-closure assertion. Also $F \in G_F$. Every miniset used to approximate a member of G_F is contained in a homothetic copy of F. For every $s > \dim_{\mathrm{A}} F$, the recentering argument in T-I.5.04 gives all these minisets the same exponent-s upper constant. Extend their estimates to larger outer radii by covering the whole miniset when the radius exceeds its diameter; if the inner radius also exceeds its diameter, one ball suffices. The resulting constant is still uniform. The covering argument in T-I.5.10 now passes the estimate to their Hausdorff limit: that argument uses only a uniform upper estimate and Hausdorff convergence, not diverging magnification ratios or an interior hit. Thus it applies to these broader microsets without changing the strict tangent definition. Consequently

$$\dim_{\mathrm{A}} F \le \sup_{X \in \mathscr{G}_F} \dim_{\mathrm{A}} X \le \dim_{\mathrm{A}} F. \tag{5.11}$$

Applying (5.2) to G_F proves (5.3).

The source for the gallery assertion is Furstenberg, "Ergodic fractal measures and dimension conservation," Definitions 1.2–1.5 and the following remark on printed p. 407. The omitted block and its conclusion are Theorem 4.2 on printed p. 418, Proposition 5.2 on printed p. 420, and Theorem 5.1 on printed p. 421. Fraser, *Assouad Dimension and Fractal Geometry*, Theorem 2.3.1 on printed p. 17 gives the exact dyadic-notation concordance. Equations (4.1)–(4.12) were established independently and do not consume this G node.

I.5.6 Intuition and strategy

The following discussion explains the geometric ideas behind the proofs.

I.5.6.1 Global averages versus the worst local scene

Box dimension asks how many small balls cover the whole set. Assouad dimension allows an adversary to choose both the location and the pair of scales. It therefore records the thickest local scene that appears anywhere and at any magnification. Lower dimension reverses the covering inequality and records the thinnest scene.

The two extreme dimensions respond differently to set inclusion. Adding points can only make the thickest scene at least as complicated, so Assouad dimension is monotone. Adding one isolated point can create a permanently thin scene, so lower dimension is not monotone. Modified lower dimension restores monotonicity by allowing the thin exceptional part to be discarded through passage to a subset.

I.5.6.2 Why lower homogeneity produces a Frostman measure

A lower covering estimate says that every ball has many well-separated descendants at a sufficiently smaller scale. Choosing a large fixed scale ratio turns those descendants into a tree with a uniform branching number. Putting equal conditional mass on the branches makes the mass of a level-n cylinder decay faster than the sth power of its diameter. Euclidean packing controls how many same-level cylinders a ball can meet.

This gives a Frostman measure without importing I.3's Frostman existence theorem. It is the mechanism behind both

$$\dim_{\mathrm{L}} F \le \dim_{\mathrm{H}} F$$

for compact sets and the lower-dimension inequality for interior weak tangents.

I.5.6.3 The spectrum records which scale gaps are responsible

The Assouad dimension permits every pair $r < R$. The spectrum forces

$$r = R^{1/\theta}.$$

Small θ produces a large logarithmic gap, while θ close to one forces nearby logarithmic scales. Iterating the prescribed relation connects one local estimate to a global box estimate. Only $O(\log\log(1/r))$ stages are needed, so fixed losses per stage disappear after division by $-\log r$.

The two-parameter comparison theorem inserts an intermediate scale. For upper estimates one multiplies two covers. For lower estimates one selects a ball which realizes a small lower covering number at the first stage, then uses an Assouad upper estimate to cover each intermediate piece. The coefficients in T-I.5.07 are exactly the proportions of the total logarithmic scale gap occupied by those two stages.

I.5.6.4 Tangents turn worst finite scales into limiting sets

If a compact set has large Assouad dimension, there are dyadic cubes with unusually many occupied descendants. A direct rescaling of those cubes need not have a useful limiting measure: mass might concentrate at a few intermediate levels. The finite-depth selection claim in T-I.5.12 finds a subcube on which no such concentration occurs for a growing number of levels.

After rescaling that subcube into a fixed interior window, compactness gives a Hausdorff limit and weak compactness gives a limiting probability. The finite-depth mass bounds become a genuine Frostman estimate in the limit. This is why one realizes not merely a tangent of large box dimension but a tangent with positive Hausdorff measure in the critical dimension.

I.5.6.5 Why the interior-window condition matters

Window intersection is not innocent for lower bounds. A rescaled set may touch the observation cube only at one boundary point even when the original set is an interval. Such a boundary remnant has dimension zero and cannot represent the original set's thinnest genuine magnification.

Requiring the tangent to meet the interior provides a fixed smaller ball which remains inside the window. The lower-homogeneity tree can then be built in that ball without being cut off by the observation boundary.

I.5.6.6 Two microset conventions must remain separate

The series' strict microsets are exact intersections

$$(\lambda_k F + t_k) \cap [0, 1]^d$$

with λ_k tending to infinity. Furstenberg's miniset convention allows an arbitrary compact subset of a magnified copy. The latter is designed to form a closed hereditary gallery and is broader.

The strict star-dimension identity is proved before the gallery theorem: the tangent-realization construction supplies a strict microset of full Assouad dimension, and the weak-tangent upper bound supplies the reverse inequality. The gallery theorem is therefore valuable extra structure, not a hidden premise.

I.5.6.7 The role of the cited gallery theorem

For a gallery, the worst dyadic occupancy exponent is attained by an actual member whose Hausdorff dimension equals that exponent. This is substantially stronger than compactness alone; Hausdorff dimension is not continuous under Hausdorff convergence. Furstenberg's theorem supplies the ergodic selection mechanism needed to prevent dimension from disappearing in the limit.

Because that mechanism is not reproduced here, G-I.5.14 has a guided architecture with an exact cited selection input, not a full internal proof. Nothing after it in the chapter uses it, and the F export remains logically closed.

I.5.7 Exercises

Every exercise uses only D/R-I.1–I.5 and A-RA. G-I.5.14 is not a permitted premise.

I.5.7.1 Exercise 1 — Intervals and cubes

Compute the Assouad dimension, lower dimension, modified lower dimension, both box dimensions, and both spectra of $[0,1]^d$.

I.5.7.2 Exercise 2 — Ahlfors-regular sets

Let μ be a probability carried by compact $F \subset \mathbb{R}^d$. Assume that for some $\alpha \geq 0$, $r_0 > 0$, and $0 < c \leq C < \infty$,

$$cr^\alpha \leq \mu(B(x,r)) \leq Cr^\alpha \qquad (x \in F, 0 < r < r_0).$$

This says that μ is α-**Ahlfors regular** on F; a compact set carrying such a measure is called α-**Ahlfors regular**.

Prove that

$$\dim_{\mathrm{L}} F = \dim_{\mathrm{H}} F = \dim_{\mathrm{P}} F = \underline{\dim}_{\mathrm{B}} F = \overline{\dim}_{\mathrm{B}} F = \dim_{\mathrm{A}} F = \alpha,$$

and that both spectra are constantly equal to α.

I.5.7.3 Exercise 3 — Finite sets and their strict microsets

Let F contain exactly $m \geq 1$ points. Prove that every dimension defined in I.5 is zero and that every strict microset has at most m points. Deduce directly, without C-I.5.13, that $\dim_* F = 0$.

I.5.7.4 Exercise 4 — An isolated point destroys lower dimension

Let

$$F = [0,1] \cup \{2\}.$$

Compute $\dim_{\mathrm{L}} F$, $\dim_{\mathrm{ML}} F$, and $\dim_{\mathrm{A}} F$. Identify precisely which dimension fails monotonicity when $[0,1]$ is enlarged to F.

I.5.7.5 Exercise 5 — Bi-Lipschitz invariance of the spectra

Let $F \subset \mathbb{R}^d$ be nonempty and compact, and let $T : F \to \mathbb{R}^e$ satisfy $K^{-1}|x-y| \leq |T(x) - T(y)| \leq K|x-y|$ for all $x, y \in F$, where $K \geq 1$. Prove, using the recentering and inner-radius comparisons in L-I.5.01, that for every $0 < \theta < 1$,

$$\dim_{\mathrm{A}}^\theta T(F) = \dim_{\mathrm{A}}^\theta F, \qquad \dim_{\mathrm{L}}^\theta T(F) = \dim_{\mathrm{L}}^\theta F.$$

I.5.7.6 Exercise 6 — Products of regular sets

Suppose compact E and F carry Ahlfors-regular probabilities of dimensions α and β. Using the max metric and the product probability, prove that $E \times F$ is $(\alpha + \beta)$-Ahlfors regular. Identify all dimensions and spectra of the product.

I.5.7.7 Exercise 7 — The polynomial sequence has an interval microset

For $p > 0$, let F_p be as in P-I.5.05. Prove that

$$E_N = (N^p F_p) \cap [0, 1] \longrightarrow [0, 1]$$

in the Hausdorff metric. Thus $[0, 1]$ is a strict microset of F_p and the star-dimension conclusion can be seen explicitly in this example.

I.5.7.8 Exercise 8 — A basic Hausdorff limit

Set

$$A_n = \{j2^{-n} \,:\, 0 \leq j \leq 2^n\} \subset [0, 1].$$

Compute an upper bound for $d_H(A_n, [0, 1])$ and prove convergence. Explain why this example does not by itself imply continuity of Hausdorff dimension.

I.5.7.9 Exercise 9 — Why boundary-only tangents are excluded

Let $F = [0, 1]$ and $T_k(x) = kx + 1$. Prove

$$T_k(F) \cap [0, 1] = \{1\}.$$

Show that if the interior condition were removed from D-I.5.06, the conclusion of T-I.5.11 would fail.

I.5.7.10 Exercise 10 — Fixed points of the spectrum inequalities

Assume the Assouad spectrum of F is constantly α. Substitute this function into both sides of the Assouad inequalities in T-I.5.07 and verify them algebraically. Do the same when both the lower and Assouad spectra are constantly α.

I.5.7.11 Exercise 11 — Strict and Furstenberg minisets differ

Let $F = [0, 1]$ and let C be the middle-third Cantor set. Prove that C is a Furstenberg miniset of F, but is not a strict miniset of F.

I.5.7.12 Exercise 12 — Strict microsets are weak tangents

Starting only from D-I.5.06–D-I.5.07, prove that every strict microset is a windowed weak tangent obtained by homotheties. Then use T-I.5.10 to reprove the upper bound $\dim_* F \leq \dim_{\mathrm{A}} F$ without using any gallery result.

Chapter I.6

Density, Porosity, Rectifiability, and Visibility

Dimension alone does not say whether a set resembles a smooth surface or contains holes at many scales. Densities, porosity, and approximate tangent planes address these different geometric questions. We separate the normalization of Hausdorff measure from dimension, distinguish rectifiable from purely unrectifiable sets, and introduce visibility. Several deeper endpoints are explicitly cited rather than proved; their hypotheses and the limits of their use are part of the statements.

The definitions below fix this chapter's conventions. Geometric intuition and proof strategy are discussed separately after the main development; exercises and their solutions provide a second route through the material.

Definitions and notation

Throughout, $d \geq 1$, Euclidean balls are closed, and $\mathscr{L}^d$ denotes Lebesgue measure on $\mathbb{R}^d$. In the rectifiability block, m is an integer with $1 \leq m \leq d$.

Definition — I.6.01: Euclidean Hausdorff normalization and densities

Let

$$\omega_m = \mathscr{L}^m(\overline{B}(0,1)).$$

The **Euclidean-normalized m-dimensional Hausdorff measure** is

$$\mathscr{H}^m = \frac{\omega_m}{2^m}\mathcal{H}^m,$$

where $\mathcal{H}^m$ is the diameter-normalized measure from D-I.1.09–D-I.1.10. The new letter is used only when integer-dimensional density constants matter.
For a locally finite Borel measure μ on $\mathbb{R}^d$, define

$$\Theta^m_*(\mu, x) = \liminf_{r\downarrow 0} \frac{\mu(\overline{B}(x,r))}{\omega_m r^m}, \qquad \Theta^{*m}(\mu, x) = \limsup_{r\downarrow 0} \frac{\mu(\overline{B}(x,r))}{\omega_m r^m}.$$

If the two values agree, their common value is $\Theta^m(\mu, x)$. For a $\mathscr{H}^m$-measurable set E, write

$$\Theta^m_*(E, x) = \Theta^m_*(\mathscr{H}^m \restriction E, x)$$

and analogously for the upper density and density. For a set whose restriction is not locally finite, these notations mean the displayed liminf and limsup of $\mathscr{H}^m(E \cap \overline{B}(x, r))/(\omega_m r^m)$ directly, with values in $[0, +\infty]$; local finiteness is not an implicit hypothesis on the set. The finite-measure hypotheses in the characterization theorems will be stated separately.
For later comparisons in this chapter, we also fix the term **Ahlfors regularity**. Let E be a nonempty compact metric set and let ν be a Borel probability with $\operatorname{supp}(\nu) = E$. For $s \geq 0$, ν is **s-Ahlfors regular on** E if there are $0 < a \leq b < \infty$ and $r_0 > 0$ such that

$$ar^s \leq \nu(\overline{B}(x, r)) \leq br^s \qquad (x \in E,\ 0 < r \leq r_0).$$

The set E is **s-Ahlfors regular** if it supports such a probability. This convention is repeated in D-I.8.06 when the property is used for self-similar measures; no result from that later chapter is needed here.

Definition — I.6.02: Set porosity

For $E \subset \mathbb{R}^d$, $x \in E$, and $r > 0$, the **point-scale porosity** is

$$\operatorname{por}(E, x, r) = \sup \left\{ a \geq 0 \ : \ \begin{array}{l} \text{there is } y \in \mathbb{R}^d \text{ such that} \\ \overline{B}(y, ar) \subset \overline{B}(x, r) \setminus E \end{array} \right\}.$$

The supremum of the empty set is zero. The **lower porosity at** x is

$$\operatorname{por}(E, x) = \liminf_{r \downarrow 0} \operatorname{por}(E, x, r).$$

For $0 < \alpha \leq 1/2$ and $r_0 > 0$, E is **uniformly (α, r_0)-porous** if

$$\operatorname{por}(E, x, r) \geq \alpha \quad (x \in E,\ 0 < r \leq r_0).$$

It is **uniformly porous** if such α, r_0 exist. The adjective "uniformly" is never suppressed in this chapter.

Definition — I.6.03: Mean porosity of sets

Let $0 < \alpha \leq 1/2$ and $0 < \eta \leq 1$. A set $E \subset \mathbb{R}^d$ is **mean (α, η)-porous** if for every $x \in E$,

$$\liminf_{N \to \infty} \frac{1}{N} \# \{0 \leq n < N \ : \ \operatorname{por}(E, x, 2^{-n}) \geq \alpha\} \geq \eta.$$

The dyadic radii are part of the definition. Removing or adding finitely many scales does not change the condition.

Definition — I.6.04: Porosity and mean porosity of measures

Let μ be a Borel probability on $\mathbb{R}^d$, $x \in \operatorname{supp}(\mu)$, $r > 0$, and $0 \le \varepsilon < 1$. Define

$$\operatorname{por}(\mu, x, r, \varepsilon) = \sup\left\{a \in [0,1] \;:\; \begin{array}{l} \text{there is } y \in \mathbb{R}^d \text{ with } \overline{B}(y, ar) \subset \overline{B}(x, r), \\ \mu(\overline{B}(y, ar)) \le \varepsilon\mu(\overline{B}(x, r)) \end{array}\right\}.$$

The supremum of the empty admissible set is zero. A radius-zero closed ball means the singleton consisting of its centre. Such an admissible ball is an **ε-hole of relative radius a**.

For $0 < \alpha \le 1/2$, $0 < \eta \le 1$, and $0 \le \varepsilon < 1$, μ is **mean $(\alpha, \eta, \varepsilon)$-porous** if for μ-almost every x,

$$\liminf_{N\to\infty} \frac{1}{N} \#\{0 \le n < N \;:\; \operatorname{por}(\mu, x, 2^{-n}, \varepsilon) \ge \alpha\} \ge \eta.$$

It is **mean (α, η)-porous** if it is mean $(\alpha, \eta, \varepsilon)$-porous for every $0 < \varepsilon < 1$. The phrase **weakly mean porous** is used only with displayed parameters $(\alpha, \eta, \varepsilon)$.

Definition — I.6.05: Dyadic hole depth and dyadic mean porosity

The levels n, j are nonnegative integers, and the depths k, ℓ are positive integers throughout this definition and the dyadic results below.

Let D_n be the half-open dyadic cubes from D-I.5.08, and let $Q_n(x)$ be the unique cube in D_n containing x. If $\mu(Q_n(x)) > 0$, set

$$\operatorname{por}_2(\mu, x, n, \varepsilon) = \min\left\{k \ge 1 \;:\; \begin{array}{l} \text{there is } R \in \mathscr{D}_{n+k},\ R \subset Q_n(x), \\ \mu(R) \le \varepsilon\mu(Q_n(x)) \end{array}\right\},$$

with minimum $+\infty$ if no such R exists. Partition atoms of zero mass are discarded; for μ-almost every x, every $Q_n(x)$ has positive mass.

For $\ell \ge 1$, $0 < \eta \le 1$, and $0 \le \varepsilon < 1$, μ is **dyadic mean $(\ell, \eta, \varepsilon)$-porous** if for μ-almost every x,

$$\liminf_{N\to\infty} \frac{1}{N} \#\{0 \le j < N \;:\; \operatorname{por}_2(\mu, x, j\ell, \varepsilon) \le \ell\} \ge \eta.$$

It is **dyadic mean (ℓ, η)-porous** if this holds for every $0 < \varepsilon < 1$. Sampling at block beginnings $j\ell$ is frozen here so the entropy proof has no overlapping-block ambiguity.

Definition — I.6.06: Rectifiable and purely unrectifiable sets and measures

A set $E \subset \mathbb{R}^d$ is **countably m-rectifiable** if there are Lipschitz maps $f_j : \mathbb{R}^m \to \mathbb{R}^d$ such that

$$\mathscr{H}^m\left(E \setminus \bigcup_{j=1}^{\infty} f_j(\mathbb{R}^m)\right) = 0.$$

It is **purely m-unrectifiable** if

$$\mathscr{H}^m(E \cap F) = 0$$

for every countably m-rectifiable set F.
A finite Borel measure μ on $\mathbb{R}^d$ is **m-rectifiable** if some countable union of Lipschitz images $f_j(\mathbb{R}^m)$ carries μ. It is **purely m-unrectifiable** if

$$\mu(f(\mathbb{R}^m)) = 0$$

for every Lipschitz $f : \mathbb{R}^m \to \mathbb{R}^d$. Absolute continuity with respect to $\mathscr{H}^m$ is not part of the measure definition.

Definition — I.6.07: Grassmannians, cones, and approximate tangent planes

The **Grassmannian** $G(d, m)$ is the set of all m-dimensional linear subspaces of $\mathbb{R}^d$. For $x \in \mathbb{R}^d$, $V \in G(d, m)$, and $s > 0$, define

$$X(x, V, s) = \{y \in \mathbb{R}^d : \operatorname{dist}(y, x + V) < s|y - x|\}.$$

Let E be $\mathscr{H}^m$-measurable and $x \in E$. A plane $V \in G(d, m)$ is an **approximate tangent m-plane to E at x** if

$$\Theta^{*m}(E, x) > 0$$

and for every $s > 0$,

$$\lim_{r \downarrow 0} r^{-m} \mathscr{H}^m\big((E \cap \overline{B}(x, r)) \setminus X(x, V, s)\big) = 0.$$

Definition — I.6.08: Directional visible and bi-visible parts

Let $K \subset \mathbb{R}^d$ and $e \in S^{d-1}$. Put

$$\mathbb{R}_+ e = \{te : t \geq 0\}.$$

The **visible part of K in direction e** is

$$\operatorname{Vis}_e(K) = \{z \in K : (z + \mathbb{R}_+ e) \cap K = \{z\}\}.$$

The **bi-visible part** is

$$\operatorname{Vis}_e^2(K) = \{z \in K : (z + \mathbb{R}e) \cap K = \{z\}\}.$$

The orthogonal projection onto the hyperplane $e^\perp$ is

$$P_{e^\perp}(z) = z - (z \cdot e)e.$$

Definition — I.6.09: Visibility from a point and radial projection

Let $K \subset \mathbb{R}^d$ and $a \in \mathbb{R}^d \setminus K$. For $z \in K$,

$$[a, z] = \{(1-t)a + tz \,:\, 0 \le t \le 1\}.$$

The **visible part of K from a** is

$$\mathrm{Vis}_a(K) = \{z \in K \,:\, [a, z] \cap K = \{z\}\}.$$

The **radial projection from a** is

$$\pi_a \,:\, \mathbb{R}^d \setminus \{a\} \longrightarrow S^{d-1}, \qquad \pi_a(z) = \frac{z-a}{|z-a|}.$$

Throughout, $d \ge 1$, Euclidean balls are closed, and $\mathscr{L}^d$ is Lebesgue measure on $\mathbb{R}^d$.

I.6.1 Density normalization

Let $1 \le m \le d$ be an integer and put

$$\omega_m = \mathscr{L}^m(\overline{B}(0,1)). \tag{1.1}$$

Recall that $\mathscr{H}^m$ from I.1 is diameter normalized. Define

$$\mathscr{H}^m = \frac{\omega_m}{2^m}\mathscr{H}^m. \tag{1.2}$$

For a locally finite Borel measure μ on $\mathbb{R}^d$, define

$$\Theta^m_*(\mu, x) = \liminf_{r\downarrow 0} \frac{\mu(\overline{B}(x,r))}{\omega_m r^m}, \qquad \Theta^{*m}(\mu, x) = \limsup_{r\downarrow 0} \frac{\mu(\overline{B}(x,r))}{\omega_m r^m}. \tag{1.3}$$

When these agree, their common value is $\Theta^m(\mu, x)$. If E is $\mathscr{H}^m$-measurable, the notation $\Theta^m_*(E, x)$, $\Theta^{*m}(E, x)$, and $\Theta^m(E, x)$ refers to the restriction $\mathscr{H}^m \restriction E$. When this restriction is not locally finite, use the same quotients directly, with extended nonnegative values, as specified in D-I.6.01.

Lemma — 1.1 (L-I.6.01: normalization and density concordance) [F]

The measures in (1.2) have the same null sets and infinite sets, and their restrictions differ by the positive scalar $\omega_m/2^m$. Consequently every comparison of $\dim_{\mathrm{H}} E$ with the integer m is unchanged. Moreover,

$$\frac{\mathscr{H}^m(E \cap \overline{B}(x,r))}{\omega_m r^m} = \frac{\mathscr{H}^m(E \cap \overline{B}(x,r))}{(2r)^m}. \tag{1.4}$$

Proof

Equation (1.2) is multiplication by the positive finite constant $\omega_m/2^m$. It therefore preserves zero sets, infinite values, and measure restrictions up to that scalar. Multiplication by a positive finite constant preserves the alternatives $\mathscr{H}^m(E) = 0$ and $\mathscr{H}^m(E) = +\infty$. The comparison theorem from I.1 therefore gives the same conclusions $\dim_{\mathrm{H}} E \le m$ or $\dim_{\mathrm{H}} E \ge m$ from either normalization. Finally,

$$\frac{(\omega_m/2^m)\mathscr{H}^m(E \cap \overline{B}(x,r))}{\omega_m r^m} = \frac{\mathscr{H}^m(E \cap \overline{B}(x,r))}{2^m r^m},$$

which is (1.4). □

I.6.2 Porosity of sets and measures

For $E \subset \mathbb{R}^d$, $x \in E$, and $r > 0$, define

$$\operatorname{por}(E,x,r) = \sup\{a \ge 0 \,:\, \text{some } \overline{B}(y,ar) \subset \overline{B}(x,r) \setminus E\}. \tag{2.1}$$

The supremum of the empty set is zero, and

$$\operatorname{por}(E,x) = \liminf_{r \downarrow 0} \operatorname{por}(E,x,r). \tag{2.2}$$

For $0 < \alpha \le 1/2$ and $r_0 > 0$, E is uniformly (α, r_0)-porous if

$$\operatorname{por}(E,x,r) \ge \alpha \quad (x \in E,\ 0 < r \le r_0). \tag{2.3}$$

It is mean (α,η)-porous, where $0 < \eta \le 1$, if for every $x \in E$,

$$\liminf_{N\to\infty} \frac{1}{N} \#\{0 \le n < N \,:\, \operatorname{por}(E,x,2^{-n}) \ge \alpha\} \ge \eta. \tag{2.4}$$

Let μ be a Borel probability, $x \in \operatorname{supp}(\mu)$, $r > 0$, and $0 \le \varepsilon < 1$. Its epsilon-porosity is

$$\operatorname{por}(\mu,x,r,\varepsilon) = \sup\left\{a \in [0,1] \,:\, \begin{array}{l} \overline{B}(y,ar) \subset \overline{B}(x,r) \text{ for some } y, \\ \mu(\overline{B}(y,ar)) \le \varepsilon\mu(\overline{B}(x,r)) \end{array}\right\}. \tag{2.5}$$

The measure is mean $(\alpha,\eta,\varepsilon)$-porous if, for μ-almost every x,

$$\liminf_{N\to\infty} \frac{1}{N} \#\{0 \le n < N \,:\, \operatorname{por}(\mu,x,2^{-n},\varepsilon) \ge \alpha\} \ge \eta. \tag{2.6}$$

It is mean (α,η)-porous if (2.6) holds for every $0 < \varepsilon < 1$.

Lemma — 2.1 (L-I.6.02: elementary porosity calculus) [F]

Let $x \in E$ and $r > 0$.

(1) $0 \le \operatorname{por}(E,x,r) \le 1/2$.

(2) If $E \subset F$ and $x \in E$, then

$$\operatorname{por}(F,x,r) \le \operatorname{por}(E,x,r). \tag{2.7}$$

(3) If $T(z) = \lambda O z + t$ is a similarity of ratio $\lambda > 0$, then

$$\operatorname{por}(T(E), T(x), \lambda r) = \operatorname{por}(E, x, r). \tag{2.8}$$

(4) A uniformly porous set has empty interior. If E is uniformly (α, r_0)-porous, then $\operatorname{cl}(E)$ is uniformly $(\alpha/4, r_0/2)$-porous.

Proof

Nonnegativity follows from the definition. Suppose

$$\overline{B}(y, ar) \subset \overline{B}(x, r) \setminus E.$$

Since $x \in E$, the point x is not in $\overline{B}(y, ar)$, and hence $|x - y| > ar$. Containment of the two balls gives

$$|x - y| + ar \leq r.$$

Thus $2ar < r$. Taking the supremum gives $\operatorname{por}(E, x, r) \leq 1/2$.

If $E \subset F$, then every ball contained in $\overline{B}(x, r) \setminus F$ is contained in $\overline{B}(x, r) \setminus E$; this proves (2.7). A similarity maps closed balls to closed balls, multiplies every radius by λ, and maps set complements bijectively. The admissible relative radii on the two sides of (2.8) are therefore identical.

If E contained an open ball, one could choose $x \in E$ and a sufficiently small r for which $\overline{B}(x, r) \subset E$. This is incompatible with a positive-radius hole in that ball, so a uniformly porous set has empty interior.

Now take $x \in \operatorname{cl}(E)$ and $0 < r \leq r_0/2$. Choose $x' \in E$ with $|x - x'| < r/4$. Since porosity is a supremum, (2.3) at x' and radius $r/2$ supplies some $b > 3\alpha/4$ and a hole

$$\overline{B}(y, br/2) \subset \overline{B}(x', r/2) \setminus E. \tag{2.9}$$

The smaller closed ball $\overline{B}(y, \alpha r/4)$ is disjoint from $\operatorname{cl}(E)$: otherwise an approximating sequence from E would eventually enter the interior of the larger ball in (2.9), since $\alpha r/4 < br/2$. Also

$$\overline{B}(y, \alpha r/4) \subset \overline{B}(x', r/2) \subset \overline{B}(x, 3r/4) \subset \overline{B}(x, r).$$

This gives the asserted $(\alpha/4, r_0/2)$ porosity. □

Theorem — 2.2 (T-I.6.03: mean porosity forces nullity or singularity) [F]

Let $0 < \alpha \leq 1/2$ and $0 < \eta \leq 1$.

(1) Every Lebesgue-measurable mean (α, η)-porous set $E \subset \mathbb{R}^d$ satisfies $\mathscr{L}^d(E) = 0$.

(2) Every mean (α, η)-porous Borel probability μ on $\mathbb{R}^d$ is singular with respect to $\mathscr{L}^d$.

Proof

Suppose first that $\mathscr{L}^d(E) > 0$. By the Lebesgue density theorem, there is $x \in E$ such that

$$\lim_{r \downarrow 0} \frac{\mathscr{L}^d(\overline{B}(x,r) \setminus E)}{\omega_d r^d} = 0. \tag{2.10}$$

Condition (2.4) supplies infinitely many n for which the supremum in (2.1) is at least α. At every such scale there is an admissible hole of relative radius greater than $\alpha/2$, and hence a concentric hole of relative radius exactly $\alpha/2$. Therefore

$$\frac{\mathscr{L}^d(\overline{B}(x,2^{-n}) \setminus E)}{\omega_d 2^{-nd}} \geq (\alpha/2)^d \tag{2.11}$$

along infinitely many scales. This contradicts (2.10), proving the first claim.

For the measure claim, write the Lebesgue decomposition

$$\mu = f\mathscr{L}^d + \mu_s. \tag{2.12}$$

Assume that the absolutely continuous part is nonzero. Since $\{0 < f < +\infty\}$ is the union of the sets

$$A_k = \{z : 1/k \leq f(z) \leq k\},$$

some A_k has positive Lebesgue measure. Choose $x \in A_k$ satisfying all of the following:

- x is a Lebesgue density point of A_k;
- the differentiation limit
$$\lim_{r \downarrow 0} \frac{\mu(\overline{B}(x,r))}{\omega_d r^d} = f(x) \tag{2.13}$$
holds; and
- (2.6) holds at x for
$$\varepsilon_0 = \frac{(\alpha/2)^d}{8k^2}. \tag{2.14}$$

The first two conditions fail only on a Lebesgue-null subset of A_k. The third fails on a μ-null set, hence on a Lebesgue-null subset of A_k because $f \geq 1/k$ there. Such an x therefore exists.

For all sufficiently small r, (2.13) and $f(x) \leq k$ give

$$\mu(\overline{B}(x,r)) \leq 2k\omega_d r^d, \tag{2.15}$$

while density of A_k at x gives

$$\mathscr{L}^d(\overline{B}(x,r) \setminus A_k) \leq \frac{1}{2}\omega_d(\alpha/2)^d r^d. \tag{2.16}$$

Let $\overline{B}(y,(\alpha/2)r) \subset \overline{B}(x,r)$. By (2.16),

$$\mathscr{L}^d(\overline{B}(y,(\alpha/2)r) \cap A_k) \geq \frac{1}{2}\omega_d(\alpha/2)^d r^d.$$

Using $f \geq 1/k$ on A_k and (2.15),

$$\frac{\mu(\overline{B}(y,(\alpha/2)r))}{\mu(\overline{B}(x,r))} \geq \frac{(1/k)(\omega_d(\alpha/2)^d r^d/2)}{2k\omega_d r^d} = \frac{(\alpha/2)^d}{4k^2} > \varepsilon_0. \tag{2.17}$$

If $\mathrm{por}(\mu,x,r,\varepsilon_0) \geq \alpha$, the definition of the supremum would supply an ε_0-hole of relative radius greater than $\alpha/2$; its concentric radius-$(\alpha/2)r$ subball would contradict (2.17). No attainment of the supremum is being assumed. Hence

$$\mathrm{por}(\mu,x,r,\varepsilon_0) < \alpha$$

at every sufficiently small scale, contradicting (2.6) and $\eta > 0$. Therefore the absolutely continuous part in (2.12) is zero. □

I.6.3 A self-contained dyadic entropy mechanism

All levels and depths in this section are integers: $n, j \geq 0$ and $k, \ell \geq 1$. All logarithms are natural unless a base is displayed.

Let D_n be the half-open dyadic cubes of side 2^{-n} and $Q_n(x)$ the unique member containing x. For $\mu(Q_n(x)) > 0$, define

$$\mathrm{por}_2(\mu,x,n,\varepsilon) = \min\left\{k \geq 1 \;:\; \begin{array}{l} R \subset Q_n(x) \text{ for some } R \in \mathscr{D}_{n+k}, \\ \mu(R) \leq \varepsilon\mu(Q_n(x)) \end{array}\right\}. \tag{3.1}$$

The minimum is $+\infty$ when the set is empty. For $\ell \geq 1$, μ is dyadic mean (ℓ,η,ε)-porous if, for μ-almost every x,

$$\liminf_{N\to\infty} \frac{1}{N}\#\{0 \leq j < N \;:\; \mathrm{por}_2(\mu,x,j\ell,\varepsilon) \leq \ell\} \geq \eta. \tag{3.2}$$

It is dyadic mean (ℓ,η)-porous if (3.2) holds for every $0 < \varepsilon < 1$.

If $Q \in D_n$ and $\mu(Q) > 0$, let

$$H_\ell(\mu,Q) = -\sum_{\substack{R\in\mathscr{D}_{n+\ell}\\ R\subset Q}} p_R \log p_R, \qquad p_R = \frac{\mu(R)}{\mu(Q)}, \tag{3.3}$$

where $0\log 0 = 0$. On a zero-mass parent we set $H_\ell(\mu,Q) = 0$. This convention makes the functions of x below defined everywhere; zero-mass dyadic atoms form a countable union of μ-measure zero.

Proof-internal Claim 3.1 (finite-alphabet information moments)

For every integer $M \geq 2$, there is $C_M < \infty$ such that every probability vector $(p_1, \ldots, p_M)$ satisfies

$$\sum_{i\,:\,p_i>0} p_i\big(\log(1/p_i)\big)^2 \leq C_M. \tag{3.4}$$

Proof

The function $t(\log(1/t))^2$, extended by zero at $t = 0$, is continuous on $[0, 1]$ and therefore has a finite maximum C. The left side of (3.4) is at most MC. □

Proof-internal Claim 3.2 (bounded-variance martingale-difference law)

A **filtration** on a probability space $(\Omega, \mathscr{F}, \mu)$ is an increasing sequence $(\mathscr{F}_j)_{j\geq 0}$ of sub-sigma-algebras of $\mathscr{F}$. A sequence $(X_j)_{j\geq 0}$ is a **martingale-difference sequence relative to this filtration** if each X_j is real-valued, $\mathscr{F}_{j+1}$-measurable, integrable, and

$$\int_A X_j \, d\mu = 0 \quad \text{for every } A \in \mathscr{F}_j. \tag{3.5}$$

Here $\mathbb{E}(Y) = \int Y \, d\mu$ denotes expectation and $\mathbb{P}(A) = \mu(A)$ probability; (3.5) is the precise zero-mean condition used in the proof. Suppose also that

$$\mathbb{E}(X_j^2) \leq C \quad (j \geq 0). \tag{3.6}$$

Then

$$\frac{1}{N} \sum_{j=0}^{N-1} X_j \longrightarrow 0 \quad \text{almost surely.} \tag{3.7}$$

Proof

Put $S_N = \sum_{j=0}^{N-1} X_j$. Martingale differences are orthogonal in L^2: if $i < j$, then X_i is $\mathscr{F}_j$-measurable. Approximate X_i in L^2 by $\mathscr{F}_j$-measurable simple functions. Equation (3.5) makes the integral of each such simple function times X_j equal to zero, and Cauchy–Schwarz passes to the L^2 limit. Hence $\mathbb{E}(X_i X_j) = 0$, and therefore

$$\mathbb{E}(S_N^2) = \sum_{j=0}^{N-1} \mathbb{E}(X_j^2) \leq CN. \tag{3.8}$$

We record the needed maximal inequality. Let

$$\tau = \min\{1 \leq k \leq N : |S_k| \geq A\},$$

with $\tau = +\infty$ if the set is empty. For $k \leq j$, the function $1_{\{\tau=k\}} S_k$ is $\mathscr{F}_j$-measurable. The same simple-function approximation and (3.5) give

$$\mathbb{E}(\mathbf{1}_{\{\tau=k\}} S_k X_j) = 0.$$

Define the everywhere finite random variable

$$Y = \sum_{k=1}^{N} \mathbf{1}_{\{\tau=k\}} S_k.$$

It equals S_τ on $\{\tau \leq N\}$ and is zero elsewhere. Summing first over $k \leq j < N$ and then over $1 \leq k \leq N$ yields

$$\mathbb{E}[Y(S_N - Y)] = 0.$$

Therefore, by Cauchy–Schwarz,

$$\begin{aligned}\mathbb{E}(Y^2) &= \mathbb{E}(YS_N)\\ &\le \mathbb{E}(Y^2)^{1/2}\mathbb{E}(S_N^2)^{1/2}.\end{aligned}$$

If $\mathbb{E}(Y^2) > 0$, division by its square root gives $\mathbb{E}(Y^2) \le \mathbb{E}(S_N^2)$; if it is zero, that inequality is immediate. Since $Y^2 \ge A^2 \mathbf{1}_{\{\tau \le N\}}$, in either case,

$$A^2 \mathbb{P}\left(\max_{1\le k\le N} |S_k| \ge A\right) \le \mathbb{E}(S_N^2) \le CN. \tag{3.9}$$

Set $N = 2^q$ and $A = \varepsilon 2^q$. The right side of the probability bound in (3.9) is $C\varepsilon^{-2}2^{-q}$, which is summable in q. By Borel–Cantelli, the displayed maximum is eventually smaller than $\varepsilon 2^q$ almost surely. Apply this for the countable sequence $\varepsilon = 1/m$, $m \ge 1$, and intersect the resulting full-measure sets. Then

$$2^{-q} \max_{1\le k\le 2^q} |S_k| \longrightarrow 0 \quad \text{almost surely.}$$

If $2^{q-1} < N \le 2^q$, then

$$\frac{|S_N|}{N} \le 2\,2^{-q} \max_{1\le k\le 2^q} |S_k|.$$

This proves (3.7). □

Lemma — 3.3 (L-I.6.04: fixed-block dyadic information law) [F]

Let μ be a Borel probability and $\ell \ge 1$. Then for μ-almost every x,

$$\lim_{N\to\infty} \frac{1}{N}\left(-\log \mu(Q_{N\ell}(x)) - \sum_{j=0}^{N-1} H_\ell(\mu, Q_{j\ell}(x))\right) = 0. \tag{3.10}$$

Proof

For μ-almost every x, every dyadic atom containing x has positive mass. On this full-measure set define the information increment

$$I_j(x) = -\log \frac{\mu(Q_{(j+1)\ell}(x))}{\mu(Q_{j\ell}(x))}. \tag{3.11}$$

On the null union of zero-mass atoms, set $I_j = 0$. Let G_j be the sigma-algebra generated by $D_{j\ell}$. The distribution of the next block inside a positive-mass parent $Q \in D_{j\ell}$ is the probability vector in (3.3). Consequently

$$\frac{1}{\mu(Q)} \int_Q I_j \, d\mu = H_\ell(\mu, Q) \qquad (Q \in \mathscr{D}_{j\ell},\ \mu(Q) > 0). \tag{3.12}$$

Put

$$X_j = I_j - H_\ell(\mu, Q_{j\ell}(x)).$$

Then (X_j) is a martingale-difference sequence. Claim 3.1, with $M = 2^{d\ell}$, bounds the conditional second moment of I_j uniformly. Also $0 \le H_\ell \le \log M$, so the second moments of X_j are uniformly bounded. Claim 3.2 yields

$$\frac{1}{N}\sum_{j=0}^{N-1} X_j \longrightarrow 0 \quad \text{for } \mu\text{-almost every } x. \tag{3.13}$$

The information increments telescope:

$$\sum_{j=0}^{N-1} I_j(x) = \log \frac{\mu(Q_0(x))}{\mu(Q_{N\ell}(x))}. \tag{3.14}$$

Since $0 < \mu(Q_0(x)) \le 1$ is independent of N, equations (3.13)–(3.14) give (3.10). □

Proof-internal Claim 3.4 (entropy with one light atom)

Let $M \ge 2$, $0 \le \varepsilon \le 1/M$, and let $(p_1, \ldots, p_M)$ be a probability vector with $p_i \le \varepsilon$ for some i. Then

$$-\sum_{j=1}^{M} p_j \log p_j \le h(\varepsilon) + (1-\varepsilon)\log(M-1), \tag{3.15}$$

where

$$h(t) = -t\log t - (1-t)\log(1-t).$$

Proof

Fix $p_i = t$. Concavity of entropy, or Jensen's inequality for $-u\log u$, shows that the remaining entropy is largest when the other `M-1` coordinates all equal $(1-t)/(M-1)$. Hence the entropy is at most

$$h(t) + (1-t)\log(M-1). \tag{3.16}$$

The derivative of the right side is

$$\log \frac{1-t}{t(M-1)},$$

which is nonnegative on $(0, 1/M]$. Continuity at zero extends monotonicity to $[0, 1/M]$. Since $t \le \varepsilon \le 1/M$, (3.16) is at most its value at ε, proving (3.15). □

Theorem — 3.5 (T-I.6.05: dyadic mean-porosity deficit) [F]

Suppose μ is dyadic mean (ℓ, η)-porous. Set

$$M = 2^{d\ell}. \tag{3.17}$$

Then

$$\overline{\dim}_{\mathrm{P}}\mu \leq d + \frac{\eta}{\ell \log 2}\log(1 - M^{-1}) \leq d - \frac{\eta}{\ell \log 2}M^{-1}. \tag{3.18}$$

Proof

Choose a sequence $\varepsilon_q \downarrow 0$ with $0 < \varepsilon_q \leq 1/M$. Intersect the full-measure sets on which (3.2) holds for every ε_q and on which Lemma 3.3 holds. Fix x in this intersection.

For a fixed q, call j good when

$$\mathrm{por}_2(\mu, x, j\ell, \varepsilon_q) \leq \ell.$$

At a good block there is a dyadic descendant of depth at most ℓ whose relative mass is at most ε_q. Any level-ℓ descendant inside that cube also has relative mass at most ε_q. Claim 3.4 therefore gives

$$H_\ell(\mu, Q_{j\ell}(x)) \leq H_q := h(\varepsilon_q) + (1 - \varepsilon_q)\log(M - 1). \tag{3.19}$$

At every other block, entropy is at most $\log M$. Since the lower density of good blocks is at least η and $H_q \leq \log M$, it follows that

$$\limsup_{N\to\infty} \frac{1}{N}\sum_{j=0}^{N-1} H_\ell(\mu, Q_{j\ell}(x)) \leq \eta H_q + (1 - \eta)\log M. \tag{3.20}$$

Combining (3.10) and (3.20),

$$\limsup_{N\to\infty} \frac{-\log \mu(Q_{N\ell}(x))}{N\ell \log 2} \leq \frac{\eta H_q + (1 - \eta)\log M}{\ell \log 2}. \tag{3.21}$$

The cube $Q_{N\ell}(x)$ is contained in $\overline{B}(x, \sqrt{d}2^{-N\ell})$. Thus

$$\mu(\overline{B}(x, \sqrt{d}\, 2^{-N\ell})) \geq \mu(Q_{N\ell}(x)).$$

Because the logarithmic denominator is negative and $\log(\sqrt{d}2^{-N\ell}) = -N\ell \log 2 + O_d(1)$, geometric sampling of local dimension from L-I.4.01 gives

$$\overline{d}_\mu(x) \leq \limsup_{N\to\infty} \frac{-\log \mu(Q_{N\ell}(x))}{N\ell \log 2}. \tag{3.22}$$

Let $q \to \infty$ in (3.21), and use (3.22). Since

$$H_q \longrightarrow \log(M - 1),$$

we obtain

$$\overline{d}_\mu(x) \leq \frac{\eta \log(M - 1) + (1 - \eta)\log M}{\ell \log 2} = d + \frac{\eta}{\ell \log 2}\log(1 - M^{-1}). \tag{3.23}$$

This holds at μ-almost every x. The packing local-to-global dictionary T-I.4.07 proves the first inequality in (3.18). The second follows from $\log(1-u) \le -u$ for $0 < u < 1$. $\square$

Corollary — 3.6 (C-I.6.06: support-to-measure porosity) [F]

Let μ be a Borel probability.
(1) If $\operatorname{supp}(\mu)$ is mean (α, η)-porous as a set, then μ is mean (α, η)-porous.
(2) Suppose that for every $j \ge 0$ and every positive-mass $Q \in D_{j\ell}$, some dyadic descendant of Q at depth at most ℓ is disjoint from $\operatorname{supp}(\mu)$. Then

$$\overline{\dim}_{\mathrm{P}}\mu \le d + \frac{1}{\ell \log 2} \log(1 - 2^{-d\ell}) < d. \tag{3.24}$$

Proof

A ball disjoint from $\operatorname{supp}(\mu)$ has μ-mass zero. Hence every set hole in the support is an epsilon-hole for every $0 < \varepsilon < 1$, with the same center and relative radius. This proves the first assertion.

Under the dyadic hypothesis, every sampled positive-mass block has a zero-mass descendant at depth at most ℓ. Thus μ is dyadic mean $(\ell, 1, \varepsilon)$-porous for every $0 < \varepsilon < 1$. Apply Theorem 3.5 with $\eta = 1$ to obtain (3.24). $\square$

Guided result — 3.7 (G-I.6.07: sharp Euclidean mean-porosity deficit) [G]

[G | currency: Published | omitted: randomized Euclidean-to-dyadic conversion and variable-depth porous-tree variational estimate | source: SHM12-WMP, Theorem 1.1 and Sections 4–5]
Let $0 < \alpha \le 1/2$ and $0 < \eta \le 1$. If μ is a mean (α, η)-porous Borel probability on $\mathbb{R}^d$, then

$$\overline{\dim}_{\mathrm{P}}\mu \le d - c_d \eta \alpha^d, \qquad c_d = \frac{2}{5 \log 2\, 24^d d^{d/2}}. \tag{3.25}$$

References. The exact statement, definitions, and omitted proof are in Shmerkin [Shm12, Theorem 1.1 and Sections 4–5, pp. 2–3 and 7–18]. A complementary proof map appears in Shmerkin [Shm11, Theorem 8 and Sections 4.2.1–4.2.3, pp. 12–16]. The source writes its measure porosity using an infimum over hole centres. Every admissible hole in D-I.6.04 makes that infimum at most the same ε. Thus our porosity is at most the source porosity, and our mean-porosity hypothesis implies the source hypothesis with unchanged α, η. No attainment of the source infimum is needed. The source's packing dimension is the essential supremum of upper local dimensions (Section 2.3, p. 8), which equals our full-mass upper packing dimension by T-I.4.07.

Guided proof architecture

Choose

$$\ell = \left\lceil \left| \log_2 \frac{\alpha}{4\sqrt{d}} \right| \right\rceil. \tag{3.26}$$

A Euclidean α-hole at scale 2^{-n} contains a dyadic cube of relative side $2^{-\ell}$ whenever the ambient dyadic cube places the base point a fixed distance from its boundary. The randomized-grid lemma in SHM12-WMP, Section 4, chooses a translated dyadic filtration and retains a fixed dimension-dependent proportion of the porous scales. The simultaneous almost-every-point and every-ε quantifiers are part of this omitted lemma; they may not be replaced by a grid chosen separately for each point.

For comparison, our fixed-block theorem gives a deficit $-\eta \log(1-2^{-d\ell})/(\ell \log 2)$ when the block frequency is η. Apart from this frequency, its scale as α tends to zero is

$$\frac{2^{-d\ell}}{\ell} \asymp_d \frac{\alpha^d}{|\log \alpha|}. \tag{3.27}$$

This is a comparison with the model proved here, not an assertion that the source's frequency over all levels equals the frequency along the fixed subsequence $j\ell$. The source controls the contribution of all dyadic levels to its variable-depth subdivision as part of the omitted argument.

The removal of ℓ^{-1} uses a nonuniform tree. When a porous parent Q has a light descendant R at depth ℓ, write

$$Q = Q_0 \supset Q_1 \supset \cdots \supset Q_\ell = R$$

for the ancestor chain. Partition Q into R and, for each $0 \le j < \ell$, the $2^d - 1$ siblings of Q_{j+1} inside Q_j. For the conditional masses $p_A = \mu(A)/\mu(Q)$ of these variable-size children, for $\mu(Q) > 0$, define

$$H(Q) = -\sum_A p_A \log p_A, \qquad \lambda(Q) = \sum_A p_A \log \frac{\ell(Q)}{\ell(A)}. \tag{3.28}$$

Here $\ell(A)$ is the side length of the cube A. The quantity $\lambda(Q)$ is the expected logarithmic contraction at this subdivision; this is the meaning of **Lyapunov increment** in this paragraph.

At a nonporous parent, use the ordinary first-level dyadic partition. The source's variable-depth information law, Theorem 3.1 and Corollary 3.2 on pp. 9–10, pairs information increments with these log-scale increments. It bounds upper local dimension by the limsup of accumulated entropy divided by accumulated logarithmic contraction. This version of the information law belongs to the externally cited variable-depth block; Lemma 3.3 alone proves only the fixed-depth version.

The omitted variable-depth variational estimate proves, uniformly over the location of R and over every sufficiently light mass p_R, that the porous partition has an entropy/Lyapunov deficit of size comparable to $2^{-d\ell}$, rather than $2^{-d\ell}/\ell$. Averaging this deficit over the retained proportion of porous scales, using (3.26), and carrying through the explicit geometric constants gives

$$d - \overline{\dim}_{\mathrm{P}} \mu \ge \frac{2}{5 \log 2 \, 24^d d^{d/2}} \eta \alpha^d,$$

which is (3.25). The randomized conversion and the variable-depth argument (its information law, variational estimate, and frequency bookkeeping) are the blocks not proved here. No later F result uses Theorem 3.7.

I.6.4 Rectifiable and purely unrectifiable objects

Let $1 \leq m \leq d$. A set $E \subset \mathbb{R}^d$ is countably m-rectifiable if there are Lipschitz maps $f_j : \mathbb{R}^m \to \mathbb{R}^d$ such that

$$\mathcal{H}^m\left(E \setminus \bigcup_{j=1}^{\infty} f_j(\mathbb{R}^m)\right) = 0. \tag{4.1}$$

It is purely m-unrectifiable if

$$\mathcal{H}^m(E \cap F) = 0 \tag{4.2}$$

for every countably m-rectifiable F.

A finite Borel measure μ is m-rectifiable if a countable union of Lipschitz images of $\mathbb{R}^m$ carries it. It is purely m-unrectifiable if

$$\mu(f(\mathbb{R}^m)) = 0 \tag{4.3}$$

for every Lipschitz $f : \mathbb{R}^m \to \mathbb{R}^d$. This definition does not include $\mu \ll \mathcal{H}^m$.

Lemma — 4.1 (L-I.6.08: elementary rectifiability calculus) [F]

(1) A countably m-rectifiable set has Hausdorff dimension at most m.
(2) A set that is both countably m-rectifiable and purely m-unrectifiable has $\mathcal{H}^m$ measure zero.
(3) A finite measure that is both m-rectifiable and purely m-unrectifiable is zero.
(4) Let $\Phi : \mathbb{R}^d \to \mathbb{R}^q$ be a bi-Lipschitz embedding. Then μ is m-rectifiable, respectively purely m-unrectifiable, if and only if $\Phi_\# \mu$ has the corresponding property.

Proof

Each Lipschitz image $f_j(\mathbb{R}^m)$ has Hausdorff dimension at most m by T-I.1.08. The exceptional set in (4.1) has zero $\mathcal{H}^m$ measure and therefore Hausdorff dimension at most m. Countable stability of Hausdorff dimension proves the first assertion.

If E has both set properties, use $F = E$ in (4.2) to obtain $\mathcal{H}^m(E) = 0$. If μ has both measure properties, choose Lipschitz images $(f_j(\mathbb{R}^m))$ whose union carries μ. Pure unrectifiability gives zero mass to every image, hence to their countable union. Thus $\mu = 0$.

For the embedding assertion, bi-Lipschitz invariance of Hausdorff dimension gives $d = \dim_{\mathrm{H}} \Phi(\mathbb{R}^d) \leq q$; hence $m \leq q$, as required for the target rectifiability convention. Suppose first that μ is carried by $\cup_j f_j(\mathbb{R}^m)$. Then $\Phi_\# \mu$ is carried by the images of the Lipschitz maps $\Phi \circ f_j$, so it is rectifiable. Conversely, suppose $\Phi_\# \mu$ is carried by $\cup_j g_j(\mathbb{R}^m)$. Put

$$A_j = g_j^{-1}(\Phi(\mathbb{R}^d)).$$

The image $\Phi(\mathbb{R}^d)$ is closed: a convergent sequence of image points has a Cauchy sequence of preimages by the inverse Lipschitz bound, and its limit maps to the given image limit. Each Lipschitz image of $\mathbb{R}^m$ is a countable union of compact images of closed balls and hence is Borel. Thus all measure evaluations and inverse images below are defined.

The map $\Phi^{-1} \circ g_j : A_j \to \mathbb{R}^d$ is Lipschitz. Each coordinate has a Lipschitz extension from A_j to $\mathbb{R}^m$: if $u : A_j \to R$ is L-Lipschitz and A_j is nonempty, then

$$\tilde{u}(x) = \inf_{a \in A_j} (u(a) + L|x - a|)$$

is finite, L-Lipschitz, and equals u on A_j; for an empty domain use the zero extension. To check finiteness, fix $a_0 \in A_j$. For every $a \in A_j$,

$$u(a_0) - L|x - a_0| \le u(a) + L|x - a|, \qquad \tilde{u}(x) \le u(a_0) + L|x - a_0|.$$

The triangle inequality also gives $\tilde{u}(x) \le \tilde{u}(x') + L|x - x'|$ and the reverse inequality, while for $x \in A_j$, the choice $a = x$ gives $\tilde{u}(x) \le u(x)$ and the Lipschitz inequality for u gives the reverse bound. Taking all coordinate extensions gives a Lipschitz $h_j : \mathbb{R}^m \to \mathbb{R}^d$. Since

$$\Phi^{-1}(g_j(\mathbb{R}^m) \cap \Phi(\mathbb{R}^d)) \subset h_j(\mathbb{R}^m),$$

the union of the images of the h_j carries μ. Hence μ is rectifiable.

Now suppose μ is purely unrectifiable and let $g : \mathbb{R}^m \to \mathbb{R}^q$ be Lipschitz. Put

$$A = g^{-1}(\Phi(\mathbb{R}^d)).$$

The map $\Phi^{-1} \circ g : A \to \mathbb{R}^d$ is Lipschitz. Extend its coordinate functions to a Lipschitz map $h : \mathbb{R}^m \to \mathbb{R}^d$. Then

$$\begin{aligned}(\Phi_\# \mu)(g(\mathbb{R}^m)) &= \mu(\Phi^{-1}(g(\mathbb{R}^m) \cap \Phi(\mathbb{R}^d))) \\ &= \mu((\Phi^{-1} \circ g)(A)) \le \mu(h(\mathbb{R}^m)) = 0.\end{aligned}$$

Thus $\Phi_\# \mu$ is purely unrectifiable. Conversely, if $\Phi_\# \mu$ is purely unrectifiable and $h : \mathbb{R}^m \to \mathbb{R}^d$ is Lipschitz, then $\Phi \circ h$ is Lipschitz and

$$\mu(h(\mathbb{R}^m)) = (\Phi_\# \mu)((\Phi \circ h)(\mathbb{R}^m)) = 0.$$

Therefore μ is purely unrectifiable. □

Theorem — 4.2 (T-I.6.09: canonical decomposition) [F]

Every finite Borel measure μ on $\mathbb{R}^d$ has a unique decomposition

$$\mu = \mu_{\mathrm{r}}^m + \mu_{\mathrm{u}}^m, \tag{4.4}$$

where μ_r^m is m-rectifiable and μ_u^m is purely m-unrectifiable. There is a Borel countable union R of Lipschitz images such that

$$\mu_{\mathrm{r}}^m = \mu \restriction R, \qquad \mu_{\mathrm{u}}^m = \mu \restriction (\mathbb{R}^d \setminus R). \tag{4.5}$$

Proof

Let C_m be the family of all countable unions of sets of the form $f(\mathbb{R}^m)$ with $f : \mathbb{R}^m \to \mathbb{R}^d$ Lipschitz. Every such image is sigma-compact:

$$f(\mathbb{R}^m) = \bigcup_{k=1}^{\infty} f(\overline{B}(0, k)),$$

and each term is compact. Hence every member of C_m is Borel. The family C_m is closed under countable unions.

Put

$$s = \sup\{\mu(S) : S \in \mathscr{C}_m\}. \tag{4.6}$$

Choose $S_j \in C_m$ with $\mu(S_j) > s - 2^{-j}$ and set

$$R = \bigcup_{j=1}^{\infty} S_j.$$

Then $R \in C_m$, $\mu(R) \geq s$, and the definition of s gives $\mu(R) = s$. Define the two restrictions in (4.5). The first is m-rectifiable.

If $f : \mathbb{R}^m \to \mathbb{R}^d$ is Lipschitz, then $R \cup f(\mathbb{R}^m)$ belongs to C_m, so

$$\mu(R \cup f(\mathbb{R}^m)) \leq s = \mu(R).$$

It follows that

$$\mu_{\mathrm{u}}^m(f(\mathbb{R}^m)) = \mu(f(\mathbb{R}^m) \setminus R) = 0.$$

Thus μ_u^m is purely unrectifiable, establishing existence.

For uniqueness, suppose

$$\mu = \nu_{\mathrm{r}} + \nu_{\mathrm{u}}$$

is another such decomposition. Choose $S \in C_m$ carrying ν_r. Pure unrectifiability gives $\nu_u(S) = 0$, so

$$\nu_{\mathrm{r}} = \mu \restriction S. \tag{4.7}$$

Since $\mu_u^m(S) = 0$, (4.7) implies $\nu_r \leq \mu_r^m$. Conversely, $\nu_u(R) = 0$ and R carries μ_r^m, so

$$\mu_{\mathrm{r}}^m = \mu \restriction R \leq \nu_{\mathrm{r}}.$$

Therefore $\nu_r = \mu_r^m$, and subtraction from μ gives $\nu_u = \mu_u^m$. □

The Grassmannian $G(d,m)$ is the set of all m-dimensional linear subspaces of $\mathbb{R}^d$. For $V \in G(d,m)$, $x \in \mathbb{R}^d$, and $s > 0$, define the cone

$$X(x,V,s) = \{y : \operatorname{dist}(y, x+V) < s|y-x|\}. \tag{4.8}$$

For a measurable E and $x \in E$, an approximate tangent m-plane to E at x is a plane $V \in G(d,m)$ such that $\Theta^{*m}(E,x) > 0$ and, for every $s > 0$,

$$\lim_{r\downarrow 0} r^{-m}\mathscr{H}^m((E \cap \overline{B}(x,r)) \setminus X(x,V,s)) = 0. \tag{4.9}$$

Cited result — 4.3 (B-I.6.10: tangent and density characterizations) [B]

[B | currency: Classical | source: MAT22-RECT, Definition 5.4 and Theorems 5.5, 5.10–5.11; PRE87 for the measure-density theorem]
Let $E \subset \mathbb{R}^d$ be $\mathscr{H}^m$-measurable with $\mathscr{H}^m(E) < +\infty$.
(1) E is countably m-rectifiable if and only if it has an approximate tangent m-plane at $\mathscr{H}^m$-almost every $x \in E$.
(2) The following are equivalent:

$$\begin{array}{ll} \text{(a)} & E \text{ is countably } m\text{-rectifiable,} \\ \text{(b)} & \Theta^m(E,x) = 1 \text{ for } \mathscr{H}^m\text{-almost every } x \in E, \\ \text{(c)} & \Theta^m(E,x) \text{ exists for } \mathscr{H}^m\text{-almost every } x \in E. \end{array} \tag{4.10}$$

(3) If ν is a finite Radon measure and

$$0 < \lim_{r\downarrow 0} r^{-m}\nu(\overline{B}(x,r)) < +\infty \tag{4.11}$$

for ν-almost every x, then ν is m-rectifiable.
References. The frozen modern statements are Mattila [Mat22, Definition 5.4 and Theorems 5.5 and 5.10–5.11, pp. 16 and 18]. Part 3 is Preiss's density theorem [Pre87]; the density-one set implication has original provenance in Mattila [Mat75]. The converse rectifiability machinery is wholly external here. This theorem is not used by an F proof.
For $\theta \in [0,\pi)$, let P_θ denote orthogonal projection onto the line spanned by $(\cos\theta, \sin\theta)$.

Cited result — 4.4 (B-I.6.11: planar Besicovitch–Federer theorem) [B]

[B | currency: Classical | source: MAT22-RECT, Theorem 4.13]
Let $E \subset \mathbb{R}^2$ be $\mathscr{H}^1$-measurable and $\mathscr{H}^1(E) < +\infty$. Then E is purely 1-unrectifiable if and only if

$$\mathscr{H}^1(P_\theta(E)) = 0 \quad \text{for Lebesgue-almost every } \theta \in [0,\pi). \tag{4.12}$$

References. See Mattila [Mat22, Theorem 4.13, p. 11]; the original provenance is Besicovitch [Bes39]. The Besicovitch projection proof is not reproduced here.

Theorem 4.4 is a terminal bridge to Volume III and is not used by an F proof.

I.6.5 Visible parts

Let $K \subset \mathbb{R}^d$ and $e \in S^{d-1}$. Put

$$\mathbb{R}_+ e = \{te \, : \, t \geq 0\}.$$

Define

$$\mathrm{Vis}_e(K) = \{z \in K \, : \, (z + \mathbb{R}_+ e) \cap K = \{z\}\}, \tag{5.1}$$

and

$$\mathrm{Vis}_e^2(K) = \{z \in K \, : \, (z + \mathbb{R}e) \cap K = \{z\}\}. \tag{5.2}$$

The orthogonal projection onto $e^\perp$ is

$$P_{e^\perp}(z) = z - (z \cdot e)e.$$

If $a \notin K$, put

$$\mathrm{Vis}_a(K) = \{z \in K \, : \, [a, z] \cap K = \{z\}\}, \qquad \pi_a(z) = \frac{z - a}{|z - a|}. \tag{5.3}$$

Lemma — 5.1 (L-I.6.12: visible-part projection identities) [F]

If $K \subset \mathbb{R}^d$ is nonempty and compact, then

$$P_{e^\perp}(\mathrm{Vis}_e(K)) = P_{e^\perp}(K) \tag{5.4}$$

for every $e \in S^{d-1}$. Consequently

$$\dim_{\mathrm{H}} P_{e^\perp}(K) \leq \dim_{\mathrm{H}} \mathrm{Vis}_e(K) \leq \dim_{\mathrm{H}} K. \tag{5.5}$$

If $a \notin K$, then

$$\pi_a(\mathrm{Vis}_a(K)) = \pi_a(K) \tag{5.6}$$

and

$$\dim_{\mathrm{H}} \pi_a(K) \leq \dim_{\mathrm{H}} \mathrm{Vis}_a(K) \leq \dim_{\mathrm{H}} K. \tag{5.7}$$

Proof

The inclusion from left to right in (5.4) is immediate. Fix $u \in P_{e^\perp}(K)$. The fiber

$$K_u = K \cap (u + \mathbb{R}e)$$

is nonempty and compact. The continuous function $z \mapsto z \cdot e$ attains its maximum on K_u; call the maximizing point z_u. It is unique within the fiber at that coordinate, and no point of K lies on $z_u + (0, +\infty)e$. Hence $z_u \in Vis_e(K)$ and $P_{e^\perp}(z_u) = u$. This proves (5.4). Orthogonal projection is 1-Lipschitz, while $Vis_e(K) \subset K$; Hausdorff-dimension monotonicity and T-I.1.08 give (5.5).

For (5.6), fix $v \in \pi_a(K)$. The set

$$T_v = \{t > 0 \,:\, a + tv \in K\}$$

is nonempty and compact in $(0, +\infty)$: compactness of K bounds it, and $\operatorname{dist}(a, K) > 0$ bounds it away from zero. Let $t_v = \min T_v$. Then $a + t_v v$ is the first point of K on the ray, so it lies in $Vis_a(K)$. This proves the nontrivial inclusion in (5.6).

Let $\delta = \operatorname{dist}(a, K) > 0$. For $y, z \in K$, the elementary normalization estimate follows by writing $u = y - a$, $v = z - a$ and using

$$\left|\frac{u}{|u|} - \frac{v}{|v|}\right| \le \frac{|u - v|}{|u|} + |v|\left|\frac{1}{|u|} - \frac{1}{|v|}\right| = \frac{|u - v| + \big||v| - |u|\big|}{|u|} \le \frac{2|u - v|}{\delta}.$$

Therefore

$$\left|\frac{y - a}{|y - a|} - \frac{z - a}{|z - a|}\right| \le \frac{2}{\delta}|y - z|. \tag{5.8}$$

Thus π_a is Lipschitz on K. Apply the same dimension argument as for (5.5) to obtain (5.7). □

Proposition — 5.2 (P-I.6.13: model visible parts) [F]

Let $e \in S^{d-1}$.

(1) If $A \subset e^\perp$ is nonempty and compact and

$$K = A + [0, 1]e,$$

then

$$\mathrm{Vis}_e(K) = A + e, \qquad \mathrm{Vis}_e^2(K) = \emptyset. \tag{5.9}$$

(2) If $A \subset e^\perp$ is compact, $f \,:\, A \to R$ is a function, and

$$K = \{u + f(u)e \,:\, u \in A\} \tag{5.10}$$

is compact, then

$$\mathrm{Vis}_e(K) = \mathrm{Vis}_e^2(K) = K. \tag{5.11}$$

Proof

Every line $u + Re$ with $u \in A$ meets the cylinder in exactly the segment $u + [0, 1]e$. Its endpoint $u + e$ is the only point whose positive e-ray does not return to the segment. No point is the unique intersection with the full line. This proves (5.9).

For the graph in (5.10), every line parallel to e meets K in at most one point, because orthogonal projection of $u + f(u)e$ is u. Hence every point of K is bi-visible and therefore visible, proving (5.11). □

Cited result — 5.3 (B-I.6.14: current visibility bounds) [B]

[B | currency: Published/Preprint | source: ORP23-VIS, Theorem 1.2; OR26-VIS, Theorem A and Remark 5.7]

(1) Let $d \geq 2$ and let $K \subset \mathbb{R}^d$ be compact. Then

$$\dim_{\mathrm{H}} \mathrm{Vis}_e(K) \leq d - \frac{1}{50d} \tag{5.12}$$

for $\mathscr{H}^{d-1}$-almost every $e \in S^{d-1}$.

(2) In $\mathbb{R}^2$, for $\sigma \in [-1, 1]$ put

$$e_\sigma = \frac{(1, \sigma)}{\sqrt{1 + \sigma^2}}, \qquad \mathrm{Vis}_\sigma(K) = \mathrm{Vis}_{e_\sigma}(K), \qquad \mathrm{Vis}^2_\sigma(K) = \mathrm{Vis}^2_{e_\sigma}(K).$$

For every compact $K \subset \mathbb{R}^2$ and $0 \leq \beta \leq 1$,

$$\dim_{\mathrm{H}} \mathrm{Vis}^2_\sigma(K) \leq \dim_{\mathrm{H}} \mathrm{Vis}_\sigma(K) \leq 2 - \frac{\beta}{2} \tag{5.13}$$

for $\mathscr{H}^\beta$-almost every $\sigma \in [-1, 1]$. Here $\mathscr{H}^\beta$ is the diameter-normalized Hausdorff measure defined in I.1; the assertion means that the exceptional slopes have $\mathscr{H}^\beta$ measure zero. This notation covers noninteger β and $\beta = 0$, unlike the integer-only normalization $\mathscr{H}^m$ in D-I.6.01. At $\beta = 0$ the upper bound is simply the ambient bound 2 for every slope.

(3) There is a compact $K \subset \mathbb{R}^2$ satisfying

$$\dim_{\mathrm{H}} \mathrm{Vis}^2_\sigma(K) \geq \frac{3}{2} \tag{5.14}$$

for every $\sigma \in [-1, 1]$. A finite union of rotated and translated copies gives the corresponding assertion for every direction in S^1.

References. Part 1 is Orponen [Orp23, Theorem 1.2, p. 2]. Parts 2–3 are Orponen and Rutar [OR26, Theorem A, p. 2, and Remark 5.7, p. 36]. The cited version 2 supersedes the former general planar visibility conjecture: the sharp almost-every-direction upper exponent for arbitrary compact planar sets is 3/2, and a compact example attains it in every direction. The point-ray incidence proof and the lower construction are external here and are assigned to III.11.

I.6.6 Intuition and strategy

The following discussion explains the geometric ideas behind the proofs.

I.6.6.1 Why a second Hausdorff-measure symbol appears

I.1 chose the clean metric normalization in which a cover contributes $(\operatorname{diam} U)^m$. Geometric measure theory usually multiplies this measure by $\omega_m/2^m$. That scalar does nothing to null sets or dimensions, but it does make an m-plane have density one. Density theorems are sensitive to that constant, so I.6 changes the symbol rather than silently changing the meaning of $\mathscr{H}^m$.

The density quotient asks whether the mass in a small ambient ball behaves like the volume of an m-dimensional ball:

$$\frac{\mu(B(x,r))}{\omega_m r^m}.$$

For rectifiable sets this quotient settles to one at almost every point. Preiss's theorem says that, conversely, a positive finite integer-dimensional density is powerful enough to force rectifiable structure. That converse is deep because a limiting scalar has to determine a geometric tangent object.

I.6.6.2 Three strengths of porosity

A set hole is literally empty. A measure hole is allowed to carry a small fraction ε of the mass. This difference matters: a measure can have full topological support and still have very uneven mass with large epsilon-holes.

Uniform porosity requires a hole at every sufficiently small scale and every point. Mean porosity allows bad scales, but demands a positive lower density of good scales. The latter is adapted to dynamical constructions: a typical orbit may return to a hole-producing configuration often without doing so at every time.

The singularity proof is local. At a Lebesgue density point of a set, no ball occupying a fixed proportion of the ambient ball can be missing infinitely often. At a point where a measure has positive Lebesgue density, every fixed relative subball contains a definite fraction of the mass. Requiring epsilon-holes for every $0 < \varepsilon < 1$ contradicts this.

I.6.6.3 A hole is an entropy deficit

Fix a dyadic block of depth ℓ. It has

$$M = 2^{d\ell}$$

descendants. The largest possible conditional entropy is $\log M$, attained by equal mass on all descendants. If one descendant has mass tending to zero, the largest possible entropy tends to $\log(M-1)$. Thus a hole removes an information branch.

The information observed along a typical nested sequence of cubes fluctuates around its conditional entropy. The martingale-difference law in the formal notes proves that the accumulated fluctuation is sublinear. Therefore a positive density of blocks with a missing branch forces an upper local- dimension deficit, and I.4 turns that pointwise estimate into an upper packing dimension estimate.

The fixed-depth argument loses a factor $1/\ell$. Since a Euclidean α-hole is detected at a dyadic depth comparable to $|\log\alpha|$, this gives only an $\alpha^d/|\log\alpha|$ scale. Shmerkin's sharp argument partitions a porous cube nonuniformly: it isolates the hole and keeps the siblings encountered along the route to it. Entropy is then divided by the actual logarithmic contraction of each branch, removing the extra $|\log\alpha|$. G-I.6.07 cites the randomized conversion to dyadic cubes and the variable-depth argument, including its information law, optimization, and frequency bookkeeping. The fixed-block calculation alone does not supply those steps.

I.6.6.4 Rectifiable measures use a broad convention

A rectifiable set is covered, apart from m-dimensional Hausdorff null mass, by countably many Lipschitz images of $\mathbb{R}^m$. A rectifiable measure in this chapter merely has to be carried by such images. It need not be absolutely continuous with respect to $\mathscr{H}^m$; in particular, an atomic measure is m-rectifiable for every integer $1 \le m \le d$ in this chapter's range.

This broad convention makes the decomposition theorem purely measure theoretic. Maximize the mass carried by countably many Lipschitz images. A countable maximizing union carries the rectifiable part, and maximality forces the complement to give zero mass to every new Lipschitz image. The deeper density theorems add geometry: under positive finite density, the broad rectifiable carrier comes with the expected m-dimensional behavior.

Pure unrectifiability is not the negation of rectifiability point by point. It is a quantified orthogonality condition: every rectifiable competitor meets the set or measure in zero m-mass. The Besicovitch–Federer theorem turns this condition into almost-everywhere collapse of planar projections.

I.6.6.5 Visibility is a section problem, not just a projection problem

On every line parallel to a viewing direction, compactness selects a last point. This immediately gives

$$P_{e^\perp}(\mathrm{Vis}_e\, K) = P_{e^\perp} K.$$

The visible part is therefore a section of the fibers of an orthogonal projection. But a section can be a very rough graph, so its dimension need not equal the dimension of its projection. Cylinders give smooth-looking top faces; arbitrary compact graphs can be entirely bi-visible.

The same first-hit principle works from a point: on each radial direction, choose the nearest point of the compact set. Radial projection is Lipschitz away from the viewpoint, yielding the elementary dimension lower bound for the visible part.

The research picture changed in 2026. The former expectation that arbitrary planar compact sets should have almost-everywhere one-dimensional visible parts is false. Orponen–Rutar prove the sharp general upper bound $3/2$ for almost every direction and construct a compact set with bi-visible parts of dimension at least $3/2$ in every direction. The connected and Ahlfors- regular subclasses remain distinct problems and are not collapsed into the general statement.

I.6.6.6 Dependency map

The elementary route has four steps:

- Lebesgue differentiation yields porosity singularity;
- finite entropy and martingale differences yield the dyadic dimension deficit;
- Lipschitz images and a maximal carrier yield the rectifiable decomposition;
- compact fibers yield the visible-part projection identities.

Four further results are stated with sources but without full proofs here:

- Shmerkin's variable-depth porosity theorem supplies G-I.6.07;
- Mattila–Preiss tangent-density theory supplies B-I.6.10;
- Besicovitch projection theory supplies B-I.6.11;
- Orponen and Orponen–Rutar visibility theory supplies B-I.6.14.

None is used to prove an F-labelled result in this chapter.

I.6.7 Exercises

These exercises use the real-analysis prerequisites and the material already proved through this chapter. None requires a proof of a cited G or B result.

I.6.7.1 Exercise 1 — Normalization conversion

Let $E \subset \mathbb{R}^d$ be $\mathscr{H}^m$-measurable. Prove directly that

$$\Theta^m_*(E,x) = \liminf_{r\downarrow 0} \frac{\mathscr{H}^m(E\cap \overline{B}(x,r))}{(2r)^m},$$

and give the analogous formula for the upper density.

I.6.7.2 Exercise 2 — Strictness hidden in a porosity supremum

Assume $\operatorname{por}(E,x,r) \ge \alpha > 0$. Prove that for every $0<\beta<\alpha$ there is a closed ball of radius βr contained in $\overline{B}(x,r)\setminus E$. Explain why the same assertion with $\beta=\alpha$ need not follow from the definition of a supremum.

I.6.7.3 Exercise 3 — Closure of a uniformly porous set

Supply all details in the proof that if E is uniformly (α, r_0)-porous, then $\operatorname{cl}(E)$ is uniformly $(\alpha/4, r_0/2)$-porous. Your proof must explicitly create positive distance between the smaller hole and E.

I.6.7.4 Exercise 4 — Mean porosity and density points

Let $E \subset \mathbb{R}^d$ be measurable. Suppose there is a set $F\subset E$ of positive Lebesgue measure such that every $x\in F$ has infinitely many scales $r_n \downarrow 0$ with

$$\operatorname{por}(E,x,r_n) \ge \alpha.$$

Prove that this is impossible.

I.6.7.5 Exercise 5 — Why "for every epsilon" matters

Let λ be normalized Lebesgue measure on $[0,1]^d$. Show that for every sufficiently small r, every ball

$$\overline{B}(y,\alpha r) \subset \overline{B}(x,r) \subset (0,1)^d$$

satisfies

$$\frac{\lambda(\overline{B}(y,\alpha r))}{\lambda(\overline{B}(x,r))} = \alpha^d.$$

Deduce that weak mean $(\alpha, 1, \varepsilon)$-porosity has no singularity content when $\varepsilon \geq \alpha^d$.

I.6.7.6 Exercise 6 — Entropy with a light coordinate

Let $M \geq 2$ and $0 \leq \varepsilon \leq 1/M$. Maximize Shannon entropy over all probability vectors $(p_1, \ldots, p_M)$ satisfying $p_1 \leq \varepsilon$. Identify all maximizers and recover (3.15).

I.6.7.7 Exercise 7 — The dyadic missing-branch model

Fix an integer $\ell \geq 1$, put $b = 2^\ell$, and set $M = b^d$. Starting from $[0,1)^d$, at each block of ℓ dyadic levels assign zero mass to the upper-corner descendant, whose digit vector is $(b-1, \ldots, b-1)$, and mass $1/(M-1)$ to every other descendant. Repeat this rule inside every positive-mass descendant. Construct the resulting Borel probability μ, including a justification that boundary points do not invalidate the half-open cube masses. Prove that it is dyadic mean $(\ell, 1)$-porous and that

$$\overline{d}_\mu(x) \leq \frac{\log(M-1)}{\ell \log 2} \quad \text{for } \mu\text{-almost every } x.$$

I.6.7.8 Exercise 8 — A maximal carrier

Let C_m be the family used in Theorem 4.2. Prove that every member of C_m is Borel and that C_m is closed under countable unions. Then prove that the supremum in (4.6) is attained by a member of C_m.

I.6.7.9 Exercise 9 — Atomic measures and the convention for rectifiability

Let

$$\mu = \sum_{j=1}^{\infty} a_j \delta_{x_j}, \qquad a_j > 0, \quad \sum_j a_j < +\infty.$$

Prove that μ is m-rectifiable for every integer $1 \leq m \leq d$. Determine its canonical decomposition from Theorem 4.2 for every such m.

I.6.7.10 Exercise 10 — Mutual exclusivity of the two measure parts

Suppose ν is m-rectifiable, τ is purely m-unrectifiable, and $\nu \leq \tau$. Prove that $\nu = 0$. Use this to give a second uniqueness proof for the decomposition (4.4).

I.6.7.11 Exercise 11 — Directional visibility of a hypograph boundary

Let $A \subset \mathbb{R}^{d-1}$ be compact and let $f : A \to [0,1]$ have compact graph. Set

$$K = \{(u,t) : u \in A,\ 0 \leq t \leq f(u)\}.$$

For $e = (0, \ldots, 0, 1)$, determine $Vis_e(K)$ and $Vis_e^2(K)$.

I.6.7.12 Exercise 12 — Radial first hits

Let $K \subset \mathbb{R}^d$ be nonempty and compact and $a \notin K$. For $v \in \pi_a(K)$, prove that

$$t(v) = \min\{t > 0 : a + tv \in K\}$$

is well defined and that $a + t(v)v \in Vis_a(K)$. Give an example showing that the first-hit map $v \mapsto a + t(v)v$ need not be continuous, even though the projection identity (5.6) always holds.

Chapter I.7

Coding Maps and Separation Conditions

Iterating finitely many contractions turns words into points of a geometric set. We construct this coding and the associated invariant measures before comparing separation conditions. The conditions describe different ways in which symbolic pieces can meet: disjointness, separation inside an open set, exact coincidence, and quantitative near coincidence must not be confused. Stopping words allow pieces of unequal symbolic length to be compared at a common geometric scale.

The definitions below fix this chapter's conventions. Geometric intuition and proof strategy are discussed separately after the main development; exercises and their solutions provide a second route through the material.

Definitions and notation

Throughout, $d \geq 1$ is an integer and $\Lambda = \{1, \ldots, N\}$, with integer $N \geq 1$, is a nonempty finite alphabet. An alphabet is a set of symbols; a word of length n is an ordered sequence of n symbols. Examples may relabel the alphabet, without identifying distinct labels. Composition order is fixed by the displayed formula below, rather than by an informal verbal convention.

Definition — I.7.01: Words, prefixes, and compositions

For $n \geq 1$, let Λ^n be the words of length n, let

$$\Lambda^* = \bigcup_{n=0}^{\infty} \Lambda^n, \qquad \Sigma = \Lambda^{\mathbb{N}},$$

and let the empty word $\varnothing$ be the unique word of length zero. If $u = u_1 \ldots u_n$, write $|u| = n$,

$$u^- = u_1 \cdots u_{n-1} \quad (n \geq 1), \qquad \omega|n = \omega_1 \cdots \omega_n.$$

Concatenation is denoted uv. For maps $(S_i)_{(i \in \Lambda)}$, put

$$S_u = S_{u_1} \circ \cdots \circ S_{u_n}, \qquad S_\varnothing = \mathrm{Id}\,.$$

Thus $S_{(uv)} = S_u \circ S_v$. A finite word u is a **prefix** of a finite or infinite word v if $v = uw$ for some word w. Two finite words are **incomparable** if neither is a prefix of the other. The longest common prefix length of $\omega, \tau \in \Sigma$ is $|\omega \wedge \tau|$, with value $+\infty$ when $\omega = \tau$. A set of finite words is an **antichain**, or **prefix-free set**, if its distinct members are incomparable.

Definition — I.7.02: Contractive IFS, compact hyperspace, and attractor

A **finite contractive iterated function system**, abbreviated **IFS**, on $\mathbb{R}^d$ is a labelled family

$$\Phi = (S_i)_{i\in\Lambda}, \qquad |S_i(x) - S_i(y)| \le c_i|x-y|, \qquad 0 \le c_i < 1.$$

Labels are retained even when two maps coincide. Put $c = \max_i c_i$. When a word-length estimate contains c^0, this zeroth power is defined to be 1, including when $c = 0$. Let $K(\mathbb{R}^d)$ be the set of nonempty compact subsets of $\mathbb{R}^d$, equipped with the Hausdorff metric d_H from D-I.5.05. The **Hutchinson operator** is

$$\mathscr{F}_\Phi(A) = \bigcup_{i\in\Lambda} S_i(A), \qquad A \in \mathscr{K}(\mathbb{R}^d).$$

An **attractor** of Φ is a set $K \in K(\mathbb{R}^d)$ satisfying

$$\mathscr{F}_\Phi(K) = K.$$

The definition does not assert existence or uniqueness; those are T-I.7.02.

Definition — I.7.03: Symbolic metric, coding, and cylinders

The symbolic space Σ carries the metric

$$d_\Sigma(\omega,\tau) = \begin{cases} 0, & \omega = \tau, \\ 2^{-|\omega\wedge\tau|}, & \omega \ne \tau. \end{cases}$$

For $u \in \Lambda^*$, the **symbolic cylinder** is

$$[u] = \{\omega \in \Sigma \,:\, \omega||u| = u\}, \qquad [\varnothing] = \Sigma.$$

For a base point $x_0 \in \mathbb{R}^d$, the **coding candidate** is

$$\pi_\Phi(\omega) = \lim_{n\to\infty} S_{\omega|n}(x_0),$$

provided the limit exists. T-I.7.03 proves uniform existence, independence of x_0, continuity, and surjectivity onto the attractor. An **address** of $x \in K$ is a sequence $\omega \in \Sigma$ with $\pi_\Phi(\omega) = x$. Once K is known, the **geometric cylinder** associated to u is

$$K_u = S_u(K).$$

If u is nonempty, the sequence $u^\infty = uuu...$ is **periodic**, and $\pi_\Phi(u^\infty)$ is its periodic point.

Definition — I.7.04: Bernoulli law and invariant measure

A **strict probability vector** is $p = (p_i)_{(i\in\Lambda)}$ with

$$p_i > 0, \qquad \sum_{i\in\Lambda} p_i = 1.$$

For $u = u_1...u_n$, put $p_u = \prod_{(k=1)}^n p_{(u_k)}$ and $p_\varnothing = 1$. The **Bernoulli law** β_p on Σ is the product probability determined by

$$\beta_p([u]) = p_u.$$

For a Borel probability ν on $\mathbb{R}^d$, define the **measure Hutchinson operator**

$$\mathscr{M}_{\Phi,p}(\nu) = \sum_{i\in\Lambda} p_i (S_i)_\# \nu.$$

A Borel probability μ is (Φ, p)**-invariant** if

$$\mathscr{M}_{\Phi,p}(\mu) = \mu.$$

Definition — I.7.05: Similarity IFS and similarity dimension

A **similarity IFS** on $\mathbb{R}^d$ is an IFS whose maps have the form

$$S_i(x) = r_i O_i x + a_i, \qquad 0 < r_i < 1, \quad O_i \in O(d), \quad a_i \in \mathbb{R}^d.$$

Here

$$O(d) = \{O \in \mathbb{R}^{d\times d} \ : \ O^\top O = I_d\}$$

is the orthogonal group. For a linear map $A : \mathbb{R}^d \to \mathbb{R}^d$, its Euclidean operator norm is

$$\|A\|_{\mathrm{op}} = \sup_{|x|=1} |Ax|.$$

For $u = u_1...u_n$, put

$$r_u = \prod_{k=1}^n r_{u_k}, \qquad O_u = O_{u_1} \cdots O_{u_n},$$

For the empty word put $r_\emptyset = 1$ and $O_\emptyset = I_d$. Thus the linear part of S_u is $r_u O_u$. The attractor is the **self-similar set** of the labelled system, and the invariant measure from T-I.7.04 is the **self-similar measure** with weights p.
The **similarity dimension** $s(\Phi)$ is the unique $s \geq 0$ satisfying

$$\sum_{i \in \Lambda} r_i^s = 1.$$

When $N = 1$, this means $s(\Phi) = 0$. Repeated labels are counted separately, so similarity dimension is a property of the labelled IFS, not only of its attractor.

Definition — I.7.06: Strong and open set conditions

Let K be the attractor of a similarity IFS.

- Φ satisfies the **strong separation condition** (**SSC**) if

$$S_i(K) \cap S_j(K) = \emptyset \quad (i \neq j).$$

- Φ satisfies the **open set condition** (**OSC**) if there is a nonempty open set $O \subset \mathbb{R}^d$ such that

$$\bigcup_{i \in \Lambda} S_i(O) \subset O, \qquad S_i(O) \cap S_j(O) = \emptyset \quad (i \neq j).$$

Such an O is a **feasible open set**.
No boundedness or intersection condition $O \cap K \neq \emptyset$ is inserted into the OSC definition.

Definition — I.7.07: Exact overlaps and weak separation

An **exact overlap** is a pair of distinct words $u, v \in \Lambda^*$ with

$$S_u = S_v.$$

Let $\mathrm{Sim}(\mathbb{R}^d)$ be the group of similarities $T(x) = rOx + a$, where $r > 0$, $O \in O(d)$, and $a \in \mathbb{R}^d$. Fix the metric

$$d_{\mathrm{Sim}}(rOx + a, qPx + b) = |\log r - \log q| + \|O - P\|_{\mathrm{op}} + \min\{1, |a - b|\}.$$

This metric induces coefficient convergence, equivalently pointwise convergence of similarities. Define

$$\mathscr{E}_\Phi = \{S_u^{-1} \circ S_v \ : \ u, v \in \Lambda^*, \ u \neq v\}.$$

The IFS satisfies the **weak separation condition** (**WSC**) if

$$\mathrm{Id} \notin \overline{\mathscr{E}_\Phi \setminus \{\mathrm{Id}\}} \quad \text{in } (\mathrm{Sim}(\mathbb{R}^d), d_{\mathrm{Sim}}).$$

Thus exact overlaps are permitted: the identity may occur in E_Φ, but it must be isolated from its nonidentity elements.

Definition — I.7.08: Stopping antichains and the finite-type convention

For $0 < t < 1$, define the stopping antichain

$$\Lambda_t = \{v \in \Lambda^* \setminus \{\varnothing\} \,:\, r_v \le t < r_{v^-}\}.$$

A **control ball** is a closed Euclidean ball B with $K \subset \mathrm{int}(B)$. For a nonempty word u, its control-ball neighbor maps are

$$\mathcal{N}_B(u) = \{S_u^{-1} \circ S_v \,:\, v \in \Lambda_{r_u},\ S_u(B) \cap S_v(B) \neq \varnothing\}.$$

The IFS is of **finite type**, in this book's **control-ball finite-neighbor convention**, if some control ball B satisfies

$$\# \bigcup_{u \in \Lambda^* \setminus \{\varnothing\}} \mathcal{N}_B(u) < \infty.$$

This convention is frozen for the series. Other finite-type and generalized finite-type definitions in the literature are not silently identified with it.

Definition — I.7.09: Cylinder-map distance and exponential separation

Write a similarity as $T(x) = A_T x + a_T$, with $A_T = r_T O_T$. Define

$$\delta(T, U) = \begin{cases} |a_T - a_U|, & A_T = A_U, \\ +\infty, & A_T \neq A_U. \end{cases}$$

For $n \ge 1$, the **level-n cylinder separation** is

$$\Delta_n(\Phi) = \min\{\delta(S_u, S_v) \,:\, u, v \in \Lambda^n,\ u \neq v\},$$

with the minimum of the empty set equal to $+\infty$. The IFS has **exponential separation** if there are $c_0 \in (0, 1)$ and infinitely many n such that

$$\Delta_n(\Phi) \ge c_0^n.$$

It has **uniform exponential separation** if the same bound holds for every $n \ge 1$. It has **super-exponential concentration of cylinders** if

$$\lim_{n \to \infty} \frac{-\log \Delta_n(\Phi)}{n} = +\infty,$$

where $-\log 0 = +\infty$ and $-\log(+\infty) = -\infty$. Exponential separation is proved in L-I.7.11 to be exactly the failure of super-exponential concentration.

Throughout, $\Lambda = \{1, \ldots, N\}$ is a nonempty finite alphabet and $d \ge 1$.

I.7.1 Words, compact sets, and the Hutchinson operator

For $n \geq 1$, let Λ^n be the words of length n, let

$$\Lambda^* = \bigcup_{n=0}^{\infty} \Lambda^n, \qquad \Sigma = \Lambda^{\mathbb{N}}, \tag{1.1}$$

and denote the empty word by $\varnothing$. If $u = u_1 \ldots u_n$, put

$$S_u = S_{u_1} \circ \cdots \circ S_{u_n}, \qquad S_\varnothing = \mathrm{Id}\,. \tag{1.2}$$

Thus $S_{(uv)} = S_u \circ S_v$. For a nonempty word, u^- is the word obtained by deleting its last letter. The prefix of $\omega \in \Sigma$ of length n is $\omega|n$.

A contractive IFS is a labelled family $\Phi = (S_i)_{(i \in \Lambda)}$ satisfying

$$|S_i(x) - S_i(y)| \leq c_i |x - y|, \qquad 0 \leq c_i < 1. \tag{1.3}$$

Put $c = \max_i c_i$. On the nonempty compact subsets of $\mathbb{R}^d$, define

$$\mathscr{F}_\Phi(A) = \bigcup_{i \in \Lambda} S_i(A). \tag{1.4}$$

In expressions indexed by word length, the zeroth power c^0 is 1, also when $c = 0$.

Lemma — 1.1 (L-I.7.01: compact-hyperspace and Hutchinson contractions) [F]

The Hausdorff metric makes $K(\mathbb{R}^d)$ complete. Moreover,

$$d_H(\mathscr{F}_\Phi(A), \mathscr{F}_\Phi(B)) \leq c\, d_H(A, B) \tag{1.5}$$

for all $A, B \in K(\mathbb{R}^d)$.

Proof

We first prove completeness. Let (A_n) be Hausdorff Cauchy and define

$$f_n(x) = \mathrm{dist}(x, A_n).$$

For all $x \in \mathbb{R}^d$,

$$|f_n(x) - f_m(x)| \leq d_H(A_n, A_m). \tag{1.6}$$

Indeed, if $a \in A_n$, choose $b \in A_m$ arbitrarily close to the distance from a to A_m; then

$$\mathrm{dist}(x, A_m) \leq |x - a| + d_H(A_n, A_m),$$

and take the infimum over a; interchange n, m for the reverse bound. Thus (f_n) converges uniformly on $\mathbb{R}^d$ to a nonnegative 1-Lipschitz function f.

The sets A_n are uniformly bounded. To see this, fix n_0 so that $d_H(A_n, A_{(n_0)}) \leq 1$ for $n \geq n_0$; the finitely many earlier sets and the closed unit neighborhood of $A_{(n_0)}$ lie in one bounded ball.

If every A_n lies in $\overline{B}(0, R)$, then $f_n(x) \geq \operatorname{dist}(x, \overline{B}(0, R))$; passage to the limit shows that the zero set below also lies in this ball. Put

$$A = \{x \in \mathbb{R}^d \ : \ f(x) = 0\}. \tag{1.7}$$

This set is closed and bounded. It is nonempty: choose a subsequence n_k such that

$$d_H(A_{n_k}, A_{n_{k+1}}) < 2^{-k}.$$

Starting from $x_1 \in A_{(n_1)}$, choose $x_{k+1} \in A_{n_{k+1}}$ with

$$|x_{k+1} - x_k| < 2^{-k} + 2^{-2k}.$$

Then (x_k) is Cauchy, say $x_k \to x$, and

$$f(x) \leq |x - x_k| + |f(x_k) - f_{n_k}(x_k)| \longrightarrow 0.$$

Hence $x \in A$. Since A is closed and bounded in $\mathbb{R}^d$, it is compact.

We claim that

$$f(x) = \operatorname{dist}(x, A). \tag{1.8}$$

If $a \in A$, then $f_n(x) \leq |x - a| + f_n(a)$; passing to the limit and then taking the infimum gives $f(x) \leq \operatorname{dist}(x, A)$. Conversely, choose $y_n \in A_n$ with

$$|x - y_n| \leq f_n(x) + 1/n.$$

The sequence (y_n) is bounded, so some subsequence converges to y. On that subsequence,

$$f(y) \leq |y - y_n| + |f(y_n) - f_n(y_n)| \longrightarrow 0.$$

Thus $y \in A$ and $\operatorname{dist}(x, A) \leq |x - y| \leq f(x)$. This proves (1.8). Uniform convergence of f_n to $\operatorname{dist}(\cdot, A)$, together with the identity

$$d_H(C, D) = \sup_{x \in \mathbb{R}^d} |\operatorname{dist}(x, C) - \operatorname{dist}(x, D)|$$

for nonempty compact C, D, now gives $d_H(A_n, A) \to 0$. For completeness, the upper bound in this identity is (1.6). Evaluating the supremum at points of C and of D gives both directed Hausdorff distances, so gives the reverse bound as well.

For (1.5), let $\rho = d_H(A, B)$. If $x = S_i(a)$ with $a \in A$, choose $b \in B$ with $|a - b| \leq \rho + \varepsilon$. Then

$$\operatorname{dist}(x, \mathscr{F}_\Phi(B)) \leq |S_i(a) - S_i(b)| \leq c(\rho + \varepsilon).$$

Take the supremum over x, let $\varepsilon \downarrow 0$, and repeat with A, B interchanged. This proves (1.5). $\square$

Theorem — 1.2 (T-I.7.02: attractor existence, uniqueness, and convergence) [F]

There is a unique $K \in K(\mathbb{R}^d)$ satisfying

$$K = \mathscr{F}_\Phi(K) = \bigcup_{i\in\Lambda} S_i(K). \tag{1.9}$$

For every $A \in K(\mathbb{R}^d)$,

$$d_H(\mathscr{F}_\Phi^n(A), K) \le c^n d_H(A, K). \tag{1.10}$$

Proof

Lemma 1.1 makes $\mathscr{F}_\Phi$ a contraction of the complete metric space $K(\mathbb{R}^d)$, with contraction constant at most $c < 1$. The contraction mapping principle gives a unique fixed point K. Iterating (1.5) and using $\mathscr{F}_\Phi^n(K) = K$ gives (1.10). □

The compact set K in Theorem 1.2 is the **attractor** of Φ.

I.7.2 Symbolic coding and geometric cylinders

For $\omega, \tau \in \Sigma$, let $|\omega \wedge \tau|$ be their longest common prefix length and put

$$d_\Sigma(\omega, \tau) = \begin{cases} 0, & \omega = \tau, \\ 2^{-|\omega\wedge\tau|}, & \omega \ne \tau. \end{cases} \tag{2.1}$$

Common prefixes satisfy

$$|\omega \wedge \tau| \ge \min\{|\omega \wedge \zeta|, |\zeta \wedge \tau|\},$$

so $d_\Sigma(\omega, \tau) \le \max(d_\Sigma(\omega, \zeta), d_\Sigma(\zeta, \tau))$. Together with symmetry and separation of distinct words, this proves the metric axioms. Requiring a finite prefix to be fixed describes its basic open neighborhoods, so this metric gives the product topology.

For $u \in \Lambda^*$, set

$$[u] = \{\omega \in \Sigma : \omega||u| = u\}, \qquad K_u = S_u(K). \tag{2.2}$$

Theorem — 2.1 (T-I.7.03: coding theorem and cylinder identities) [F]

For every $x_0 \in \mathbb{R}^d$, the limit

$$\pi_\Phi(\omega) = \lim_{n\to\infty} S_{\omega|n}(x_0) \tag{2.3}$$

exists uniformly in ω and is independent of x_0. It defines a continuous surjection $\pi_\Phi : \Sigma \to K$ satisfying

$$\pi_\Phi(i\omega) = S_i(\pi_\Phi(\omega)), \qquad \pi_\Phi([u]) = K_u. \tag{2.4}$$

Furthermore, if $\omega \ne \tau$,

$$|\pi_\Phi(\omega) - \pi_\Phi(\tau)| \le \mathrm{diam}(K) c^{|\omega\wedge\tau|}. \tag{2.5}$$

If $c > 0$, π_Φ is Hölder with exponent

$$\gamma = \frac{\log c}{\log(1/2)} > 0; \tag{2.6}$$

if $c = 0$, it is locally constant. Finally, the fixed points of the maps S_u, $u \neq \emptyset$, are dense in K.

Proof

Fix x_0 and put

$$M = \max_{i\in\Lambda} |S_i(x_0) - x_0|.$$

For $x_n(\omega) = S_{(\omega|n)}(x_0)$, equation (1.2) gives

$$\begin{aligned} |x_{n+1}(\omega) - x_n(\omega)| &= |S_{\omega|n}(S_{\omega_{n+1}}(x_0)) - S_{\omega|n}(x_0)| \\ &\le c^n M. \end{aligned} \tag{2.7}$$

The geometric series proves uniform convergence. If y_0 is another base point, then

$$|S_{\omega|n}(x_0) - S_{\omega|n}(y_0)| \le c^n |x_0 - y_0| \longrightarrow 0, \tag{2.8}$$

so the limit is independent of the base point.

The metric space Σ is compact. Indeed, from any sequence of infinite words, successively choose subsequences on which the first, second, and later coordinates are constant; the diagonal subsequence converges in d_Σ. If two words share their first $n \ge 1$ letters, their approximating sequences agree at time n, and (2.7) gives the explicit tail estimate

$$|\pi_\Phi(\omega) - \pi_\Phi(\tau)| \le \frac{2Mc^n}{1-c}. \tag{2.9}$$

Thus π_Φ is uniformly continuous. Therefore $L = \pi_\Phi(\Sigma)$ is compact.

Passing to the limit in

$$S_{i(\omega|n)}(x_0) = S_i(S_{\omega|n}(x_0))$$

gives the first identity in (2.4). Consequently

$$L = \bigcup_{i\in\Lambda} S_i(L).$$

Uniqueness in Theorem 1.2 gives $L = K$, proving surjectivity. Repeating the first identity in (2.4) gives

$$\pi_\Phi(u\omega) = S_u(\pi_\Phi(\omega)).$$

As ω ranges over Σ, this proves the cylinder identity in (2.4).

If $\omega \neq \tau$ and their common prefix u has length n, their images lie in $S_u(K)$. Thus

$$|\pi_\Phi(\omega) - \pi_\Phi(\tau)| \le \operatorname{diam}(S_u(K)) \le c^n \operatorname{diam}(K),$$

which is (2.5). For $c > 0$, $c^n = (2^{-n})^\gamma$; this is the Hölder estimate. If $c = 0$, every S_i is constant, so $\pi_\Phi(\omega)$ depends only on ω_1 and is locally constant.

Periodic sequences are dense in Σ: if $u = \omega|n$, then u^∞ shares the first n symbols with ω. The point $\pi_\Phi(u^\infty)$ satisfies

$$S_u(\pi_\Phi(u^\infty)) = \pi_\Phi(u^\infty).$$

It is the unique fixed point of the contraction S_u. Continuity and surjectivity of π_Φ now show that these fixed points are dense in K. □

I.7.3 Invariant Bernoulli measures

Let $p = (p_i)_{(i\in\Lambda)}$ satisfy $p_i > 0$ and $\sum_i p_i = 1$. Put

$$p_u = \prod_{k=1}^{|u|} p_{u_k}, \qquad \beta_p([u]) = p_u, \tag{3.1}$$

where β_p is the Bernoulli product probability on Σ. For a Borel probability ν on $\mathbb{R}^d$, put

$$\mathscr{M}_{\Phi,p}(\nu) = \sum_{i\in\Lambda} p_i (S_i)_\# \nu. \tag{3.2}$$

Theorem — 3.1 (T-I.7.04: invariant Bernoulli measures) [F]

There is a unique invariant Borel probability, and it is

$$\mu_{\Phi,p} = (\pi_\Phi)_\# \beta_p. \tag{3.3}$$

For every initial Borel probability ν,

$$\mathscr{M}^n_{\Phi,p}(\nu) \Longrightarrow \mu_{\Phi,p}, \tag{3.4}$$

where the arrow means convergence of integrals against every bounded continuous real-valued function on $\mathbb{R}^d$, and $\operatorname{supp}(\mu_{(\Phi,p)}) = K$.

Proof

Let $iota_i : \Sigma \to \Sigma$ prepend the symbol i. The Bernoulli law obeys

$$\beta_p = \sum_{i\in\Lambda} p_i (\iota_i)_\# \beta_p. \tag{3.5}$$

It suffices to verify (3.5) on symbolic cylinders. Its value on the empty cylinder is $\sum_i p_i = 1$. If $u = jv$ is nonempty, then

$$\sum_i p_i(\iota_i)_\# \beta_p([u]) = p_j \beta_p([v]) = p_u = \beta_p([u]). \tag{3.5a}$$

The cylinders form a pi-system generating the Borel sigma-algebra. By (2.4),

$$\pi_\Phi \circ \iota_i = S_i \circ \pi_\Phi.$$

Push (3.5) forward by π_Φ to obtain

$$\mu_{\Phi,p} = \sum_i p_i (S_i)_\# \mu_{\Phi,p}, \tag{3.6}$$

so (3.3) is invariant.
Iteration of (3.2) gives

$$\mathscr{M}^n_{\Phi,p}(\nu) = \sum_{u\in\Lambda^n} p_u (S_u)_\# \nu. \tag{3.7}$$

Fix $x_0 \in \mathbb{R}^d$ and let

$$\sigma_n = \sum_{u\in\Lambda^n} p_u \delta_{S_u(x_0)}. \tag{3.8}$$

If g is bounded and continuous, then

$$\int g\, d\sigma_n = \int_\Sigma g(S_{\omega|n}(x_0))\, d\beta_p(\omega) \longrightarrow \int_\Sigma g(\pi_\Phi(\omega))\, d\beta_p(\omega) \tag{3.9}$$

by Theorem 2.1 and bounded convergence. Hence $\sigma_n \Longrightarrow \mu_{(\Phi,p)}$.
For a bounded Lipschitz g, with Lipschitz constant L, (3.7)–(3.8) give

$$\left|\int g\, d\mathscr{M}^n_{\Phi,p}(\nu) - \int g\, d\sigma_n\right| \le \int \min\{2\|g\|_\infty, Lc^n|x - x_0|\}\, d\nu(x). \tag{3.10}$$

The integrand tends pointwise to zero and is bounded by $2\|g\|_\infty$, so dominated convergence makes (3.10) tend to zero without a moment assumption on ν.

We verify the same conclusion for every bounded continuous g, which proves weak convergence directly. The set

$$C = \overline{\{S_u(x_0) \,:\, u \in \Lambda^*\}}$$

is compact: (2.7) places every point in one fixed bounded ball, and the closure is closed. For fixed x,

$$\sup_{u\in\Lambda^n} |S_u(x) - S_u(x_0)| \le c^n|x - x_0| \longrightarrow 0.$$

Uniform continuity of g on the closed unit neighborhood of C therefore shows that

$$h_n(x) = \sup_{u\in\Lambda^n} |g(S_u(x)) - g(S_u(x_0))| \longrightarrow 0.$$

Since $0 \le h_n \le 2\|g\|_\infty$, dominated convergence gives

$$\left|\iint g\, d\mathscr{M}^n_{\Phi,p}(\nu) - \int g\, d\sigma_n\right| \leq \int h_n\, d\nu \longrightarrow 0. \tag{3.11}$$

Together with (3.9), this proves (3.4).

For later uniqueness checks, bounded Lipschitz functions determine Borel probabilities on $\mathbb{R}^d$. If two probabilities have equal integrals against all such functions and F is closed, then

$$g_m(x) = \max\{0, 1 - m\,\mathrm{dist}(x, F)\}$$

is bounded Lipschitz and decreases pointwise to 1_F when F is nonempty; the empty set needs no test function. Dominated convergence shows that the probabilities agree on every closed set, hence on every Borel set by the uniqueness theorem for finite measures on a generating intersection-stable class. This also gives uniqueness of the weak limit in (3.4).

If ν is invariant, then $\mathscr{M}^n_{\Phi,p}(\nu) = \nu$ for all n; (3.4) forces $\nu = \mu_{(\Phi,p)}$, proving uniqueness.

Finally, strict positivity of all p_i gives $\mathrm{supp}(\beta_p) = \Sigma$, because every nonempty symbolic open set contains a positive-mass cylinder. The measure in (3.3) is carried by K. Conversely, if $y = \pi_\Phi(\omega) \in K$ and U is an open neighborhood of y, then $\pi_\Phi^{-1}(U)$ is an open neighborhood of ω, so it contains a positive-mass cylinder. Hence $\mu_{(\Phi,p)}(U) > 0$, and y lies in the support.
□

I.7.4 Similarities and the universal upper bound

Assume now that every map is a similarity

$$S_i(x) = r_i O_i x + a_i, \qquad 0 < r_i < 1, \qquad O_i \in O(d). \tag{4.1}$$

Here $O(d) = \{O \in \mathbb{R}^{d\times d} : O^T O = I_d\}$. The operator norm used below is $\|A\|_{(op)} = \sup_{(|x|=1)} |Ax|$.

For $u = u_1 ... u_n$, put

$$r_u = \prod_{k=1}^{n} r_{u_k}. \tag{4.2}$$

The attractor and invariant measures are respectively called self-similar sets and self-similar measures.

Lemma — 4.1 (L-I.7.05: similarity dimension and cylinder weights) [F]

There is a unique $s \geq 0$ such that

$$\sum_{i\in\Lambda} r_i^s = 1. \tag{4.3}$$

For every $n \geq 1$,

$$\sum_{u\in\Lambda^n} r_u^s = 1. \tag{4.4}$$

The vector $p_i = r_i^s$ is a strict probability vector and $p_u = r_u^s$.

Proof

If $N = 1$, equation (4.3) has the unique nonnegative solution $s = 0$. If $N \geq 2$, the function

$$g(t) = \sum_{i\in\Lambda} r_i^t$$

is continuous and strictly decreasing on $[0, +\infty)$, with $g(0) = N > 1$ and $g(t) \to 0$. The intermediate value theorem gives a unique positive solution.

Using (4.2) and distributing the finite product,

$$\sum_{u\in\Lambda^n} r_u^s = \left(\sum_{i\in\Lambda} r_i^s\right)^n = 1.$$

The statements about p follow from the same identity. □

The number in (4.3) is the **similarity dimension** $s(\Phi)$. Labels are counted even when their maps coincide.

Theorem — 4.2 (T-I.7.06: universal similarity-dimension upper bound) [F]

For every similarity IFS with attractor K,

$$\dim_{\mathrm{H}} K \leq \min\{d, s(\Phi)\}, \qquad \mathscr{H}^{s(\Phi)}(K) < +\infty. \tag{4.5}$$

Proof

Write $s = s(\Phi)$, $D = \mathrm{diam}(K)$, and $r_{\max} = \max_i r_i$. Iterating (1.9),

$$K = \bigcup_{u\in\Lambda^n} K_u. \tag{4.6}$$

Since similarities scale diameters exactly,

$$\mathrm{diam}(K_u) = r_u D \leq r_{\max}^n D. \tag{4.7}$$

Suppose first that $s > 0$. The covers (4.6) have mesh tending to zero, and Lemma 4.1 gives

$$\sum_{u\in\Lambda^n} (\mathrm{diam}\, K_u)^s = D^s \sum_{u\in\Lambda^n} r_u^s = D^s. \tag{4.8}$$

Thus $\mathscr{H}^s(K) \leq D^s < \infty$, and the Hausdorff-dimension comparison from I.1 gives $\dim_{\mathrm{H}} K \leq s$. If $s = 0$, then $N = 1$; the attractor of the single contraction is its unique fixed point, so $\mathscr{H}^0(K) = 1$ and $\dim_{\mathrm{H}} K = 0$. Finally, every subset of $\mathbb{R}^d$ has Hausdorff dimension at most d. This proves (4.5). □

I.7.5 Strong, open, weak, and finite separation

The **strong separation condition** is

$$S_i(K) \cap S_j(K) = \emptyset \quad (i \neq j). \tag{5.1}$$

The **open set condition** means that some nonempty open $O \subset \mathbb{R}^d$ satisfies

$$\bigcup_i S_i(O) \subset O, \qquad S_i(O) \cap S_j(O) = \emptyset \quad (i \neq j). \tag{5.2}$$

The open set O is called feasible.

Lemma — 5.1 (L-I.7.07: SSC and injective coding) [F]

The coding map of a similarity IFS is injective if and only if the SSC holds.

Proof

Assume SSC and suppose $\pi_\Phi(\omega) = \pi_\Phi(\tau)$. If the words differ, let q be their maximal common prefix and write

$$\omega = qi\omega', \qquad \tau = qj\tau', \qquad i \neq j.$$

By (2.4),

$$\pi_\Phi(\omega) = S_q(x), \qquad \pi_\Phi(\tau) = S_q(y)$$

for some $x \in S_i(K)$ and $y \in S_j(K)$. The similarity S_q is injective, so equality would give $x = y$, contradicting (5.1). Thus the coding is injective.

Conversely, if SSC fails, choose

$$z \in S_i(K) \cap S_j(K), \qquad i \neq j.$$

There are $x, y \in K$ with $z = S_i(x) = S_j(y)$. By surjectivity of the coding, choose $\omega, \tau \in \Sigma$ with $\pi_\Phi(\omega) = x$ and $\pi_\Phi(\tau) = y$. Then

$$z = \pi_\Phi(i\omega) = \pi_\Phi(j\tau),$$

and the two symbolic sequences are distinct. Hence the coding is not injective. □

Lemma — 5.2 (L-I.7.08: strong separation implies the OSC) [F]

Every SSC similarity IFS satisfies the OSC.

Proof

If $N = 1$, any sufficiently large open ball centered at the fixed point and mapped into itself is feasible. Assume $N \geq 2$. The first-level pieces are finitely many disjoint compact sets, so

$$\delta = \min_{i \neq j} \operatorname{dist}(S_i(K), S_j(K)) > 0. \tag{5.3}$$

Put $r_{\max} = \max_i r_i$ and choose $\varepsilon > 0$ with

$$2r_{\max}\varepsilon < \delta. \tag{5.4}$$

Let

$$O = \{x : \operatorname{dist}(x, K) < \varepsilon\}. \tag{5.5}$$

If $x \in O$, choose $y \in K$ with $|x - y| < \varepsilon$. Then

$$\operatorname{dist}(S_i(x), K) \leq |S_i(x) - S_i(y)| < r_i\varepsilon < \varepsilon,$$

so $S_i(O) \subset O$. Moreover, $S_i(O)$ lies in the open $r_i\varepsilon$-neighborhood of $S_i(K)$. Equation (5.4) makes these neighborhoods pairwise disjoint. Hence O is feasible. □

Write a similarity as $T(x) = r_T O_T x + a_T$ and equip the similarity group with

$$d_{\mathrm{Sim}}(T, U) = |\log r_T - \log r_U| + \|O_T - O_U\|_{\mathrm{op}} + \min\{1, |a_T - a_U|\}. \tag{5.6}$$

This metric gives coefficient convergence. It is also exactly the topology of pointwise convergence within the similarity group. Indeed, coefficient convergence gives uniform convergence on every bounded set. Conversely, if $T_n(x) = r_n O_n x + a_n$ converges pointwise to $T(x) = rOx + a$, then evaluation at $0, e_1, \ldots, e_d$ gives

$$a_n \longrightarrow a, \qquad r_n O_n e_j \longrightarrow rOe_j \quad (1 \leq j \leq d).$$

Thus $r_n = |r_n O_n e_1| \to r$, and then $O_n \to O$ in operator norm.

Define

$$\mathscr{E}_\Phi = \{S_u^{-1} \circ S_v : u, v \in \Lambda^*, \ u \neq v\}. \tag{5.7}$$

An **exact overlap** is an equality $S_u = S_v$ for distinct words. The **weak separation condition** is

$$\mathrm{Id} \notin \overline{\mathscr{E}_\Phi \setminus \{\mathrm{Id}\}}. \tag{5.8}$$

The closure is in (5.6). In particular, an isolated identity produced by an exact overlap is allowed.

Lemma — 5.3 (L-I.7.09: OSC excludes exact overlaps and implies WSC) [F]

Every OSC similarity IFS has no exact overlaps and satisfies (5.8).

Proof

Let O be feasible. If u, v are incomparable, then

$$S_u(O) \cap S_v(O) = \emptyset. \tag{5.9}$$

Indeed, cancel their common prefix. The remaining words begin with distinct letters, and repeated use of $S_i(O) \subset O$ places their open images inside two disjoint first-level images; applying the cancelled prefix preserves disjointness.

Suppose $S_u = S_v$ with $u \neq v$. Neither word can be a strict prefix of the other: if $v = uw$ with w nonempty, then equality of contraction ratios would give $r_u = r_u r_w$, impossible since $0 < r_w < 1$. Thus u, v are incomparable. But equality of maps makes $S_u(O) = S_v(O)$, contradicting (5.9). Hence there is no exact overlap.

Suppose next that WSC fails. There are words $u_k \neq v_k$ such that

$$g_k = S_{u_k}^{-1} \circ S_{v_k} \neq \mathrm{Id}, \qquad g_k \longrightarrow \mathrm{Id} \tag{5.10}$$

in (5.6). In particular $r_{(v_k)}/r_{(u_k)} \to 1$. For all large k, neither word is a strict prefix of the other. Indeed, if $v_k = u_k w$ with w nonempty, this ratio is at most $r_{\max} < 1$; in the reverse prefix case it is at least $r_{\max}^{-1} > 1$.

Cancel the maximal common prefix and call the remaining incomparable words a_k, b_k. Then

$$g_k = S_{a_k}^{-1} \circ S_{b_k}. \tag{5.11}$$

Choose a closed ball $C = \overline{B}(x, \rho) \subset O$ with $\rho > 0$. Coefficient convergence of similarities is uniform convergence on C, so (5.10) gives

$$g_k(x) \in C$$

for all large k. Thus $C \cap g_k(C)$ is nonempty. Applying $S_{(a_k)}$ and using (5.11),

$$S_{a_k}(C) \cap S_{b_k}(C) \neq \emptyset,$$

contrary to (5.9). Therefore WSC holds. □

For $0 < t < 1$, define the stopping antichain

$$\Lambda_t = \{v \in \Lambda^* \setminus \{\emptyset\} : r_v \leq t < r_{v^-}\}. \tag{5.12}$$

If B is a closed ball with $K \subset \mathrm{int}(B)$, put

$$\mathcal{N}_B(u) = \{S_u^{-1} \circ S_v : v \in \Lambda_{r_u},\ S_u(B) \cap S_v(B) \neq \emptyset\}. \tag{5.13}$$

The IFS is **finite type** in the control-ball convention if, for some such B,

$$\# \bigcup_{u \neq \emptyset} \mathcal{N}_B(u) < \infty. \tag{5.14}$$

Lemma — 5.4 (L-I.7.10: finite type implies WSC) [F]

Condition (5.14) implies (5.8).

Proof

Assume (5.14) and suppose WSC fails. Choose nonidentity relative maps converging to the identity as in (5.10). After passing to large k and, when necessary, replacing a relative map by its inverse, write

$$h_k = S_{u_k}^{-1} \circ S_{v_k} \longrightarrow \mathrm{Id}, \qquad r_{v_k} \leq r_{u_k}. \tag{5.15}$$

Both words are nonempty for large k, because a nonempty-word similarity or its inverse has contraction ratio outside a fixed neighborhood of 1 when paired with the empty word. Let $r_{\min} = \min_i r_i$ and $r_{\max} = \max_i r_i$. From (5.15),

$$\frac{r_{v_k}}{r_{u_k}} \longrightarrow 1.$$

For large k this ratio exceeds $r_{\max}$. If the last letter of v_k is j, then

$$r_{v_k} \leq r_{u_k} < \frac{r_{v_k}}{r_j} = r_{v_k^-}.$$

Hence $v_k \in \Lambda_{r_{u_k}}$.

Convergence in (5.15) is uniform on the control ball B, so

$$B \cap h_k(B) \neq \emptyset$$

for all large k. Applying $S_{(u_k)}$, we obtain

$$S_{u_k}(B) \cap S_{v_k}(B) \neq \emptyset.$$

Thus $h_k \in \mathcal{N}_B(u_k)$. The union in (5.14) is finite, so its nonidentity elements have positive distance from the identity. This contradicts $h_k \to Id$. □

I.7.6 Exact overlap and exponential separation

If $T(x) = A_T x + a_T$, define

$$\delta(T,U) = \begin{cases} |a_T - a_U|, & A_T = A_U, \\ +\infty, & A_T \neq A_U, \end{cases} \tag{6.1}$$

and

$$\Delta_n(\Phi) = \min_{\substack{u,v\in\Lambda^n \\ u\neq v}} \delta(S_u, S_v). \tag{6.2}$$

The empty minimum is $+\infty$. Exponential separation means that $\Delta_n \geq c_0^n$ for some $c_0 \in (0,1)$ along infinitely many levels; **uniform** exponential separation requires the inequality at every level.

Lemma — 6.1 (L-I.7.11: WSC and exponential separation) [F]

Exponential separation holds if and only if super-exponential concentration of cylinders fails. If Φ satisfies WSC and has no exact overlaps, then there is $c_0 \in (0,1)$ such that

$$\Delta_n(\Phi) \geq c_0^n \quad (n \geq 1). \tag{6.3}$$

Consequently OSC implies uniform exponential separation. If an exact overlap occurs, then $\Delta_n = 0$ for all sufficiently large n, so super-exponential concentration holds and exponential separation fails.

Proof

Put

$$q_n = \frac{-\log \Delta_n(\Phi)}{n}$$

with the conventions in D-I.7.09. If exponential separation holds with constant c_0, then $q_n \leq -\log c_0$ along infinitely many levels, so q_n does not tend to $+\infty$. Conversely, if q_n does not tend to $+\infty$, there is a finite $M > 0$ and an unbounded set of levels on which $q_n < M$. On those levels,

$$\Delta_n(\Phi) > e^{-Mn} > e^{-(M+1)n}. \tag{6.3a}$$

Thus exponential separation holds with $c_0 = e^{-(M+1)}$. This proves the first assertion.

By WSC there is $\eta \in (0,1]$ such that every nonidentity member g of E_Φ satisfies

$$d_{\mathrm{Sim}}(g, \mathrm{Id}) \geq \eta. \tag{6.4}$$

Let $u \neq v$ have length n and suppose S_u, S_v have the same linear part A. Write

$$S_u(x) = Ax + a_u, \qquad S_v(x) = Ax + a_v.$$

Absence of exact overlaps gives $a_u \neq a_v$, and

$$S_u^{-1} S_v(x) = x + A^{-1}(a_v - a_u). \tag{6.5}$$

This is a nonidentity translation. Equations (5.6) and (6.4) imply

$$|A^{-1}(a_v - a_u)| \geq \eta.$$

Since A is a similarity of ratio r_u,

$$|a_v - a_u| \geq \eta r_u \geq \eta r_{\min}^n. \tag{6.6}$$

Take $c_0 = \eta r_{\min}$ when $\eta < 1$, and $c_0 = r_{\min}$ when $\eta = 1$. Because $\eta \geq \eta^n$ for $0 < \eta \leq 1$, (6.6) implies (6.3); pairs with different linear parts contribute $+\infty$ and cause no problem. The OSC consequence follows from Lemma 5.3.

It remains to treat exact overlaps. Suppose $S_u = S_v$ for distinct finite words. Neither word is a strict prefix of the other, by the contraction-ratio argument in Lemma 5.3. The words uv and vu are distinct: if, say, $|u| \leq |v|$ and $uv = vu$, comparison of the first $|u|$ letters would make u a prefix of v. Yet

$$S_{uv} = S_u \circ S_v = S_v \circ S_u = S_{vu}. \tag{6.7}$$

Thus an exact overlap occurs at the common level $|u| + |v|$. Appending the same arbitrary word to both sides produces an exact overlap at every later level. Hence $\Delta_n = 0$ and $q_n = +\infty$ eventually. This gives super-exponential concentration and is incompatible with exponential separation. □

Proposition — 6.2 (P-I.7.12: model separation systems) [F]

The following labelled systems on R distinguish the separation conditions.

(1) The middle-third Cantor system

$$C_0(x) = x/3, \qquad C_2(x) = x/3 + 2/3$$

satisfies SSC, OSC, finite type, WSC, and uniform exponential separation.

(2) The dyadic interval system

$$D_0(x) = x/2, \qquad D_1(x) = x/2 + 1/2$$

has attractor $[0, 1]$ and satisfies OSC but not SSC.

(3) The three-label system

$$S_0(x) = x/3, \qquad S_1(x) = x/3, \qquad S_2(x) = x/3 + 2/3$$

is finite type and WSC but has an exact overlap and fails both OSC and exponential separation. Its similarity dimension is 1, whereas its attractor is the middle-third Cantor set, of dimension $\log 2/\log 3$.

Consequently SSC is strictly stronger than OSC, OSC is strictly stronger than WSC, and WSC permits exact overlaps. Here finite type always has the control-ball meaning in D-I.7.08.

Proof

For the middle-third system

$$C_0(x) = x/3, \qquad C_2(x) = x/3 + 2/3, \tag{6.8}$$

the interval $[0, 1]$ is mapped into the disjoint intervals $[0, 1/3]$ and $[2/3, 1]$. The attractor C is contained in their union, and its two first-level pieces are disjoint. Thus SSC holds, and Lemmas 5.2, 5.3, and 6.1 give OSC, WSC, and uniform exponential separation.

This system is also finite type. Use the control ball $B = [-1, 2]$. At a fixed level n, every cylinder map has the form

$$C_u(x) = 3^{-n}x + t_u, \qquad 3^n t_u \in 2\mathbb{Z}. \tag{6.9}$$

All contraction ratios are 3^{-n}, so comparable stopping words have the same length. A relative map is translation by

$$m = 3^n(t_v - t_u) \in 2\mathbb{Z}.$$

If $C_u(B)$ meets $C_v(B)$, then B meets $B + m$, so $|m| \leq 3$. Hence the only possible neighbor maps are translations by $-2, 0, 2$, a finite family.

For the dyadic interval system

$$D_0(x) = x/2, \qquad D_1(x) = x/2 + 1/2, \tag{6.10}$$

the attractor is $[0, 1]$, since this interval is invariant and the attractor is unique. The open set $(0, 1)$ is feasible, because its images are $(0, 1/2)$ and $(1/2, 1)$. Thus OSC holds. SSC fails because

$$D_0([0, 1]) \cap D_1([0, 1]) = \{1/2\}. \tag{6.11}$$

Finally consider the labelled three-map system

$$S_0(x) = x/3, \qquad S_1(x) = x/3, \qquad S_2(x) = x/3 + 2/3. \tag{6.12}$$

Deleting the repeated label leaves (6.8), so uniqueness of the attractor shows that the attractor is again C. The same calculation (6.9), now with duplicate symbolic descriptions of some geometric maps, shows finite type. Lemma 5.4 gives WSC. Since $S_0 = S_1$, there is an exact overlap; Lemma 5.3 rules out OSC and Lemma 6.1 rules out exponential separation.

The labelled similarity dimension solves $3(1/3)^s = 1$, hence equals 1. The accepted middle-third Cantor calculation from I.1 gives

$$\dim_{\mathrm{H}} C = \frac{\log 2}{\log 3} < 1. \tag{6.13}$$

The dyadic system proves that SSC is strictly stronger than OSC, while the duplicate Cantor system proves that OSC is strictly stronger than WSC and that WSC permits exact overlaps. □

I.7.7 The terminal open-set theorem

Guided result — 7.1 (G-I.7.13: Moran–Hutchinson theorem under the OSC) [G]

[G | currency: Classical/Published | omitted: the bounded-overlap input (7.5a) for comparable OSC cylinders | source: HUT81-IFS, Sections 5.1–5.3, especially Theorem 5.3(1)]

Let Φ satisfy the OSC and let $s = s(\Phi)$. Then

$$\dim_{\mathrm{H}} K = s, \qquad 0 < \mathscr{H}^s(K) < +\infty. \tag{7.1}$$

For $p_i = r_i^s$,

$$\mu_{\Phi,p} = \frac{\mathscr{H}^s \restriction K}{\mathscr{H}^s(K)}. \tag{7.2}$$

There are $0 < a \le b < \infty$ and $r_0 > 0$ such that

$$ar^s \le \mu_{\Phi,p}(\overline{B}(x,r)) \le br^s \quad (x \in K,\ 0 < r \le r_0). \tag{7.3}$$

Reference. See Hutchinson [Hut81, Sections 5.2–5.3, Theorem 5.3(1), pp. 18–20 of the author's retypeset copy].

Guided proof architecture

If $N = 1$, then $s = 0$, $K = \{x_*\}$ consists of the fixed point of the sole similarity, and (7.1)–(7.3) follow directly with $\mu_{(\Phi,p)} = \delta_{(x_*)}$ and $a = b = 1$. Assume henceforth that $N \ge 2$; then $s > 0$.

The upper half of (7.1) is already internal: Theorem 4.2 gives

$$\mathscr{H}^s(K) < +\infty, \qquad \dim_{\mathrm{H}} K \le s. \tag{7.4}$$

Use the canonical weights $p_i = r_i^s$. For a scale $0 < r < 1$, the stopping antichain Λ_r satisfies

$$r_{\min} r < r_u \le r \quad (u \in \Lambda_r), \qquad \sum_{u \in \Lambda_r} r_u^s = 1. \tag{7.5}$$

Here are the details behind the stopping operation. Choose an integer m with $r_{\max}^m \le r$. Every infinite word has a first prefix whose ratio is at most r, and that prefix has length at most m. Two such prefixes cannot be comparable, because ratios strictly decrease upon appending a letter. Thus Λ_r is finite and its symbolic cylinders partition Σ. Taking their Bernoulli probabilities proves the sum in (7.5). The ratio bounds follow from $r_u = r_{(u^-)} r_j$ and $r_{(u^-)} > r$, where j is the last letter of u.

We state the one external input precisely. For the fixed OSC system there is a finite constant M_Φ such that

$$\#\mathscr{J}(x,r) \le M_\Phi, \qquad \mathscr{J}(x,r) = \{u \in \Lambda_r \,:\, K_u \cap \overline{B}(x,r) \ne \varnothing\}, \quad x \in \mathbb{R}^d, \quad 0 < r < 1. \tag{7.5a}$$

This is the bounded-overlap input from HUT81, Theorem 5.3(1), proof parts (a)–(b), together with Section 5.2(3). Its packing argument is not reproduced here. It does not assert disjointness of the compact sets K_u. For the source application, the feasible open set may be taken bounded: choose R so large that

$$R > \frac{\max_i |a_i|}{1 - r_{\max}} \quad \text{and} \quad O \cap B(0,R) \ne \varnothing.$$

Then $S_i(B(0,R)) \subset B(0,R)$, so $O' = O \cap B(0,R)$ remains feasible. For $z \in O'$, every $S_{(\omega|n)}(z)$ lies in O'; the coding theorem gives $K \subset \mathrm{cl}(O')$. Hence $K_u \subset \mathrm{cl}(S_u(O'))$. The source packing bound for closures of these disjoint stopping open sets therefore gives exactly (7.5a), with the source's ambient dimension equal to d and its minimum ratio equal to $r_{\min}$. Open-ball bounds also give closed-ball bounds by a fixed radius enlargement, changing only the constant.

All further deductions can now be made explicitly. If $\pi_\Phi(\omega) \in \overline{B}(x,r)$, its stopping prefix belongs to $\mathcal{J}(x,r)$. Thus

$$\begin{aligned} \mu_{\Phi,p}(\overline{B}(x,r)) &\le \sum_{u\in\mathcal{J}(x,r)} \beta_p([u]) \\ &= \sum_{u\in\mathcal{J}(x,r)} r_u^s \le M_\Phi r^s \qquad (x \in \mathbb{R}^d,\ 0 < r < 1). \end{aligned} \tag{7.6}$$

In particular, no equality $\mu_{(\Phi,p)}(K_u) = r_u^s$ has been assumed. The mass-distribution principle yields $\mathscr{H}^s(K) > 0$ and $\dim_{\mathrm{H}} K \ge s$. Together with (7.4), this proves (7.1). Since $s > 0$, a singleton has zero $\mathscr{H}^s$ measure, so we now know $D = \mathrm{diam}(K) > 0$.

For $0 < \rho < D$, take $t = \rho/D$ and let $u \in \Lambda_t$ be the stopping prefix of an address of $x \in K$. Both $x \in K_u$ and $\mathrm{diam}(K_u) \le \rho$ hold, so $K_u \subset \overline{B}(x,\rho)$. Consequently

$$\mu_{\Phi,p}(\overline{B}(x,\rho)) \ge \mu_{\Phi,p}(K_u) \ge \beta_p([u]) = r_u^s > \left(\frac{r_{\min}}{D}\right)^s \rho^s. \tag{7.7}$$

Equations (7.6)–(7.7) prove (7.3) on $0 < \rho \le r_0$ with $r_0 = \min(1,D)/2$. Enlarging the upper constant if necessary ensures $a \le b$.

For the measure identity, put $\lambda = \mathscr{H}^s \restriction K$. Similarity scaling from I.1 and the canonical weights give

$$\sum_i \mathscr{H}^s(K_i) = \sum_i r_i^s \mathscr{H}^s(K) = \mathscr{H}^s(K).$$

Since $K = \cup_i K_i$ and all these measures are finite, the nonnegative function $\sum_i 1_{(K_i)} - 1_K$ has integral zero with respect to $\mathscr{H}^s$. It follows that $\mathscr{H}^s(K_i \cap K_j) = 0$ for $i \ne j$. For every Borel set E,

$$\begin{aligned} \lambda(E) &= \sum_i \mathscr{H}^s(E \cap K_i) \\ &= \sum_i r_i^s \mathscr{H}^s(S_i^{-1}(E) \cap K) = \sum_i r_i^s (S_i)_\# \lambda(E). \end{aligned}$$

Normalize by the positive finite number $\mathscr{H}^s(K)$ and apply the invariant probability uniqueness in Theorem 3.1 to obtain (7.2). Only the input (7.5a) remains unproved here; its consequences above are proved relative to that input. The result remains G, and no F result uses it.

I.7.8 Intuition and strategy

The following discussion explains the geometric ideas behind the proofs.

I.7.8.1 One contraction principle, three incarnations

An ordinary contraction repeatedly pulls points toward one fixed point. An IFS applies several contractions at once. On compact sets this becomes the Hutchinson operator

$$A \longmapsto \bigcup_i S_i(A),$$

and its fixed point is the attractor. On probability measures the analogous operator averages the pushed-forward measures,

$$\nu \longmapsto \sum_i p_i (S_i)_{\#}\nu.$$

The compact-set construction records where the geometry lives; the measure construction records how often the symbolic process visits each part. Both are controlled by the same exponential contraction of finite compositions.

I.7.8.2 Symbolic space is a universal address book

An infinite word tells the system which map to apply at every stage. Two words with a long common prefix have passed through the same long chain of copies, so their images are close. This is the geometric meaning of the symbolic metric and the Hölder estimate for the coding map.

The coding map is always onto, but it need not be one-to-one. Nonuniqueness of addresses has several strengths. Two first-level pieces may merely touch, as in the dyadic interval. More complicated cylinders may overlap without being identical. At the strongest algebraic extreme, two distinct words may define exactly the same map.

Periodic words correspond to fixed points of finite compositions. Their density is the symbolic analogue of approximating an infinite instruction list by a repeating finite block.

I.7.8.3 Labels and geometry must not be conflated

The IFS is a labelled family. If two labels carry the same map, the attractor does not notice the duplicate, but the symbolic space and similarity equation do. Thus

$$\sum_i r_i^s = 1$$

can predict a dimension larger than the actual Hausdorff dimension. The universal upper bound remains correct because it counts every labelled cylinder in a cover; overlaps only make that cover redundant.

This distinction is the elementary prototype of dimension drop. Later volumes ask when redundancy is forced by exact algebraic coincidences and when it can arise from subtler concentration of distinct cylinders.

I.7.8.4 The separation ladder measures different phenomena

The strong separation condition says the compact first-level pieces are disjoint. It is a statement about the attractor itself. The open set condition asks instead for disjoint images of an auxiliary open set. Pieces of the attractor may then meet on their boundaries.

Weak separation is formulated in the space of similarity maps. It forbids distinct relative cylinder maps from accumulating at the identity, but it allows the identity itself to occur as an isolated exact overlap. This is why weak separation is not simply a diluted form of disjointness.

Finite type makes the comparable-scale overlap patterns genuinely finite. The control ball is a bookkeeping device: normalize a cylinder back to unit scale and list the neighboring cylinders it can see. A finite list cannot contain nonidentity maps converging to the identity, which explains the route

$$\text{finite type} \longrightarrow \text{weak separation.}$$

The chapter's definition is deliberately pinned to one control-ball convention. Other definitions called "finite type" in the literature may require additional hypotheses before they become equivalent.

I.7.8.5 Exponential separation prevents near coincidences

At a fixed level, two cylinder maps can be compared by translation only when their linear parts agree. This is the comparison relevant to Hochman's separation scale. Weak separation says that, after normalizing by one cylinder, a nonzero translation cannot be arbitrarily small. Scaling back to level n loses at most the smallest contraction ratio to the power n, so a uniform exponential lower bound follows when exact overlaps are absent.

An exact overlap at unequal levels is not harmless. Composing the equal maps in opposite orders creates an equal-map pair at one common level, and then common suffixes propagate the coincidence to every later level. This is why exact overlap and exponential separation are incompatible.

Super-exponential concentration is a limiting-rate condition, whereas exponential separation asks for a lower bound on infinitely many levels. Negating convergence of $-\log \Delta_n/n$ to infinity produces precisely such an infinite subsequence. The formal notes record this quantifier equivalence explicitly.

I.7.8.6 What the open set theorem adds

The easy half of the Moran–Hutchinson theorem is a cylinder cover: at level n, the sum of the s-powers of all cylinder diameters is constant. The difficult half prevents too many comparable cylinders from crowding one ball. A fixed ball inside a bounded feasible open set supplies a smaller ball inside every image of that open set. These are auxiliary open pieces, not balls inside the fractal K_u. Their disjointness and Euclidean packing control local multiplicity of the geometric cylinders.

That bounded-overlap mechanism gives both positive Hausdorff measure and an upper r^s mass bound. The opposite mass bound comes from following one address down to scale r. Together they make the canonical self-similar measure Ahlfors regular. The exact bounded-overlap block is cited to Hutchinson as (7.5a); the mass estimates and measure identity are then deduced explicitly. Positivity of the attractor diameter is established before dividing by it.

I.7.8.7 Dependency map

The fully proved route is

Hausdorff hyperspace completeness
$\longrightarrow$ attractor fixed point
$\longrightarrow$ symbolic coding and cylinders
$\longrightarrow$ invariant Bernoulli measure
similarity weights $\longrightarrow$ universal dimension upper bound
SSC $\longrightarrow$ OSC $\longrightarrow$ WSC
finite type $\longrightarrow$ WSC
WSC + no exact overlap $\longrightarrow$ uniform exponential separation

The one cited classical input is

OSC bounded overlap and lower measure estimate $\longrightarrow$ G-I.7.13.

No F-labelled result depends on that bridge or on material from I.8.

I.7.9 Exercises

These exercises use the permitted real-analysis and dynamical-systems prerequisites and the results already proved through this chapter. None requires the cited bounded-overlap lemma in the final open-set theorem.

I.7.9.1 Exercise 1 — Word and cylinder calculus

Let $u, v \in \Lambda^*$. Prove

$$K_{uv} = S_u(K_v) \subset K_u$$

and, for a similarity IFS,

$$\operatorname{diam}(K_{uv}) = r_u r_v \operatorname{diam}(K).$$

Show also that if neither word is a prefix of the other and the OSC holds, then $S_u(O) \cap S_v(O) = \emptyset$ for every feasible open set O.

I.7.9.2 Exercise 2 — An a posteriori attractor estimate

Let Φ be a contractive IFS with maximal contraction constant c. Prove that every nonempty compact A satisfies

$$d_H(A, K) \le \frac{d_H(A, \mathscr{F}_\Phi(A))}{1-c}.$$

Deduce

$$d_H(\mathscr{F}_\Phi^n(A), K) \le \frac{c^n}{1-c} d_H(A, \mathscr{F}_\Phi(A)).$$

I.7.9.3 Exercise 3 — The exact middle-third coding formula

Use the alphabet $\{0, 2\}$ and maps

$$C_j(x) = \frac{x + j}{3} \qquad (j \in \{0, 2\}).$$

Prove that

$$\pi_C(\omega) = \sum_{k=1}^{\infty} \frac{\omega_k}{3^k}.$$

Derive the Hölder estimate directly from the series, and explain why the image is exactly the set of numbers in $[0, 1]$ admitting a ternary expansion using only the digits 0 and 2.

I.7.9.4 Exercise 4 — Periodic points of affine cylinders

Let u be a nonempty finite word. Suppose the similarity composition S_u has affine form

$$S_u(x) = A_u x + a_u.$$

Prove that $I_d - A_u$ is invertible and that the point coded by u^∞ is

$$\pi_\Phi(u^\infty) = (I_d - A_u)^{-1} a_u.$$

I.7.9.5 Exercise 5 — Quantitative convergence for measures with a first moment

Fix $x_0 \in \mathbb{R}^d$, put

$$M = \max_i |S_i(x_0) - x_0|,$$

and suppose $\int |x - x_0| d\nu(x) < \infty$. If g is bounded and Lipschitz with Lipschitz constant L, prove

$$\left| \int g \, d\mathcal{M}^n_{\Phi,p}(\nu) - \int g \, d\mu_{\Phi,p} \right| \le Lc^n \left(\int |x - x_0| \, d\nu(x) + \frac{M}{1 - c} \right).$$

I.7.9.6 Exercise 6 — What changes when some weights vanish

Let $p_i \ge 0$, $\sum_i p_i = 1$, and put

$$\Lambda_+ = \{i \,:\, p_i > 0\}.$$

Let K_+ be the attractor of the sub-IFS $(S_i)_{(i \in \Lambda_+)}$. Prove that the measure operator still has a unique invariant probability and that its support is exactly K_+. Give an example in which K_+ is a proper subset of the attractor K of the full labelled IFS.

I.7.9.7 Exercise 7 — Arbitrarily large labelled dimension drop

Fix $0 < r < 1$ and let all $N \geq 2$ labelled maps on R equal $S_i(x) = rx$. Determine the attractor, its Hausdorff dimension, and the labelled similarity dimension. Show that the difference between the last two quantities can be made arbitrarily large by increasing N.

I.7.9.8 Exercise 8 — All addresses in the dyadic interval

For

$$D_0(x) = x/2, \qquad D_1(x) = x/2 + 1/2,$$

prove

$$\pi_D(\omega) = \sum_{k=1}^{\infty} \frac{\omega_k}{2^k}.$$

Here a **dyadic point** means a number $k/2^m$, with k an integer and m a nonnegative integer. An **address** of x is a word $\omega \in \Sigma$ with $\pi_D(\omega) = x$. Show that every non-dyadic point of $(0, 1)$ has one address, every dyadic point of $(0, 1)$ has exactly two addresses, and each endpoint has one address.

I.7.9.9 Exercise 9 — Quantitative feasible neighborhoods under SSC

Assume $N \geq 2$ and SSC. Let

$$\delta = \min_{i \neq j} \operatorname{dist}(S_i(K), S_j(K)), \qquad r_{\max} = \max_i r_i.$$

Prove that for every

$$0 < \varepsilon < \frac{\delta}{2r_{\max}},$$

the open ε-neighborhood of K is feasible for the OSC.

I.7.9.10 Exercise 10 — Similarity conjugacy invariance

Let $Q(x) = qRx + b$ be an invertible similarity and define

$$\tilde{S}_i = Q \circ S_i \circ Q^{-1}.$$

Prove that the conjugated IFS has attractor $Q(K)$ and coding map $Q \circ \pi_\Phi$. Prove that SSC, OSC, exact overlap, WSC, exponential separation, and uniform exponential separation are each invariant under this conjugacy. Show more precisely that

$$\Delta_n(\widetilde{\Phi}) = q\Delta_n(\Phi).$$

I.7.9.11 Exercise 11 — Finite type for the dyadic interval

For the dyadic IFS in Exercise 8, use the control ball $B = [-1, 2]$ and prove directly that

$$\bigcup_{u \neq \emptyset} \mathcal{N}_B(u)$$

consists of translations by integers in a subset of $\{-3, -2, -1, 0, 1, 2, 3\}$. Deduce finite type and WSC.

I.7.9.12 Exercise 12 — Propagation of an arbitrary exact overlap

Suppose $S_u = S_v$ for distinct finite words, which may have different lengths. Prove that neither word is a prefix of the other. Prove that uv and vu are distinct words of the same length but define the same map. Finally, show that for every $n \geq |u| + |v|$ there are distinct words $a_n, b_n \in \Lambda^n$ with $S_{(a_n)} = S_{(b_n)}$.

I.7.9.13 Exercise 13 — The separation-rate quantifiers

Let (δ_n) take values in $[0, +\infty]$ and use the logarithmic conventions of D-I.7.09. Prove directly that the following are equivalent:

1. there is $c \in (0, 1)$ such that $\delta_n \geq c^n$ for infinitely many n;
2. $(-\log \delta_n)/n$ does not tend to $+\infty$.

Construct a sequence satisfying these conditions but for which no exponential lower bound holds at every level.

I.7.9.14 Exercise 14 — Stopping antichains and canonical weights

Let $s = s(\Phi)$ and $p_i = r_i^s$. Prove that Λ_t is a finite prefix-free set and that every $\omega \in \Sigma$ has exactly one prefix in Λ_t. Deduce, without using G-I.7.13, that

$$\sum_{u \in \Lambda_t} r_u^s = 1.$$

Chapter I.8

Self-Similar Sets and Measures

For self-similar constructions, the contraction ratios predict a dimension, and the probabilities predict the dimension of a measure. Separation determines when these predictions are valid. This chapter combines coding, typical word frequencies, and covering estimates to prove dimension formulas in the stated regimes. It also explains why the geometry of coding fibers and the distinction between a set and a measure become important when overlaps are present.

The definitions below fix this chapter's conventions. Geometric intuition and proof strategy are discussed separately after the main development; exercises and their solutions provide a second route through the material.

Definitions and notation

Throughout, $\Phi = (S_i)_{(i\in\Lambda)}$ is a labelled similarity IFS on $\mathbb{R}^d$, $p = (p_i)_{(i\in\Lambda)}$ is a strict probability vector, K is the attractor, and $\mu = \mu_{(\Phi,p)}$ is the self-similar measure from T-I.7.04. All logarithms are natural.

Definition — I.8.01: Entropy, Lyapunov exponent, and natural dimension

The **symbol information** and **contraction information** are

$$I_p(i) = -\log p_i, \qquad L_\Phi(i) = -\log r_i.$$

The **Shannon entropy** of p and the **Lyapunov exponent** of (Φ, p) are

$$h(p) = -\sum_{i\in\Lambda} p_i \log p_i, \qquad \chi(\Phi, p) = -\sum_{i\in\Lambda} p_i \log r_i.$$

Since every $0 < r_i < 1$, one has $\chi(\Phi, p) > 0$. The **similarity dimension of the self-similar measure** is

$$\mathrm{sdim}(\mu_{\Phi,p}) = \frac{h(p)}{\chi(\Phi, p)},$$

and its **natural dimension** is

$$\dim_{\mathrm{nat}}(\mu_{\Phi,p}) = \min\left\{d, \frac{h(p)}{\chi(\Phi, p)}\right\}.$$

These are labelled-system quantities. Neither definition asserts equality with a geometric dimension of μ.

Definition — I.8.02: Frequencies, information sums, and typical sequences

For $\omega \in \Sigma$, $i \in \Lambda$, and $n \geq 1$, define the empirical frequency

$$f_{n,i}(\omega) = \frac{1}{n}\#\{1 \leq k \leq n : \omega_k = i\}.$$

The length-n **information sum** and **contraction sum** are

$$I_n(\omega) = -\log p_{\omega|n} = \sum_{k=1}^{n} I_p(\omega_k),$$

$$L_n(\omega) = -\log r_{\omega|n} = \sum_{k=1}^{n} L_\Phi(\omega_k).$$

The **Bernoulli typical set** is

$$\Omega_p = \{\omega \in \Sigma : f_{n,i}(\omega) \to p_i \text{ for every } i \in \Lambda\}.$$

Definition — I.8.03: Contraction-weighted prefix distance

An **ultrametric** is a metric ρ satisfying the strong triangle inequality

$$\rho(x, z) \leq \max\{\rho(x, y), \rho(y, z)\}.$$

Put $r_\varnothing = 1$. The **contraction-weighted prefix distance** on Σ is

$$\rho_\Phi(\omega, \tau) = \begin{cases} 0, & \omega = \tau, \\ r_{\omega\wedge\tau}, & \omega \neq \tau, \end{cases}$$

where $\omega \wedge \tau$ is their longest common prefix. L-I.8.04 proves that this is an ultrametric, that coding is Lipschitz from this metric, and that it is bi-Lipschitz under SSC.

Definition — I.8.04: Coding fibers and zero-dimensional fibers

For $x \in K$, the **coding fiber** of x is

$$\mathrm{Fib}_\Phi(x) = \pi_\Phi^{-1}(\{x\}) \subset \Sigma.$$

The IFS has the **zero-dimensional-fiber property** if

$$\dim_{\mathrm{H}}^{d_\Sigma} \mathrm{Fib}_\Phi(x) = 0 \qquad (x \in K),$$

where the superscript records that Hausdorff dimension is computed in the symbolic metric d_Σ of D-I.7.03.

Definition — I.8.05: Dimension drop and full-dimensional measures

In this chapter, unadorned Hausdorff and packing dimensions of a probability mean their upper, full-mass versions from D-I.4.03:

$$\dim_{\mathrm{H}} \mu := \overline{\dim}_{\mathrm{H}}\mu, \qquad \dim_{\mathrm{P}} \mu := \overline{\dim}_{\mathrm{P}}\mu.$$

For an exact-dimensional measure all four measure dimensions agree, by Chapter I.4. The convention above also makes statements meaningful before exact dimensionality has been established.
The self-similar measure has **dimension drop from its natural dimension** if

$$\dim_{\mathrm{H}} \mu_{\Phi,p} < \dim_{\mathrm{nat}}(\mu_{\Phi,p}).$$

It is **full-dimensional in its attractor** if

$$\dim_{\mathrm{H}} \mu_{\Phi,p} = \dim_{\mathrm{H}} K.$$

These properties are distinct. In an overlapping labelled system a measure may be full-dimensional in K while still having dimension drop from the labelled natural dimension.

Definition — I.8.06: Ahlfors-regular measures and sets

Let E be a nonempty compact metric set and let ν be a Borel probability with $\mathrm{supp}(\nu) = E$. For $s \geq 0$, the measure ν is s-**Ahlfors regular on** E if there are constants $0 < a \leq b < \infty$ and $r_0 > 0$ such that

$$ar^s \leq \nu(\overline{B}(x,r)) \leq br^s \quad (x \in E,\ 0 < r \leq r_0).$$

The set E is s-**Ahlfors regular** if it supports an s-Ahlfors-regular Borel probability. This is the probability-normalized convention used in the chapter.

Throughout, $\Phi = (S_i)_{(i\in\Lambda)}$ is a labelled similarity IFS on $\mathbb{R}^d$, p is a strict probability vector, and

$$\mu = \mu_{\Phi,p} = (\pi_\Phi)_\# \beta_p. \tag{0.1}$$

All logarithms are natural. Put

$$r_{\min} = \min_i r_i, \qquad r_{\max} = \max_i r_i. \tag{0.2}$$

I.8.1 Bernoulli information and contraction averages

Define

$$h = h(p) = -\sum_i p_i \log p_i, \qquad \chi = \chi(\Phi, p) = -\sum_i p_i \log r_i > 0. \tag{1.1}$$

For $\omega \in \Sigma$, let

$$I_n(\omega) = -\log p_{\omega|n}, \qquad L_n(\omega) = -\log r_{\omega|n}. \tag{1.2}$$

Lemma — 1.1 (L-I.8.01: finite-alphabet Bernoulli strong law) [F]

One has $\beta_p(\Omega_p) = 1$. For β_p-almost every ω,

$$\frac{I_n(\omega)}{n} \longrightarrow h, \qquad \frac{L_n(\omega)}{n} \longrightarrow \chi, \tag{1.3}$$

and therefore

$$\frac{\log p_{\omega|n}}{\log r_{\omega|n}} \longrightarrow \frac{h}{\chi}. \tag{1.4}$$

Proof

We first prove the needed strong law. Let $f : \Lambda \to R$ be arbitrary, put

$$m_f = \sum_i p_i f(i), \qquad X_k(\omega) = f(\omega_k) - m_f, \qquad S_n = \sum_{k=1}^n X_k. \tag{1.5}$$

The variables X_k are independent and identically distributed (iid), centered, meaning $\mathbb{E}(X_k) = 0$, and bounded. On expanding the fourth power, independence and centering make every term vanish unless its indices occur in one block of size four or two blocks of size two. Hence

$$\mathbb{E}_{\beta_p}(S_n^4) = \sum_{k=1}^n \mathbb{E}(X_k^4) + 6 \sum_{1 \le j < k \le n} \mathbb{E}(X_j^2)\mathbb{E}(X_k^2) \le C_f n^2 \tag{1.6}$$

for a finite constant C_f. Markov's inequality gives, for every $\varepsilon > 0$,

$$\beta_p\{|S_n| > \varepsilon n\} \le \frac{C_f}{\varepsilon^4 n^2}. \tag{1.7}$$

The right side is summable in n. Borel–Cantelli, first with $\varepsilon = 1$, then simultaneously with $\varepsilon = 1/m$, $m \ge 1$, yields

$$\frac{1}{n}\sum_{k=1}^n f(\omega_k) \longrightarrow \sum_i p_i f(i) \quad \text{for } \beta_p\text{-almost every } \omega. \tag{1.8}$$

Apply (1.8) to the finitely many indicator functions $f = 1_{\{i\}}$. Their common probability-one set is exactly Ω_p, so $\beta_p(\Omega_p) = 1$. Apply it also to

$$f(i) = -\log p_i \quad \text{and} \quad f(i) = -\log r_i.$$

These functions are bounded because the alphabet is finite and every p_i and r_i is positive. Their sums are I_n and L_n, and their expectations are h and χ; this proves (1.3). Since $\chi > 0$, division of the two limits gives (1.4). □

Lemma — 1.2 (L-I.8.02: entropy/Lyapunov comparison) [F]

If $s = s(\Phi)$, then

$$\frac{h}{\chi} \le s. \tag{1.9}$$

Equality holds if and only if $p_i = r_i^s$ for every i.

Proof

Put $q_i = r_i^s$. Lemma 4.1 of I.7 gives $q_i > 0$ and $\sum_i q_i = 1$. For $x > 0$, the inequality $\log x \le x - 1$ gives

$$\sum_i p_i \log \frac{q_i}{p_i} \le \sum_i p_i \left(\frac{q_i}{p_i} - 1 \right) = \sum_i q_i - \sum_i p_i = 0. \tag{1.10}$$

Therefore

$$0 \le \sum_i p_i \log \frac{p_i}{q_i} = \sum_i p_i \log p_i - s \sum_i p_i \log r_i = -h + s\chi. \tag{1.11}$$

Division by $\chi > 0$ proves (1.9). Equality in (1.10) requires $q_i/p_i = 1$ for every i, because every $p_i > 0$ and equality in $\log x \le x - 1$ occurs only at $x = 1$. This is exactly the asserted equality case. □

Theorem — 1.3 (T-I.8.03: universal natural-dimension upper bound) [F]

Without any separation condition,

$$\dim_{\mathrm{H}} \mu_{\Phi,p} \le \min \left\{ d, \frac{h}{\chi} \right\}. \tag{1.12}$$

Proof

If $\operatorname{diam}(K) = 0$, the measure is a point mass and the claim is immediate. Assume $D = \operatorname{diam}(K) > 0$. Fix $0 < \varepsilon < \chi$. For $m \ge 1$, let

$$E_{m,\varepsilon} = \{ \omega : I_n(\omega) \le n(h + \varepsilon), \quad L_n(\omega) \ge n(\chi - \varepsilon) \text{ for every } n \ge m \}. \tag{1.13}$$

Lemma 1.1 gives

$$\beta_p \left(\bigcup_{m=1}^{\infty} E_{m,\varepsilon} \right) = 1. \tag{1.14}$$

Each $E_{(m,\varepsilon)}$ is an intersection of finite unions of symbolic cylinders, hence is closed in the compact space Σ. Its coding image is therefore compact, and the union of these images is a Borel carrier.

For fixed m and $n \geq m$, let W_n be the words $u \in \Lambda^n$ for which $[u] \cap E_{(m,\varepsilon)}$ is nonempty. From the information inequality in (1.13),

$$p_u \geq e^{-n(h+\varepsilon)} \quad (u \in W_n).$$

The level-n cylinders are disjoint in symbolic space and their Bernoulli masses sum to one, so

$$\#W_n \leq e^{n(h+\varepsilon)}. \tag{1.15}$$

The contraction inequality in (1.13) and exact similarity scaling give

$$\operatorname{diam}(K_u) = Dr_u \leq De^{-n(\chi-\varepsilon)} \quad (u \in W_n). \tag{1.16}$$

For every

$$t > \frac{h+\varepsilon}{\chi-\varepsilon},$$

the cylinders $(K_u)_{(u\in W_n)}$ cover $\pi_\Phi(E_{(m,\varepsilon)})$ and satisfy

$$\sum_{u\in W_n} (\operatorname{diam} K_u)^t \leq D^t \exp\{n(h+\varepsilon) - nt(\chi-\varepsilon)\} \longrightarrow 0. \tag{1.17}$$

Thus

$$\dim_{\mathrm{H}} \pi_\Phi(E_{m,\varepsilon}) \leq \frac{h+\varepsilon}{\chi-\varepsilon}.$$

The union over m is a full μ-measure carrier by (0.1) and (1.14), so the same bound holds for $\dim_{\mathrm{H}} \mu$. Letting $\varepsilon \downarrow 0$ gives $\dim_{\mathrm{H}} \mu \leq h/\chi$. The ambient bound $\dim_{\mathrm{H}} \mu \leq d$ follows because μ is carried by $\mathbb{R}^d$. This proves (1.12). □

I.8.2 Weighted symbolic geometry under strong separation

For distinct ω, τ, put

$$\rho_\Phi(\omega,\tau) = r_{\omega\wedge\tau}, \qquad \rho_\Phi(\omega,\omega) = 0, \tag{2.1}$$

with $r_\varnothing = 1$.

Lemma — 2.1 (L-I.8.04: weighted symbolic metric and SSC coding) [F]

The function ρ_Φ is an ultrametric. For distinct ω, τ,

$$|\pi_\Phi(\omega) - \pi_\Phi(\tau)| \leq D\rho_\Phi(\omega,\tau), \qquad D = \operatorname{diam}(K). \tag{2.2}$$

If $N \geq 2$ and SSC holds, then

$$\delta = \min_{i \neq j} \operatorname{dist}(S_i(K), S_j(K)) > 0 \tag{2.3}$$

and

$$\delta \rho_\Phi(\omega, \tau) \leq |\pi_\Phi(\omega) - \pi_\Phi(\tau)| \leq D \rho_\Phi(\omega, \tau). \tag{2.4}$$

Consequently $\pi_\Phi : (\Sigma, \rho_\Phi) \to K$ is bi-Lipschitz, and

$$\pi_\Phi^{-1}(K_u) = [u], \qquad \mu(K_u) = p_u. \tag{2.5}$$

For $N = 1$, Σ and K are singletons and (2.5) remains valid.

Proof

Positivity off the diagonal and symmetry of ρ_Φ are immediate. The strong triangle inequality is immediate if two of the three sequences coincide. Otherwise, the two words $\omega \wedge \eta$ and $\eta \wedge \tau$ are both prefixes of η, so one is a prefix of the other. Their shorter one is a prefix of both ω and τ. Since contraction products strictly decrease when a word is extended,

$$\rho_\Phi(\omega, \tau) \leq \max\{\rho_\Phi(\omega, \eta), \rho_\Phi(\eta, \tau)\}. \tag{2.6}$$

Thus ρ_Φ is an ultrametric. It gives the same Borel sets as the original symbolic metric: for distinct sequences whose common prefix has length n,

$$r_{\min}^n \leq \rho_\Phi(\omega, \tau) \leq r_{\max}^n, \qquad d_\Sigma(\omega, \tau) = 2^{-n}. \tag{2.6a}$$

These bounds show that convergence in either metric is equivalent to the common-prefix length tending to infinity. The topologies therefore agree, so the Bernoulli law is a Borel probability for the weighted metric as well.

Let $u = \omega \wedge \tau$. Both coded points lie in $S_u(K)$, whose diameter is $r_u D$; this proves (2.2). Under SSC, write

$$\omega = ui\omega', \qquad \tau = uj\tau', \qquad i \neq j.$$

The two images lie respectively in $S_u(S_i(K))$ and $S_u(S_j(K))$. Similarity scaling and (2.3) give

$$|\pi_\Phi(\omega) - \pi_\Phi(\tau)| \geq r_u \delta = \delta \rho_\Phi(\omega, \tau),$$

which proves (2.4).

The coding is onto by T-I.7.03 and injective by L-I.7.07, so (2.4) makes it a bi-Lipschitz bijection. The cylinder identity gives $\pi_\Phi([u]) = K_u$. If $\pi_\Phi(\omega) \in K_u$, choose $\tau \in [u]$ with the same image; injectivity gives $\omega = \tau$, so $\pi_\Phi^{-1}(K_u) = [u]$. Now (0.1) yields

$$\mu(K_u) = \beta_p([u]) = p_u.$$

The one-label case follows directly from the definitions. □

Theorem — 2.2 (T-I.8.05: exact dimensionality under SSC) [F]

If SSC holds, then for μ-almost every x,

$$\lim_{r\downarrow 0}\frac{\log\mu(\overline{B}(x,r))}{\log r}=\frac{h}{\chi}. \tag{2.7}$$

In particular,

$$\dim_{\mathrm{H}}\mu=\dim_{\mathrm{P}}\mu=\frac{h}{\chi}. \tag{2.8}$$

Proof

The case $N=1$ is a point mass and both sides of (2.7) are zero. Assume $N\geq 2$. For every ω and $n\geq 1$, strict decrease of prefix ratios gives

$$\overline{B}_{\rho_\Phi}(\omega,r_{\omega|n})=[\omega|n]. \tag{2.9}$$

Indeed, a sequence in the cylinder has weighted distance at most $r_{(\omega|n)}$, while a sequence outside it has common prefix of length less than n and hence strictly larger weighted distance.

Given sufficiently small $r>0$, there is a unique n with

$$r_{\omega|n}\leq r<r_{\omega|n-1}. \tag{2.10}$$

Equations (2.9)–(2.10) show that the weighted closed ball has Bernoulli mass

$$\beta_p(\overline{B}_{\rho_\Phi}(\omega,r))=p_{\omega|n}. \tag{2.11}$$

Moreover,

$$0\leq|\log r_{\omega|n}-\log r_{\omega|n-1}|\leq\max_i|\log r_i|. \tag{2.12}$$

For ω in the probability-one set of Lemma 1.1, $-\log r_{(\omega|n)}\ n\chi$; hence (2.10)–(2.12) and (1.4) give

$$\lim_{r\downarrow 0}\frac{\log\beta_p(\overline{B}_{\rho_\Phi}(\omega,r))}{\log r}=\frac{h}{\chi}. \tag{2.13}$$

Put $x=\pi_\Phi(\omega)$. The bi-Lipschitz estimates (2.4) give

$$\pi_\Phi\Big(\overline{B}_{\rho_\Phi}(\omega,r/D)\Big)\subset\overline{B}(x,r)\cap K\subset\pi_\Phi\Big(\overline{B}_{\rho_\Phi}(\omega,r/\delta)\Big). \tag{2.14}$$

The coding is bijective, so applying (0.1) to (2.14) sandwiches the Euclidean ball mass between the two symbolic ball masses. Multiplying the radius by a fixed positive constant does not change the limit in (2.13), because $\log(cr)/\log r\to 1$. This proves (2.7) at the images of the full β_p-measure set, hence at μ-almost every point.

The exact-dimensional local-to-global result C-I.4.08 now gives (2.8). $\square$

I.8.3 The complete Moran package under SSC

Theorem — 3.1 (T-I.8.06: SSC Moran theorem and canonical regularity) [F]

Assume SSC, put $s = s(\Phi)$, and take the canonical weights $p_i = r_i^s$. There are $0 < a \leq b < \infty$ and $r_0 > 0$ such that

$$ar^s \leq \mu(\overline{B}(x,r)) \leq br^s \quad (x \in K,\ 0 < r \leq r_0). \tag{3.1}$$

Furthermore,

$$\dim_{\mathrm{H}} K = \dim_{\mathrm{P}} K = \underline{\dim}_{\mathrm{B}} K = \overline{\dim}_{\mathrm{B}} K = s, \tag{3.2}$$

$$0 < \mathscr{H}^s(K) < +\infty, \qquad \mu = \frac{\mathscr{H}^s \restriction K}{\mathscr{H}^s(K)}. \tag{3.3}$$

Proof

If $N = 1$, then $s = 0$, $K = \{x_*\}$ is a singleton, and all assertions follow with $\mu = \delta_{x_*}$. Assume $N \geq 2$, so $s > 0$, $D = \operatorname{diam}(K) > 0$, and the SSC gap δ in (2.3) is positive.

For $0 < t < 1$, every infinite word has a unique first prefix u satisfying

$$r_u \leq t < r_{u^-}. \tag{3.4}$$

Such a prefix exists because $r_{(\omega|n)} \leq r_{\max}^n \to 0$. If $r_{\max}^m \leq t$, every first crossing occurs by level m, so Λ_t is finite. If one stopping word were a proper prefix of another, then the parent of the longer word would have ratio at most t, contradicting (3.4). Thus the stopping words are prefix-free, the stopping cylinders partition Σ, and

$$r_{\min} t < r_u \leq t, \qquad \mu(K_u) = r_u^s. \tag{3.5}$$

Distinct words in Λ_t are incomparable. Let q be their longest common prefix and let their next letters be $i \neq j$. Since q is a prefix of each parent word in (3.4), $r_q > t$. Therefore

$$\operatorname{dist}(K_u, K_v) \geq r_q \operatorname{dist}(S_i(K), S_j(K)) > \delta t. \tag{3.6}$$

Fix $0 < r < 1$ and use $t = r$. If a ball $\overline{B}(x,r)$ meets m stopping pieces, choose one point from its intersection with each piece. By (3.6) these points are pairwise more than δr apart. The closed balls of radius $\delta r/2$ about them have disjoint interiors and lie in $\overline{B}(x,(1+\delta/2)r)$. Comparison of Euclidean volumes gives

$$m \leq M_{d,\delta} := \left(1 + \frac{2}{\delta}\right)^d. \tag{3.7}$$

The stopping pieces cover K; hence (3.5) and (3.7) imply

$$\mu(\overline{B}(x,r)) \leq M_{d,\delta} r^s. \tag{3.8}$$

For the lower estimate, choose $0 < r < D$ and an address ω of x. Let $u \in \Lambda_{(r/D)}$ be its stopping prefix. Then

$$K_u \subset \overline{B}(x, r)$$

because $x \in K_u$ and $\operatorname{diam}(K_u) = r_u D \le r$. Equations (2.5) and (3.5) give

$$\mu(\overline{B}(x, r)) \ge r_u^s > \left(\frac{r_{\min}}{D}\right)^s r^s. \tag{3.9}$$

Equations (3.8)–(3.9), with

$$r_0 = \frac{1}{2}\min\{1, D\}, \quad a = (r_{\min}/D)^s, \quad b = M_{d,\delta},$$

prove (3.1).

The upper mass estimate and the mass-distribution principle give $\mathscr{H}^s(K) > 0$; T-I.7.06 gives $\mathscr{H}^s(K) < \infty$ and $\dim_{\mathrm{H}} K \le s$. Hence

$$\dim_{\mathrm{H}} K = s. \tag{3.10}$$

For completeness, take a maximal r-separated subset $\{x_1, ..., x_m\}$ of K. The balls $\overline{B}(x_j, r/3)$ are pairwise disjoint, and the lower bound in (3.1) gives

$$1 \ge \sum_{j=1}^{m} \mu(\overline{B}(x_j, r/3)) \ge ma(r/3)^s. \tag{3.11}$$

Maximality makes the radius-r balls centered at the x_j cover K, so $N_r(K) \le a^{-1}3^s r^{-s}$. Thus $\overline{\dim}_{\mathrm{B}} K \le s$. The accepted inequalities

$$\dim_{\mathrm{H}} K \le \dim_{\mathrm{P}} K \le \overline{\dim}_{\mathrm{B}} K, \qquad \dim_{\mathrm{H}} K \le \underline{\dim}_{\mathrm{B}} K \le \overline{\dim}_{\mathrm{B}} K$$

and (3.10) prove (3.2).

It remains to identify the measure. Define

$$\nu = \frac{\mathscr{H}^s \restriction K}{\mathscr{H}^s(K)}.$$

SSC makes the first-level pieces disjoint. Similarities scale $\mathscr{H}^s$ by r_i^s, so for every Borel set A,

$$\nu(A) = \sum_i r_i^s \nu(S_i^{-1}(A)). \tag{3.12}$$

Thus ν is invariant for the canonical weights. Uniqueness in T-I.7.04 gives $\nu = \mu$, proving (3.3). □

Corollary — 3.2 (C-I.8.07: variational and unique full-dimensional weight) [F]

Under SSC,

$$\dim_{\mathrm{H}} \mu_{\Phi,p} \le \dim_{\mathrm{H}} K = s(\Phi), \tag{3.13}$$

with equality if and only if $p_i = r_i^s$ for all i. Consequently

$$s(\Phi) = \max_{p_i>0,\ \sum_i p_i=1} \dim_{\mathrm{H}} \mu_{\Phi,p}, \tag{3.14}$$

and the canonical vector is the unique maximizing strict Bernoulli weight.

Proof

Theorem 2.2 gives $\dim_{\mathrm{H}} \mu = h/\chi$, Lemma 1.2 gives $h/\chi \le s$ with equality only at the canonical vector, and Theorem 3.1 gives $\dim_{\mathrm{H}} K = s$. Combining these statements proves (3.13)–(3.14), including uniqueness. □

I.8.4 Explicit entropy loss from duplicate labels

Proposition — 4.1 (P-I.8.08: separated and duplicate-label models) [F]

The following three systems distinguish geometric dimension from labelled information.

(1) For $C_0(x) = x/3$ and $C_2(x) = x/3 + 2/3$ with weights $(t, 1-t)$, where $0 < t < 1$, the self-similar measure satisfies

$$\dim_{\mathrm{H}} \mu = \frac{-t\log t - (1-t)\log(1-t)}{\log 3}.$$

(2) For $S_0 = S_1 = C_0$ and $S_2 = C_2$ with strict weights (p_0, p_1, p_2), put $q = p_0 + p_1$. Then

$$\begin{aligned}\dim_{\mathrm{H}} \mu &= \frac{-q\log q - p_2\log p_2}{\log 3} \\ &< \frac{-p_0\log p_0 - p_1\log p_1 - p_2\log p_2}{\log 3} = \mathrm{sdim}(\mu_{\Phi,p}).\end{aligned}$$

This system has dimension drop from its natural dimension. If $q = p_2 = 1/2$, it is nevertheless full-dimensional in its Cantor attractor.

(3) If all N labelled maps on $\mathbb{R}^d$ equal $S(x) = rx$, with $0 < r < 1$, their self-similar measure is δ_0 for every strict weight vector. With uniform weights,

$$\dim_{\mathrm{H}} \mu = 0, \qquad \mathrm{sdim}(\mu_{\Phi,p}) = \frac{\log N}{\log(1/r)},$$

so the labelled similarity dimension is unbounded as N grows with r fixed.

Proof

For the middle-third Cantor maps $C_0(x) = x/3$ and $C_2(x) = x/3 + 2/3$, SSC holds and both contraction ratios equal $1/3$. With weights $(t, 1-t)$, Theorem 2.2 gives

$$\dim_{\mathrm{H}} \mu = \frac{-t\log t - (1-t)\log(1-t)}{\log 3}. \tag{4.1}$$

Now use the duplicate system

$$S_0 = S_1 = C_0, \qquad S_2 = C_2, \tag{4.2}$$

with strict weights (p_0, p_1, p_2), and put $q = p_0 + p_1$. Its measure operator is

$$\mathscr{M}_{\Phi,p}(\nu) = q(C_0)_{\#}\nu + p_2(C_2)_{\#}\nu. \tag{4.3}$$

Thus uniqueness in T-I.7.04 identifies the measure with the separated two-map Cantor measure of weights (q, p_2). Applying (4.1) gives

$$\dim_{\mathrm{H}} \mu = \frac{-q\log q - p_2\log p_2}{\log 3}. \tag{4.4}$$

The labelled entropy is larger. Indeed, with $\alpha = p_0/q \in (0,1)$,

$$\begin{aligned} h(p_0, p_1, p_2) &= h(q, p_2) + q\big[-\alpha\log\alpha - (1-\alpha)\log(1-\alpha)\big] \\ &> h(q, p_2). \end{aligned} \tag{4.5}$$

All three labelled ratios are $1/3$. The finite Gibbs inequality with the uniform three-label vector gives $h(p_0, p_1, p_2) \le \log 3$. Thus the labelled similarity dimension is at most the ambient dimension one; taking the ambient minimum does not alter it. Division of (4.5) by $\log 3$ therefore proves strict drop from the labelled natural dimension. If $q = p_2 = 1/2$, then (4.4) equals $\log 2/\log 3$, the dimension of the Cantor attractor, while the strict inequality in (4.5) remains. The measure is therefore full-dimensional in its attractor and simultaneously has labelled natural-dimension drop.

Finally, suppose all N labelled maps equal $S(x) = rx$. Their measure operator is simply

$$\mathscr{M}_{\Phi,p}(\nu) = S_{\#}\nu,$$

whose unique invariant probability is the point mass δ_0. Its Hausdorff dimension is zero. For uniform weights the labelled measure similarity dimension is

$$\frac{\log N}{\log(1/r)}, \tag{4.6}$$

which tends to infinity with N. □

I.8.5 The terminal open-set extension

Guided result — 5.1 (G-I.8.09: entropy/Lyapunov exact dimensionality under OSC) [G]

[G | currency: Published | omitted: projection-entropy calculus, conditional Shannon–McMillan–Breiman theorem, and the Feng–Hu local-dimension theorem | sources: FH-01, Theorem 2.8 and Corollary 4.16; HUT81-IFS, Section 5.3 comparable-cylinder lemma]

If Φ satisfies the OSC, then for every strict p,

$$\lim_{r\downarrow 0} \frac{\log \mu_{\Phi,p}(\overline{B}(x,r))}{\log r} = \frac{h(p)}{\chi(\Phi,p)} \quad \text{for } \mu_{\Phi,p}\text{-almost every } x, \tag{5.1}$$

and

$$\dim_{\mathrm{H}} \mu_{\Phi,p} = \dim_{\mathrm{P}} \mu_{\Phi,p} = \frac{h(p)}{\chi(\Phi,p)}. \tag{5.2}$$

References. The projection-entropy statement is Feng and Hu [FH09, Definition 2.1, p. 5, Theorem 2.8, p. 7, and Corollary 4.16, p. 25]. The OSC fiber estimate uses Hutchinson's comparable-cylinder packing lemma [Hut81, Section 5.3, pp. 19–20], already isolated in G-I.7.13.

Guided proof architecture

If $N = 1$, the measure is a point mass and $h(p) = 0$, so the conclusion follows directly. Henceforth assume $N \geq 2$.

The external HUT81 bound (7.5a) in I.7 gives a constant M_Φ such that, for every $0 < t < 1$ and $z \in K$,

$$\#\{u \in \Lambda_t \ :\ z \in K_u\} \leq M_\Phi. \tag{5.3}$$

Every address in $\mathrm{Fib}_\Phi(z)$ has exactly one prefix in Λ_t, so

$$\mathrm{Fib}_\Phi(z) \subset \bigcup_{\substack{u\in\Lambda_t\\ z\in K_u}} [u]. \tag{5.4}$$

If $u \in \Lambda_t$, then $r_u \leq t$ and $r_u \geq r_{\min}^{|u|}$. Therefore

$$|u| \geq \frac{\log t}{\log r_{\min}}, \qquad \mathrm{diam}_{d_\Sigma}[u] \leq 2^{-|u|} \leq t^\alpha, \qquad \alpha = \frac{\log 2}{\log(1/r_{\min})} > 0. \tag{5.5}$$

For every $q > 0$, (5.3)–(5.5) give

$$\mathscr{H}^q_{t^\alpha}(\mathrm{Fib}_\Phi(z); d_\Sigma) \leq M_\Phi t^{\alpha q} \longrightarrow 0. \tag{5.6}$$

Thus every coding fiber has symbolic Hausdorff dimension zero. The source uses cylinder scales N^{-k} rather than 2^{-k} at depth k. Replacing $\log 2$ by $\log N$ in (5.5)–(5.6) proves the same zero-dimensional statement in that metric. Fibers outside K are empty and require no additional cover.

For precision, let $\sigma(\omega_1, \omega_2, \dots) = (\omega_2, \omega_3, \dots)$ be the left shift, let m be a shift-invariant Borel probability on Σ, let $\mathscr{P} = \{[i] : i \in \Lambda\}$ be the first-coordinate partition, and let $\gamma = \mathscr{B}(\mathbb{R}^d)$, the Borel sigma-algebra. Here $\pi_\Phi^{-1}\gamma = \{\pi_\Phi^{-1}(A) : A \in \gamma\}$ is the pull-back sigma-algebra. For a sub-sigma-algebra $\mathscr{A}$ of $\mathscr{B}(\Sigma)$, define the **conditional information** and **conditional entropy** by

$$\mathbf{I}_m(\mathscr{P} \mid \mathscr{A})(x) = -\sum_{P \in \mathscr{P}} \mathbf{1}_P(x) \log \mathbb{E}_m(\mathbf{1}_P \mid \mathscr{A})(x), \qquad H_m(\mathscr{P} \mid \mathscr{A}) = \int \mathbf{I}_m(\mathscr{P} \mid \mathscr{A})\, dm. \tag{5.7}$$

The **Feng–Hu projection entropy** appearing only in this G architecture is

$$h_\pi(\sigma, m) = H_m(\mathscr{P} \mid \sigma^{-1}\pi_\Phi^{-1}\gamma) - H_m(\mathscr{P} \mid \pi_\Phi^{-1}\gamma). \tag{5.8}$$

In (5.7), a term outside its partition element is defined to be zero. If $q_P = \mathbb{E}_m(\mathbf{1}_P \mid \mathscr{A})$, then $m(P \cap \{q_P = 0\}) = \int_{\{q_P = 0\}} q_P\, dm = 0$. Thus the logarithm on P is finite almost everywhere; its value on the remaining null set can be chosen arbitrarily. Conditional expectation and truncation give

$$H_m(\mathscr{P} \mid \mathscr{A}) = \int \sum_{P \in \mathscr{P}} -q_P \log q_P\, dm, \qquad 0 \le H_m(\mathscr{P} \mid \mathscr{A}) \le \log N, \tag{5.8a}$$

where $0 \log 0 = 0$. The upper bound follows from the finite Gibbs inequality against the uniform vector (or its limit when some coordinates vanish). Consequently both terms in (5.8) are well-defined finite numbers independent of the chosen versions. Formula (5.8) is FH-01, Definition 2.1; its dynamical calculus remains part of the omitted I.9 block.

The Bernoulli law is shift-invariant and ergodic by the allowed foundational dynamical-systems results. Write $h(\sigma, m)$ for its measure-theoretic entropy. Independence gives $h(\sigma, \beta_p) = h(p)$ for the Bernoulli shift. Feng–Hu Corollary 4.16, applied to the fiber estimate above, now gives

$$h_\pi(\sigma, \beta_p) = h(\sigma, \beta_p) = h(p). \tag{5.9}$$

To check the differential hypothesis of Feng–Hu Theorem 2.8, note that similarities extend to differentiable bijections of $\mathbb{R}^d$. They preserve a sufficiently large closed ball. For every unit vector v,

$$|DS_{\omega|n}(y)v| = r_{\omega|n}, \qquad -\frac{1}{n}\log r_{\omega|n} \longrightarrow \chi(\Phi, p) \quad \text{for } \beta_p\text{-almost every } \omega. \tag{5.9a}$$

Here $DS_u(y)$ is the derivative at y; the equality holds at every y. Thus the smallest and largest differential expansion factors have the same positive logarithmic contraction limit, as required by that theorem. Its ergodic conclusion is the local-dimension formula $d_\mu(x) = h_\pi(\sigma, \beta_p)/\chi(\Phi, p)$ almost everywhere. Substituting (5.9) yields (5.1). The accepted exact-dimensional dictionary from I.4 yields (5.2).

These projection-entropy arguments are intentionally omitted here. Chapter I.9 supplies the full projection-entropy development, the conditional Shannon–McMillan–Breiman theorem, and the proof of Feng–Hu Theorem 2.8. The HUT81 proof of (5.3) is the terminal geometric block isolated in G-I.7.13. No F result uses Theorem 5.1.

I.8.6 Intuition and strategy

The following discussion explains the geometric ideas behind the proofs.

I.8.6.1 Two clocks run along a random address

For a word $\omega|n$, its probability and geometric scale are

$$p_{\omega|n} \quad \text{and} \quad r_{\omega|n}.$$

Their negative logarithms are additive. The information clock accumulates $-\log p_{(\omega_k)}$; the contraction clock accumulates $-\log r_{(\omega_k)}$. Along a Bernoulli-typical address, the two clocks advance at average speeds

$$h(p) \quad \text{and} \quad \chi(\Phi, p).$$

Dividing the speeds predicts the amount of information carried per unit of geometric magnification:

$$\frac{h(p)}{\chi(\Phi, p)}.$$

The fourth-moment proof in the proofs is intentionally elementary. It shows exactly where the probability-one statement enters instead of hiding it behind a general ergodic theorem.

I.8.6.2 Why the ratio is always an upper bound

A typical level has at most about $\exp(nh)$ relevant words, and each relevant cylinder has diameter at most about $\exp(-n\chi)$. Covering the typical image by those cylinders gives the dimension quotient h/χ. Overlaps can only make the geometric cover more redundant, so they cannot invalidate the upper bound.

The ambient truncation by d is kept in the definition of natural dimension because a labelled system can carry arbitrarily much symbolic entropy while living in $\mathbb{R}^d$. Under genuine separation the quotient is already at most the set similarity dimension and hence at most d.

I.8.6.3 The weighted symbolic metric uses the correct ruler

The fixed metric d_Σ treats every symbolic generation as the same amount of magnification. That is convenient for topology but ignores unequal contraction ratios. The weighted distance

$$\rho_\Phi(\omega,\tau) = r_{\omega\wedge\tau}$$

measures the actual geometric scale of the last common cylinder.

Under strong separation, the coding map neither collapses nor expands this ruler by more than fixed constants. Symbolic cylinders become exact metric balls, their masses are exactly p_u, and the strong law transfers directly to Euclidean local dimension. This is the cleanest setting in which the entropy/Lyapunov formula is a literal change of metric.

I.8.6.4 Canonical weights balance mass and diameter at every word

The similarity dimension s solves $\sum_i r_i^s = 1$. Choosing

$$p_i = r_i^s$$

makes every cylinder satisfy

$$p_u = r_u^s.$$

Under SSC, comparable cylinders are separated at a comparable Euclidean scale. A ball can meet only boundedly many of them, giving the upper r^s mass bound; following one address gives the lower bound. Thus the canonical measure is Ahlfors regular, not merely exact dimensional.

The Gibbs inequality says that no other Bernoulli vector has a larger entropy/Lyapunov quotient. Under SSC the canonical vector is therefore the unique Bernoulli measure of full attractor dimension.

I.8.6.5 Duplicate labels reveal where entropy disappears

Suppose two labels carry the same map. The symbolic process records which label was chosen, but the geometric point records only their common image. If their weights are p_0, p_1, geometry sees the merged weight

$$q = p_0 + p_1.$$

The entropy chain rule becomes

$$h(p_0, p_1, p_2) = h(q, p_2) + q\, h\!\left(\frac{p_0}{q}, \frac{p_1}{q}\right).$$

The conditional term is positive and is lost under coding. This gives a fully explicit dimension drop: the measure can still fill the Cantor attractor dimension while falling strictly below its labelled natural dimension.

I.8.6.6 OSC and the first appearance of projection entropy

OSC allows boundary identifications, so coding need not be globally injective. Nevertheless, its bounded-overlap geometry makes each coding fiber symbolically zero-dimensional.

Projection entropy measures the information that survives coding. In the Feng–Hu theorem, the fiber bound forces it to equal classical Bernoulli entropy.

The resulting OSC formula is stated in I.8 as a terminal sourced bridge. Its projection-entropy and conditional Shannon–McMillan–Breiman proof is not compressed here: it is the full-proof task of I.9. This keeps the elementary SSC mechanism visible while marking exactly what must change when addresses are no longer unique.

I.8.6.7 Dependency map

The full-proof route is

fourth-moment Bernoulli law
$\longrightarrow$ information and contraction averages
$\longrightarrow$ arbitrary-overlap natural-dimension upper bound
weighted symbolic ultrametric + SSC
$\longrightarrow$ bi-Lipschitz coding
$\longrightarrow$ exact dimension h/χ
$\longrightarrow$ canonical Ahlfors regularity
$\longrightarrow$ unique full-dimensional Bernoulli weight
duplicate-label merging $\longrightarrow$ explicit entropy loss

The cited OSC extension follows the route

OSC comparable-cylinder bound
$\longrightarrow$ zero-dimensional coding fibers
$\longrightarrow$ Feng–Hu projection entropy formula
$\longrightarrow$ G-I.8.09.

No F-labelled result uses that bridge.

I.8.7 Exercises

These exercises use the preceding full proofs, earlier chapters, and the assumed real-analysis and foundational dynamical-systems material. The Feng–Hu theorem and the omitted OSC packing argument are not prerequisites.

I.8.7.1 Exercise 1 — The exact fourth-moment count

Let $X_1, X_2, \ldots$ be independent, identically distributed, centered random variables with

$$\mathbb{E}(X_1^2) = \sigma^2, \qquad \mathbb{E}(X_1^4) = m_4 < +\infty.$$

Prove that

$$\mathbb{E}\left(\sum_{k=1}^{n} X_k\right)^4 = nm_4 + 3n(n-1)\sigma^4.$$

Deduce the strong law for bounded iid variables using only Markov's inequality and Borel–Cantelli.

I.8.7.2 Exercise 2 — Asymptotic equipartition on a finite alphabet

For $\varepsilon > 0$, let

$$A_n(\varepsilon) = \{u \in \Lambda^n \,:\, e^{-n(h+\varepsilon)} \le p_u \le e^{-n(h-\varepsilon)}\}.$$

Prove that

$$\sum_{u \in A_n(\varepsilon)} p_u \longrightarrow 1$$

and that, for all sufficiently large n,

$$\frac{1}{2} e^{n(h-\varepsilon)} \le \# A_n(\varepsilon) \le e^{n(h+\varepsilon)}.$$

I.8.7.3 Exercise 3 — The equicontractive entropy maximum

Assume $r_i = r$ for all i and there are N labels. Prove

$$\frac{h(p)}{\chi(\Phi, p)} \le \frac{\log N}{\log(1/r)},$$

with equality if and only if $p_i = 1/N$ for every i.

I.8.7.4 Exercise 4 — Comparing the two symbolic metrics

Let

$$\gamma_{\min} = \frac{\log r_{\min}}{\log(1/2)}, \qquad \gamma_{\max} = \frac{\log r_{\max}}{\log(1/2)}.$$

Prove that $\gamma_{\min} \ge \gamma_{\max} > 0$ and, for distinct ω, τ,

$$d_\Sigma(\omega, \tau)^{\gamma_{\min}} \le \rho_\Phi(\omega, \tau) \le d_\Sigma(\omega, \tau)^{\gamma_{\max}}.$$

Deduce that ρ_Φ and d_Σ induce the same topology. Prove directly that

$$\overline{B}_{\rho_\Phi}(\omega, r_{\omega|n}) = [\omega|n].$$

I.8.7.5 Exercise 5 — A pointwise frequency formula under SSC

Assume SSC. Suppose ω has limiting digit frequencies

$$f_{n,i}(\omega) \longrightarrow q_i \qquad (i \in \Lambda),$$

where $q = (q_i)$ is a probability vector that may have zero entries. Prove that at $x = \pi_\Phi(\omega)$ the local dimension of $\mu_{(\Phi,p)}$ exists and is

$$\frac{-\sum_i q_i \log p_i}{-\sum_i q_i \log r_i}.$$

I.8.7.6 Exercise 6 — Typical and atypical points in a Cantor measure

For the middle-third Cantor measure with weights $(t, 1-t)$, let ω have limiting frequency θ of the digit 0. Compute the local dimension at $\pi(\omega)$. Specialize your formula to a typical point and to the periodic address $(02)^\infty$.

I.8.7.7 Exercise 7 — Ambient truncation can be substantial

Let N labelled maps on $\mathbb{R}^d$ all equal $S(x) = rx$. For uniform weights, compute $\operatorname{sdim}(\mu)$, $\dim_{\mathrm{nat}}(\mu)$, and $\dim_{\mathrm{H}} \mu$. Show that the first can be arbitrarily large, the second never exceeds d, and the third remains zero.

I.8.7.8 Exercise 8 — Separation at a stopping scale

Assume $N \geq 2$ and SSC, let δ be the first-level gap, and fix $0 < t < 1$. Prove that distinct $u, v \in \Lambda_t$ satisfy

$$\operatorname{dist}(K_u, K_v) > \delta t.$$

Deduce directly that a Euclidean ball of radius t meets at most

$$\left(1 + \frac{2}{\delta}\right)^d$$

stopping pieces.

I.8.7.9 Exercise 9 — Ahlfors regularity and covering numbers

Let E be a nonempty compact metric set, let ν be a Borel probability supported on E, and suppose that $s \geq 0$, $0 < a \leq b < \infty$, and $r_0 > 0$ satisfy

$$ar^s \leq \nu(\overline{B}(x, r)) \leq br^s \quad (x \in E,\ 0 < r \leq r_0).$$

Prove that there are $0 < c \leq C < \infty$ such that

$$cr^{-s} \leq N_r(E) \leq Cr^{-s}$$

for every sufficiently small r.

I.8.7.10 Exercise 10 — Which separated Bernoulli measures are Ahlfors regular?

Assume $N \geq 2$ and SSC. Suppose $\mu_{(\Phi,p)}$ is t-Ahlfors regular on K for some $t > 0$. Prove that

$$p_i = r_i^t \qquad (i \in \Lambda),$$

and hence $t = s(\Phi)$ and p is canonical. Prove the converse from T-I.8.06.

I.8.7.11 Exercise 11 — Merging arbitrary classes of duplicate labels

Partition Λ into classes C such that the maps are identical exactly within each class. Put

$$q_C = \sum_{i \in C} p_i .$$

Assume the reduced IFS containing one map from each class satisfies SSC. Prove that the original self-similar measure equals the reduced self-similar measure with weights (q_C), and derive

$$h(p) = h(q) + \sum_C q_C h((p_i/q_C)_{i \in C}) .$$

Compute the actual dimension and the labelled entropy/Lyapunov upper bound.

I.8.7.12 Exercise 12 — Allowing zero weights under SSC

Let $p_i \geq 0$, $\sum_i p_i = 1$, and use the convention $0 \log 0 = 0$. Assume the full IFS satisfies SSC and put $\Lambda_+ = \{i : p_i > 0\}$. Prove that the invariant measure is carried by the sub-attractor K_+ and is exact dimensional with

$$\dim_H \mu = \frac{-\sum_{i \in \Lambda_+} p_i \log p_i}{-\sum_{i \in \Lambda_+} p_i \log r_i} .$$

I.8.7.13 Exercise 13 — Similarity conjugacy and the dimension formula

Let Q be an invertible similarity and $\tilde{S}_i = Q \circ S_i \circ Q^{-1}$. Prove that the contraction ratios, $h(p)$, $\chi(\Phi, p)$, and the natural dimension are unchanged. Prove that

$$\mu_{\widetilde{\Phi}, p} = Q_\# \mu_{\Phi, p}$$

and that exact dimensionality and its value are preserved.

Chapter I.9

Conformal IFS and Projection Entropy

Conformal systems retain infinitesimal similarity without requiring a fixed contraction ratio at every point. Distortion estimates replace exact scaling, and pressure organizes the accumulated contraction along words. Projection entropy measures the information retained after symbolic points are mapped to the geometric attractor. We specify the smoothness, invariance, and conditioning conventions before using them, so that the transition from symbolic to geometric dimension has an explicit mathematical meaning.

The definitions below fix this chapter's conventions. Geometric intuition and proof strategy are discussed separately after the main development; exercises and their solutions provide a second route through the material.

Definitions and notation

All logarithms are natural. The shift alphabet Λ is finite and nonempty, $\Sigma = \Lambda^{\mathbb{N}}$, and $P = ([i])_{(i\in\Lambda)}$ is the first-coordinate partition. Terms imported from A-RA or A-DS are nevertheless normalized here whenever their convention enters a displayed formula.

The coding map and attractor are those of I.7. For distinct sequences let $k(\omega,\tau)$ be the number of initial coordinates in which they agree and put $d_\Sigma(\omega,\tau) = 2^{-k(\omega,\tau)}$; put $d_\Sigma(\omega,\omega) = 0$. Hölder continuity on Σ refers to this metric: there exist $C < \infty$ and $\beta > 0$ such that $|f(\omega) - f(\tau)| \le C d_\Sigma(\omega,\tau)^\beta$ for all sequences. A probability m on Σ is invariant when $m(\sigma^{-1}E) = m(E)$ for every Borel E. Projection entropy below is defined for invariant m. For a finite partition ξ, write $H_m(\xi) = -\sum_{(C\in\xi)} m(C)\log m(C)$ with $0\log 0 = 0$.

Definition — I.9.01: Smooth and conformal iterated systems

Let X be a nonempty compact subset of $\mathbb{R}^d$. A C^1 **IFS on** X is a finite tuple $\Phi = (S_i)_{(i\in\Lambda)}$ for which some open neighborhood U of X has the following properties: every S_i extends to a C^1 diffeomorphism from U onto $S_i(U) \subset U$, $S_i(X) \subset X$, and the maps have a common Lipschitz constant strictly below one on U (after shrinking to a fixed neighborhood of X if necessary). It is $C^{1+\alpha}$ if the derivatives are uniformly α-Hölder

on a fixed neighborhood of X, where $0 < \alpha \le 1$.
Explicitly, the Hölder condition means that on some fixed neighborhood V of X with closure contained in U, one constant $H < \infty$ satisfies $\|DS_i(x) - DS_i(y)\| \le H|x-y|^\alpha$ for all i and $x, y \in V$.
For an invertible linear map $A : \mathbb{R}^d \to \mathbb{R}^d$, put

$$\|A\| = \sup_{|v|=1} |Av|, \qquad \mathfrak{m}(A) = \inf_{|v|=1} |Av|.$$

The map A is **conformal** if $|Av| = \|A\||v|$ for every v. The IFS is **conformal** if every $DS_i(x)$ is conformal for $x \in U$. For a word u and $x \in X$, its derivative scale is

$$a_u(x) = \|DS_u(x)\|, \qquad S_u = S_{u_1} \circ \cdots \circ S_{u_n}.$$

Definition — I.9.02: Bounded distortion and conformal cylinders

Let K be the attractor. The IFS has **derivative bounded distortion** if a constant $C_{\mathrm{bd}} \ge 1$ satisfies

$$C_{\mathrm{bd}}^{-1} \le \frac{a_u(x)}{a_u(y)} \le C_{\mathrm{bd}} \quad (u \in \Lambda^*,\ x, y \in X).$$

It has **cylinder bounded distortion** if, after enlarging C_{bd} if necessary,

$$C_{\mathrm{bd}}^{-1} a_u(z)|x-y| \le |S_u(x) - S_u(y)| \le C_{\mathrm{bd}} a_u(z)|x-y|$$

for all words u, all $x, y \in K$, and all $z \in X$. A **conformal cylinder** is $K_u = S_u(K)$. Its upper derivative scale is $q_u = \sup_{(x\in X)} a_u(x)$.

Definition — I.9.03: Pressure and Gibbs states

For a continuous potential $\varphi : \Sigma \to \mathbb{R}$, its Birkhoff sum and level-n partition sum are

$$S_n\varphi(\omega) = \sum_{j=0}^{n-1} \varphi(\sigma^j\omega), \qquad Z_n(\varphi) = \sum_{u\in\Lambda^n} \exp\left(\sup_{\omega\in[u]} S_n\varphi(\omega)\right).$$

When the limit exists, the **pressure** is

$$P(\varphi) = \lim_{n\to\infty} \frac{1}{n} \log Z_n(\varphi).$$

An invariant probability m is an **equilibrium state** for φ if

$$h_m(\sigma) + \int \varphi\, dm = P(\varphi).$$

It is a **Gibbs measure** for φ if some $C_G \ge 1$ satisfies

$$C_G^{-1} \leq \frac{m([u])}{\exp(S_n\varphi(\omega) - nP(\varphi))} \leq C_G$$

for every n, $u \in \Lambda^n$, and $\omega \in [u]$.
For a conformal IFS, the **geometric potential** is

$$\phi(\omega) = \log \|DS_{\omega_1}(\pi(\sigma\omega))\| < 0.$$

The geometric pressure function is $P_\Phi(t) = \mathscr{P}(t\phi)$. Its zero, when it exists uniquely, is the **pressure dimension** s_Φ.

Definition — I.9.04: Conditional information and entropy

Let (Ω, F, m) be a probability space, ξ a finite measurable partition, and $A \subset F$ a sub-sigma-algebra. For $C \in \xi$, choose a version

$$p_C^{\mathscr{A}} = \mathbb{E}_m(\mathbf{1}_C \mid \mathscr{A}).$$

With $0 \log 0 = 0$, define

$$I_m(\xi \mid \mathscr{A}) = -\sum_{C \in \xi} \mathbf{1}_C \log p_C^{\mathscr{A}}, \qquad H_m(\xi \mid \mathscr{A}) = \int I_m(\xi \mid \mathscr{A})\, dm.$$

All identities are modulo m-null sets and hence do not depend on the chosen versions. The notation $\xi(x)$ means the unique cell containing x, $\hat{\xi}$ means the sigma-algebra generated by the cells, and $\xi \vee \eta$ is the partition of nonempty cell intersections. A join of sigma-algebras is the sigma-algebra generated by their union. For a measurable map T, $T^{-1}\xi$ and $T^{-1}A$ consist of the corresponding inverse images (empty partition cells are discarded). In a pull-back formula the partition and conditioning algebra on the target of T are evaluated under the target law T_*m.
The information function is evaluated on the cell containing the point. On each cell C, the set where $p_C^A = 0$ is null, since its measure equals the integral of p_C^A over that set. Values on this exceptional set can be assigned arbitrarily. Moreover, $H_m(\xi|A) = \sum_{(C\in\xi)} \int -p_C^A \log p_C^A dm \leq H_m(\xi) < \infty$. Thus the difference defining projection entropy is integrable.

Definition — I.9.05: Projection entropy and conformal Lyapunov exponent

Let $A_\pi = \pi^{-1}B(\mathbb{R}^d)$ and let

$$\Delta_\pi = I_m(\mathscr{P} \mid \sigma^{-1}\mathscr{A}_\pi) - I_m(\mathscr{P} \mid \mathscr{A}_\pi).$$

The **projection entropy** and **local projection entropy** are

$$h_\pi(\sigma, m) = \int \Delta_\pi\, dm, \qquad h_\pi(\sigma, m, \omega) = \mathbb{E}_m(\Delta_\pi \mid \mathscr{I})(\omega),$$

where $I = \{E : \sigma^{-1}E = E \bmod m\}$ is the invariant sigma-algebra.
For a conformal IFS put

$$\ell_\Phi(\omega) = -\log \|DS_{\omega_1}(\pi(\sigma\omega))\|, \qquad \lambda_\Phi(m,\omega) = \mathbb{E}_m(\ell_\Phi \mid \mathscr{I})(\omega).$$

If m is ergodic, its Lyapunov exponent is the constant $\chi_\Phi(m) = \int \ell_\Phi dm > 0$.

Definition — I.9.06: Fiber-prefix multiplicity

For $z \in K$ and $n \geq 1$, define

$$N_n(z) = \#\{u \in \Lambda^n : z \in K_u\}, \qquad M_n = \sup_{z \in K} N_n(z).$$

The coding has **uniformly subexponential fiber-prefix multiplicity** if

$$\lim_{n\to\infty} \frac{1}{n} \log M_n = 0.$$

This is a quantitative property of prefixes, not merely the assertion that each fiber is countable or has symbolic Hausdorff dimension zero.

All balls are closed Euclidean balls and all logarithms are natural. Let $\Lambda = \{1, ..., N\}$, $\Sigma = \Lambda^{\mathbb{N}}$, and let σ be the left shift. The case $N = 1$ is retained. Write

$$\mathscr{P} = \{[i] : i \in \Lambda\}, \qquad \mathscr{P}_0^{n-1} = \bigvee_{j=0}^{n-1} \sigma^{-j}\mathscr{P}. \tag{I.9.1}$$

The entropy and invariant-sigma-algebra conventions are those of A-DS. Every fractal-specific object used below is defined in the chapter.

We make the working conventions explicit. If $\omega \neq \tau$ have exactly k common initial coordinates, put $d_\Sigma(\omega,\tau) = 2^{-k}$; put $d_\Sigma(\omega,\omega) = 0$. A function f is Hölder on this space if $|f(\omega) - f(\tau)| \leq C d_\Sigma(\omega,\tau)^\beta$ for fixed $C < \infty$ and $\beta > 0$. Its Birkhoff sum is $S_n f = \sum_{(j=0)}^{n-1} f \circ \sigma^j$. The Bowen bound for f means that the oscillation of $S_n f$ on every length-n cylinder has one bound independent of n and the cylinder. This is the meaning used in Lemma 1.3.

For a finite partition ξ, $\xi(x)$ is its cell containing x, $\hat{\xi}$ is the sigma-algebra generated by its cells, and $\xi \vee \eta$ is the partition into nonempty intersections. Joins of sigma-algebras are generated by their union; pull-backs mean inverse images. We use $H_m(\xi) = -\sum_{(C\in\xi)} m(C) \log m(C)$ with $0 \log 0 = 0$. Unless a general probability space is explicitly specified, m is an invariant Borel probability on Σ, so $m(\sigma^{-1}E) = m(E)$ for every Borel set E.

I.9.1 Smooth conformal systems

Definition — 1.1 (C^1 and conformal IFS)

Let $X \subset \mathbb{R}^d$ be nonempty and compact. A finite C^1 **IFS on** X is a tuple $\Phi = (S_i)_{(i\in\Lambda)}$ for which there are an open neighborhood $U \supset X$ and C^1 diffeomorphisms

$$S_i : U \longrightarrow S_i(U) \subset U \tag{I.9.2}$$

such that $S_i(X) \subset X$ and, after shrinking U to a fixed neighborhood of X,

$$\sup_{x\in U} \|DS_i(x)\| \le \vartheta, \qquad |S_i(x) - S_i(y)| \le \vartheta|x-y| \quad (x, y \in U, i \in \Lambda) \tag{I.9.3}$$

for one $0 < \vartheta < 1$. It is $C^{1+\alpha}$, $0 < \alpha \le 1$, if the derivatives have a common α-Hölder constant on a fixed open neighborhood of X whose closure is contained in U.
For an invertible linear map A, define

$$\|A\| = \sup_{|v|=1} |Av|, \qquad \mathfrak{m}(A) = \inf_{|v|=1} |Av|. \tag{I.9.4}$$

The linear map is **conformal** if

$$|Av| = \|A\||v| \quad (v \in \mathbb{R}^d). \tag{I.9.5}$$

The IFS is conformal if every $DS_i(x)$ is conformal on U. Thus $\mathfrak{m}(DS_i(x)) = \|DS_i(x)\| > 0$.
Let K be the attractor and $\pi : \Sigma \to K$ the coding map. For a word $u = u_1...u_n$, put

$$S_u = S_{u_1} \circ \cdots \circ S_{u_n}, \quad K_u = S_u(K), \quad a_u(x) = \|DS_u(x)\|, \quad q_u = \sup_{x\in X} a_u(x). \tag{I.9.6}$$

The set K_u is the **conformal cylinder** indexed by u.

Definition — 1.2 (bounded distortion)

The system has **derivative bounded distortion** if some $C_{\mathrm{bd}} \ge 1$ satisfies

$$C_{\mathrm{bd}}^{-1} \le \frac{a_u(x)}{a_u(y)} \le C_{\mathrm{bd}} \quad (u \in \Lambda^*,\ x, y \in X). \tag{I.9.7}$$

It has **cylinder bounded distortion** if, possibly after increasing C_{bd},

$$C_{\mathrm{bd}}^{-1} a_u(z)|x-y| \le |S_u(x) - S_u(y)| \le C_{\mathrm{bd}} a_u(z)|x-y| \tag{I.9.8}$$

for $u \in \Lambda^*$, $x, y \in K$, and $z \in X$.

Lemma — 1.3 (L-I.9.01: uniform conformal distortion) [F]

Every finite $C^{1+\alpha}$ conformal IFS has derivative and cylinder bounded distortion. There is $C \geq 1$ such that

$$C^{-1} q_u \operatorname{diam} K \leq \operatorname{diam} K_u \leq C q_u \operatorname{diam} K, \tag{I.9.9}$$

and

$$C^{-1} q_u q_v \leq q_{uv} \leq q_u q_v. \tag{I.9.10}$$

The geometric potential

$$\phi(\omega) = \log \|DS_{\omega_1}(\pi(\sigma\omega))\| \tag{I.9.11}$$

is Hölder. It also satisfies the Bowen bound

$$\sup_{n\geq 1} \sup_{\omega|_n = \tau|_n} |S_n\phi(\omega) - S_n\phi(\tau)| < \infty. \tag{I.9.12}$$

Proof

Set $a_i(x) = \|DS_i(x)\|$. Compactness, invertibility, and finiteness give

$$0 < a_* := \min_{i,x\in X} a_i(x) \leq \max_{i,x\in X} a_i(x) < 1. \tag{I.9.13}$$

After enlarging the Hölder constant, there is $H < \infty$ such that

$$|\log a_i(x) - \log a_i(y)| \leq H|x-y|^\alpha \quad (i \in \Lambda,\ x, y \in X). \tag{I.9.14}$$

Conformality makes norms multiplicative along the derivative chain. Hence, for $u = u_1 \ldots u_n$,

$$a_u(x) = \prod_{j=1}^{n} a_{u_j}(S_{u_{j+1}\cdots u_n} x). \tag{I.9.15}$$

Using (I.9.3) and (I.9.14),

$$\begin{aligned} |\log a_u(x) - \log a_u(y)| &\leq H \sum_{j=1}^{n} |S_{u_{j+1}\cdots u_n} x - S_{u_{j+1}\cdots u_n} y|^\alpha \\ &\leq H(\operatorname{diam} X)^\alpha \sum_{k=0}^{n-1} \vartheta^{\alpha k} \\ &\leq D := \frac{H(\operatorname{diam} X)^\alpha}{1-\vartheta^\alpha}. \end{aligned} \tag{I.9.16}$$

Exponentiation proves (I.9.7) with $C_{\mathrm{bd}} = e^D$.

If K is a singleton, the two sides of (I.9.8)–(I.9.9) vanish and the cylinder assertions are immediate. We prove them for non-singleton K. Uniform $C^{1+\alpha}$ regularity and (I.9.13) give $\delta > 0$ and $L < \infty$ such that, whenever $x, y \in X$ and $0 < |x-y| \leq \delta$,

$$\left| \frac{|S_i(x) - S_i(y)|}{a_i(y)|x-y|} - 1 \right| \le L|x-y|^\alpha. \tag{I.9.17}$$

Indeed, the numerator before division differs from $|DS_i(y)(x-y)| = a_i(y)|x-y|$ by at most a common constant times $|x-y|^{1+\alpha}$. Choose k_0 so large that

$$\vartheta^{k_0} \operatorname{diam} K < \delta, \qquad L\vartheta^{\alpha k_0} (\operatorname{diam} K)^\alpha \le \frac{1}{2}. \tag{I.9.18}$$

For each of the finitely many words w with $|w| \le k_0$, compactness and injectivity give constants $0 < c_0 \le C_0 < \infty$ for which

$$c_0 a_w(y)|x-y| \le |S_w(x) - S_w(y)| \le C_0 a_w(y)|x-y| \tag{I.9.19}$$

for $x, y \in K$. For each of these finitely many diffeomorphisms S_w, uniform differentiability near K and the positive minimum of its conorm give the analogue of (I.9.17) near the diagonal. Take the minimum of the finitely many resulting thresholds and bounds. Away from the diagonal, the quotient in (I.9.19) has a positive minimum on the compact set of pairs a fixed positive distance from the diagonal.

Now let $|u| > k_0$. Apply (I.9.17) successively to all outer letters after the last k_0 letters have acted. At the step with a tail of length k, the relative error has absolute value at most

$$\epsilon_k = L\vartheta^{\alpha k} (\operatorname{diam} K)^\alpha, \qquad k \ge k_0. \tag{I.9.20}$$

The series $\sum_{(k \ge k_0)} \varepsilon_k$ converges and every $\varepsilon_k \le 1/2$. Therefore

$$0 < \prod_{k=k_0}^{\infty} (1 - \epsilon_k) \le \prod (1 + \delta_k) \le \prod_{k=k_0}^{\infty} (1 + \epsilon_k) < \infty \tag{I.9.21}$$

whenever $|\delta_k| \le \varepsilon_k$. Combining (I.9.19)–(I.9.21) with the chain identity (I.9.15) yields

$$c_1 a_u(y)|x-y| \le |S_u(x) - S_u(y)| \le C_1 a_u(y)|x-y|, \tag{I.9.22}$$

uniformly in u, x, y. Equation (I.9.16) replaces y by any $z \in X$, proving (I.9.8).

Taking the supremum over $x, y \in K$ in (I.9.8) proves (I.9.9). The chain rule gives $q_{(uv)} \le q_u q_v$. If x realizes q_v, then (I.9.7) gives

$$a_{uv}(x) = a_u(S_v x) a_v(x) \ge C_{\mathrm{bd}}^{-1} q_u q_v, \tag{I.9.23}$$

which proves (I.9.10).

If ω and τ agree through coordinate $n \ge 1$, then

$$|\pi(\sigma\omega) - \pi(\sigma\tau)| \le (\operatorname{diam} K) \vartheta^{n-1}. \tag{I.9.24}$$

Equations (I.9.14) and (I.9.24) show that ϕ is Hölder for the standard prefix metric. Finally, if $\omega|_n = \tau|_n = u$, then conformality gives

$$S_n\phi(\omega) = \log a_u(\pi(\sigma^n \omega)), \quad S_n\phi(\tau) = \log a_u(\pi(\sigma^n \tau)). \tag{I.9.25}$$

The bound (I.9.12) follows from (I.9.16). This completes the proof. □

I.9.2 Pressure and geometric Gibbs states

Definition — 2.1 (partition sum, pressure, equilibrium, and Gibbs property)

For a continuous $\varphi : \Sigma \to R$, define

$$S_n\varphi(\omega) = \sum_{j=0}^{n-1} \varphi(\sigma^j \omega), \quad Z_n(\varphi) = \sum_{u \in \Lambda^n} \exp\left(\sup_{\omega \in [u]} S_n\varphi(\omega)\right). \tag{I.9.26}$$

Its pressure is

$$P(\varphi) = \lim_{n\to\infty} \frac{1}{n} \log Z_n(\varphi), \tag{I.9.27}$$

once existence is proved. An invariant probability m is an equilibrium state when

$$h_m(\sigma) + \int \varphi \, dm = P(\varphi). \tag{I.9.28}$$

It is a Gibbs measure when some $C_G \geq 1$ gives

$$C_G^{-1} \leq \frac{m([u])}{\exp(S_n\varphi(\omega) - nP(\varphi))} \leq C_G \tag{I.9.29}$$

for $u \in \Lambda^n$ and $\omega \in [u]$.
For the geometric potential (I.9.11), set

$$P_\Phi(t) = P(t\phi). \tag{I.9.30}$$

The unique zero of this function is denoted s_Φ and called the pressure dimension.

Theorem — 2.2 (T-I.9.02: pressure and Gibbs package) [F]

Provenance: the allowed *Dynamical Systems Foundations*, reader edition F01–F16, Theorem F14.4 (printed pp. 388–392), supplies the equilibrium/Gibbs theorem. The qualitative consequence of Theorem F14.7 (printed pp. 406–407) supplies mixing, as explained below. The geometric statement and every dimensional deduction are local. These locators refer to the 14 September 2026 reader edition of that companion textbook.
The limit (I.9.27) exists for every continuous potential on the finite full shift. For the geometric potential of Lemma 1.3, $P_\Phi : R \to R$ is finite, continuous, and strictly decreasing, and has a unique zero $s_\Phi \geq 0$.
Let m_t be the equilibrium/Gibbs state for $t\phi$ furnished by the exact A-DS theorem just cited. It is mixing and

$$P_\Phi(t) = h_{m_t}(\sigma) - t\chi_\Phi(m_t), \quad \chi_\Phi(m_t) = -\int \phi \, dm_t > 0, \tag{I.9.31}$$

while uniformly in $u \in \Lambda^n$, $x \in X$,

$$m_t([u]) \asymp e^{-nP_\Phi(t)} a_u(x)^t. \tag{I.9.32}$$

At the pressure zero,

$$h_{m_{s_\Phi}}(\sigma) = s_\Phi \chi_\Phi(m_{s_\Phi}). \tag{I.9.33}$$

Proof

The Birkhoff decomposition gives

$$\sup_{[uv]} S_{n+k}\varphi \le \sup_{[u]} S_n\varphi + \sup_{[v]} S_k\varphi. \tag{I.9.34}$$

Consequently $a_n = \log Z_n(\varphi)$ is subadditive. Put $a_0 = 0$. For fixed $k \ge 1$, write $n = qk + r$, where $0 \le r < k$. Repeated subadditivity gives

$$\frac{a_n}{n} \le \frac{qa_k + a_r}{qk + r}. \tag{I.9.34a}$$

The finitely many possible a_r are bounded, so $\limsup_n a_n/n \le a_k/k$. Take the infimum over k; the reverse inequality $\inf_k a_k/k \le \liminf_n a_n/n$ follows directly from the definition of the infimum. Hence

$$P(\varphi) = \inf_{n\ge 1} \frac{1}{n} \log Z_n(\varphi), \tag{I.9.35}$$

which proves existence and finiteness.

For finiteness, explicitly,

$$\log N - \|\varphi\|_\infty \le \frac{1}{n} \log Z_n(\varphi) \le \log N + \|\varphi\|_\infty.$$

For bounded potentials,

$$|P(\varphi) - P(\psi)| \le \|\varphi - \psi\|_\infty. \tag{I.9.36}$$

Indeed, the corresponding partition sums differ by factors between $e^{-n\|\varphi-\psi\|}$ and $e^{n\|\varphi-\psi\|}$. Hence P_Φ is continuous. By (I.9.3) and compactness there are constants

$$-\infty < \log a_* \le \phi(\omega) \le \log \vartheta < 0. \tag{I.9.37}$$

If $t > s$, then

$$P_\Phi(t) \le P_\Phi(s) + (t - s)\log\vartheta < P_\Phi(s), \tag{I.9.38}$$

so the function is strictly decreasing. Moreover,

$$P_\Phi(0) = \log N, \qquad P_\Phi(t) \le \log N + t\log\vartheta. \tag{I.9.39}$$

For $N \ge 2$, continuity and (I.9.39) give exactly one positive zero. For $N = 1$, $P_\Phi(0) = 0$ and strict decrease gives the unique zero $s_\Phi = 0$.

Lemma 1.3 proves that $t\phi$ is Hölder. The A-DS theorem cited in the statement above supplies the unique equilibrium state m_t and (I.9.29). Its cylinder-sum pressure agrees with (I.9.27): on every cylinder the difference between a representative Birkhoff sum and its supremum is bounded independently of n, by Lemma 1.3, so the sums have the same exponential rate.

For completeness, the passage from the allowed qualitative correlation limit to measure-theoretic mixing is as follows. Theorem F14.7 in A-DS implies, for Hölder f and bounded measurable g,

$$\int f(g \circ \sigma^n)\, dm_t \longrightarrow \left(\int f\, dm_t\right)\left(\int g\, dm_t\right).$$

No rate is needed here. For arbitrary Borel sets $A, B \subset \Sigma$, choose a finite union C of cylinders with $m_t(AtriangleC) < \varepsilon$. Such approximations exist because the finite-cylinder algebra generates $B(\Sigma)$. The indicator of C is locally constant and therefore Hölder. Apply the correlation limit with $f = 1_C$ and $g = 1_B$. Replacing C by A changes the intersection measure and the product of measures by at most ε each, giving

$$\limsup_{n\to\infty} |m_t(A \cap \sigma^{-n} B) - m_t(A) m_t(B)| \le 2\epsilon.$$

Letting ε decrease to zero proves mixing and hence ergodicity. Substitution of $\varphi = t\phi$ into (I.9.28) gives (I.9.31). Equations (I.9.25), (I.9.29), and derivative bounded distortion give (I.9.32), including negative t. Finally set $t = s_\Phi$ in (I.9.31) to obtain (I.9.33). No other thermodynamic theorem is used. □

I.9.3 Euclidean differentiation at moving centers

The proof of exact dimensionality needs differentiation for an arbitrary finite Borel measure, not only for Lebesgue measure. We prove the exact form used later.

Lemma — 3.1 (Besicovitch selection in $\mathbb{R}^d$) [F]

There is an integer $B_d < \infty$ with the following property. Let E be a bounded subset of $\mathbb{R}^d$, and let $\mathscr{B}$ be a family of nondegenerate closed balls with uniformly bounded radii such that every $x \in E$ is the center of at least one ball in $\mathscr{B}$. There is a finite or countable subfamily $\mathscr{G}$ which covers E and satisfies

$$\sum_{B \in \mathscr{G}} \mathbf{1}_B(y) \le B_d \quad (y \in \mathbb{R}^d). \tag{I.9.40}$$

One may take $B_d = 15^d$.

Proof

Choose one ball $B(x, r(x))$ from $\mathscr{B}$ for every $x \in E$. It is enough to select from this assigned family. We begin with the geometric estimate which controls interactions between different selection stages. Call a finite family $(B(x_j, r_j))_{(j=1)}^M$ a Besicovitch family if it has a common point and no center belongs to any other ball. Thus distinct centers satisfy

$$|x_i - x_j| > r_i, \qquad |x_i - x_j| > r_j. \tag{I.9.41}$$

Let y be a common point. If some $x_j = y$, then (I.9.41) permits no other member, so suppose that every $a_j = |x_j - y|$ is positive and put $v_j = (x_j - y)/a_j$. If, say, $a_i \geq a_j$, then $a_i \leq r_i$ and the cosine rule give

$$a_i^2 \leq r_i^2 < |x_i - x_j|^2 = a_i^2 + a_j^2 - 2a_i a_j v_i \cdot v_j, \tag{I.9.42}$$

and hence $v_i \cdot v_j < a_j/(2a_i) \leq 1/2$. Therefore $|v_i - v_j| > 1$. The balls $B(v_j, 1/2)$ are pairwise disjoint and their union is contained in $B(0, 3/2)$. Comparing Euclidean volumes proves

$$M\,|B(0,1/2)| \leq |B(0,3/2)|, \qquad M \leq 3^d. \tag{I.9.42a}$$

We now construct the covering. Put $E_0 = E$. If $E_{(k-1)}$ is nonempty, set

$$R_k = \sup\{r(x) : x \in E_{k-1}\}. \tag{I.9.42b}$$

Starting with $W_1 = E_{(k-1)}$, choose $x_j \in W_j$ with $r(x_j) > R_k/2$, whenever such a point exists, and set

$$W_{j+1} = W_j \setminus B(x_j, r(x_j)). \tag{I.9.42c}$$

This loop terminates after finitely many choices. Indeed, for $i < j$, $x_j \notin B[x_i, r(x_i)]$, and consequently

$$|x_i - x_j| > r(x_i) > R_k/2. \tag{I.9.42d}$$

An infinite $R_k/2$-separated subset of the bounded set E cannot exist: the balls $B(x_j, R_k/4)$ would be disjoint, have the same positive volume, and lie in one bounded ball. Denote the finite set chosen at stage k by S_k, and put

$$E_k = E_{k-1} \setminus \bigcup_{x \in S_k} B(x, r(x)). \tag{I.9.42e}$$

When the loop stops, every $z \in E_k$ satisfies $r(z) \leq R_k/2$. Thus, if E_k is nonempty, $R_{(k+1)} \leq R_k/2$. Moreover, for $x \in S_k$ and $z \in E_k$, the construction and $r(x) > R_k/2 \geq r(z)$ give

$$|x - z| > r(x) = \max\{r(x), r(z)\}. \tag{I.9.42f}$$

The selected balls cover E. Otherwise a fixed uncovered $z \in E$ would belong to every E_k, whence $0 < r(z) \leq R_k/2 \leq R_1/2^k$ for all k, a contradiction. The selected family $\mathscr{G} = \{B(x, r(x)) : x \in S_k,\ k \geq 1\}$ is finite or countable.

It remains to prove its multiplicity bound. Fix $y \in \mathbb{R}^d$. At one stage k, the centers of all selected balls containing y lie in $B(y, R_k)$ and, by (I.9.42d), are mutually more than $R_k/2$ apart. The balls of radius $R_k/4$ about those centers are disjoint and lie in $B(y, 5R_k/4)$; volume comparison therefore shows that at most 5^d members from S_k contain y.

Let $\Gamma(y)$ be the set of stages contributing at least one such member, and choose one center $x_k \in S_k$ for every $k \in \Gamma(y)$. For two different stages, (I.9.42f), applied when the earlier stage was completed, shows that neither selected center belongs to the other selected ball. All these balls contain y, so (I.9.42a) gives $|\Gamma(y)| \le 3^d$. Consequently

$$\sum_{B \in \mathcal{G}} \mathbf{1}_B(y) \le 3^d 5^d = 15^d. \tag{I.9.42g}$$

This is (I.9.40) with $B_d = 15^d$. □

Lemma — 3.2 (centered differentiation for finite Borel measures) [F]

Let ν be a finite Borel measure on $\mathbb{R}^d$ and $f \in L^1(\nu)$. Then

$$\lim_{r \downarrow 0} \frac{1}{\nu(B(x,r))} \int_{B(x,r)} f \, d\nu = f(x) \tag{I.9.43}$$

for ν-almost every x. Ratios with zero denominator are set equal to zero; for ν-almost every $x \in \operatorname{spt}(\nu)$, all displayed denominators are positive.

Proof

For $g \in L^1(\nu)$ define

$$\mathscr{M}_0 g(x) = \limsup_{r \downarrow 0} \frac{1}{\nu(B(x,r))} \int_{B(x,r)} |g| \, d\nu. \tag{I.9.44}$$

Choose a Borel representative of g; changing it on a ν-null set does not change any of its ball integrals. This function is Borel measurable. For fixed $r > 0$, both $x \mapsto \int_(B(x,r))|g|d\nu$ and $x \mapsto \nu(B(x,r))$ are Borel by parameter integration of the Borel function $1_{(|x-y| \le r)}$. Continuity from above as rational radii decrease to r shows that the limsup in (I.9.44) may be computed using rational radii. Outside $\operatorname{spt}(\nu)$ the zero-denominator convention is eventually constant.

Fix $t > 0$ and a compact subset F of $\{\mathscr{M}_0 g > t\}$. At every point of F there are arbitrarily small centered balls B with

$$t\nu(B) < \int_B |g| \, d\nu. \tag{I.9.45}$$

Apply Lemma 3.1 to such balls. Since the chosen family covers F and has multiplicity at most B_d,

$$\begin{aligned} t\nu(F) &\le t \sum_{B \in \mathcal{G}} \nu(B) \le \sum_{B \in \mathcal{G}} \int_B |g| \, d\nu \\ &\le B_d \|g\|_{L^1(\nu)}. \end{aligned} \tag{I.9.46}$$

Regularity of ν and exhaustion by compact subsets give

$$\nu\{\mathscr{M}_0 g > t\} \le \frac{B_d}{t} \|g\|_{L^1(\nu)}. \tag{I.9.47}$$

Choose $h \in C_c(\mathbb{R}^d)$ with $\|f - h\|_1 < \varepsilon$; this density is ARA-KB1, Proposition 9.9(a). Uniform continuity gives

$$\lim_{r\downarrow 0} \frac{1}{\nu(B(x,r))} \int_{B(x,r)} h\, d\nu = h(x) \tag{I.9.48}$$

for $x \in \operatorname{spt}(\nu)$. Therefore the limsup of the absolute difference in (I.9.43) is at most

$$\mathscr{M}_0(f-h)(x) + |f(x) - h(x)|. \tag{I.9.49}$$

Equations (I.9.47) and Chebyshev's inequality show that the set where (I.9.49) exceeds $2t$ has measure at most

$$\frac{(B_d + 1)\epsilon}{t}. \tag{I.9.50}$$

Let $\varepsilon \downarrow 0$ and then take rational $t > 0$. This proves (I.9.43). $\square$

Lemma — 3.3 (conditional density and logarithmic envelope) [F]

Let (Ω, F, m) be a probability space, let $\theta : \Omega \to \mathbb{R}^d$ be measurable, and put

$$B_\theta(x,r) = \theta^{-1} B(\theta x, r). \tag{I.9.51}$$

For every $A \in F$,

$$\lim_{r\downarrow 0} \frac{m(A \cap B_\theta(x,r))}{m(B_\theta(x,r))} = \mathbb{E}_m(\mathbf{1}_A \mid \theta^{-1}\mathscr{B}(\mathbb{R}^d))(x) \tag{I.9.52}$$

for m-almost every x.
Let ξ be a finite partition. Define

$$G_{\xi,\theta}(x) = -\log \inf_{0<r<1} \frac{m(\xi(x) \cap B_\theta(x,r))}{m(B_\theta(x,r))}. \tag{I.9.53}$$

Then $G_{(\xi,\theta)} \geq 0$, it is measurable, and

$$\int G_{\xi,\theta}\, dm \leq H_m(\xi) + 1 + \log B_d. \tag{I.9.54}$$

In particular, when $A = \xi(x)$, its negative logarithm in (I.9.52), restricted to $0 < r < 1$, is dominated by $G_{(\xi,\theta)}(x)$ almost everywhere. For a fixed event A, use the partition $\{A, \Omega \setminus A\}$ and restrict this assertion to $x \in A$. No such assertion is made on the complement: the conditional probability of A can vanish there.

Proof

Let $\nu = \theta_* m$ and $\nu_A = \theta_*(1_A m)$. Since $\nu_A \ll \nu$, ARA-KB1, Theorem 9.16 gives $f_A = d\nu_A/d\nu$. For every Borel $D \subset \mathbb{R}^d$,

$$\int_{\theta^{-1}D} f_A \circ \theta\, dm = \int_D f_A\, d\nu = \nu_A(D) = m(A \cap \theta^{-1}D). \tag{I.9.55}$$

Thus the defining property of conditional expectation gives

$$f_A \circ \theta = \mathbb{E}_m(\mathbf{1}_A \mid \theta^{-1}\mathscr{B}(\mathbb{R}^d)) \quad m\text{-a.e.} \tag{I.9.56}$$

Lemma 3.2 applied to $f_A \in L^1(\nu)$ proves (I.9.52).

For a cell $C \in \xi$, let

$$g_C(z) = \inf_{0<r<1} \frac{\nu_C(B(z,r))}{\nu(B(z,r))}, \qquad \nu_C = \theta_*(\mathbf{1}_C m). \tag{I.9.57}$$

Fix $0 < t < 1$ and $R > 0$, and put $E_R = \{z \in \operatorname{spt}(\nu) \cap B(0,R) \,:\, g_C(z) < t\}$. For every $z \in E_R$, the definition of the infimum supplies a radius $0 < r_z < 1$ such that

$$\nu_C(B(z,r_z)) < t\nu(B(z,r_z)). \tag{I.9.57a}$$

Apply Lemma 3.1 to these witness balls. If $\mathscr{G}_R$ is the selected cover, its multiplicity is at most B_d; hence

$$\nu_C(E_R) \le \sum_{B\in\mathscr{G}_R} \nu_C(B) \le t \sum_{B\in\mathscr{G}_R} \nu(B) \le tB_d\nu(\mathbb{R}^d) = tB_d. \tag{I.9.57b}$$

Letting $R \to \infty$ and ignoring the ν_C-null complement of $\operatorname{spt}(\nu)$ gives

$$\nu_C\{z \,:\, g_C(z) < t\} \le B_d t. \tag{I.9.58}$$

Since (I.9.53) equals

$$-\sum_{C\in\xi} \mathbf{1}_C(x) \log g_C(\theta x), \tag{I.9.59}$$

the layer-cake formula and (I.9.58) give

$$\begin{aligned} \int G_{\xi,\theta}\, dm &= \int_0^\infty \sum_{C\in\xi} m(C \cap \{g_C \circ \theta < e^{-t}\})\, dt \\ &\le \sum_{C\in\xi} \int_0^\infty \min\{m(C), B_d e^{-t}\}\, dt \\ &= \sum_{C\in\xi} m(C)\left(1 + \log \frac{B_d}{m(C)}\right) \\ &= H_m(\xi) + 1 + \log B_d. \end{aligned} \tag{I.9.60}$$

Terms with $m(C) = 0$ are zero. Dyadic outer approximations of moving balls show measurability of the ratios and their infimum; equivalently, restrict the infimum to rational radii using continuity from above. This proves every claim. □

Lemma — 3.4 (L-I.9.03: differentiation package) [F]

Lemmas 3.1–3.3 together prove L-I.9.03. □

I.9.4 Conditional information and projection entropy

Definition — 4.1 (conditional information)

The projection-entropy normalization below is FH-01, Definition 2.1.
For a finite measurable partition ξ and sub-sigma-algebra $\mathscr{A}$, put

$$p_C^{\mathscr{A}} = \mathbb{E}_m(\mathbf{1}_C \mid \mathscr{A}), \tag{I.9.61}$$

and define

$$I_m(\xi \mid \mathscr{A}) = -\sum_{C\in\xi} \mathbf{1}_C \log p_C^{\mathscr{A}}, \quad H_m(\xi \mid \mathscr{A}) = \int I_m(\xi \mid \mathscr{A})\, dm. \tag{I.9.62}$$

Here $0 \log 0 = 0$; on C, the event $p_C^A = 0$ has measure zero by the defining integral identity for conditional expectation.
These information functions are integrable. Indeed, conditional expectation and finite-partition entropy give

$$H_m(\xi \mid \mathscr{A}) = \sum_{C\in\xi} \int -p_C^{\mathscr{A}} \log p_C^{\mathscr{A}}\, dm \le -\sum_{C\in\xi} m(C) \log m(C) = H_m(\xi) < \infty,$$

where the inequality is Jensen's inequality for the concave function $-t \log t$ on $[0,1]$. In particular, the difference in (I.9.64) belongs to $L^1(m)$, so its conditional expectation in (I.9.65) is well defined.
Let

$$\mathscr{A}_\pi = \pi^{-1}\mathscr{B}(\mathbb{R}^d), \qquad \mathscr{I} = \{E : \sigma^{-1}E = E \pmod{m}\}. \tag{I.9.63}$$

Define

$$\Delta_\pi = I_m(\mathscr{P} \mid \sigma^{-1}\mathscr{A}_\pi) - I_m(\mathscr{P} \mid \mathscr{A}_\pi), \tag{I.9.64}$$

$$h_\pi(\sigma, m) = \int \Delta_\pi\, dm, \qquad h_\pi(\sigma, m, \omega) = \mathbb{E}_m(\Delta_\pi \mid \mathscr{I})(\omega). \tag{I.9.65}$$

These are the projection entropy and local projection entropy.

Lemma — 4.2 (finite conditional-information calculus) [F]

For finite partitions ξ, η on a probability space and a sub-sigma-algebra $\mathscr{A}$, the chain rule is

$$I_m(\xi \vee \eta \mid \mathscr{A}) = I_m(\xi \mid \mathscr{A}) + I_m(\eta \mid \hat{\xi} \vee \mathscr{A}), \tag{I.9.66}$$

For the next identity, let $T : (\Omega, F) \to (Y, G)$ be measurable, let m be a probability on Ω, and let ξ and $\mathscr{A}$ be respectively a finite measurable partition and a sub-sigma-algebra on Y. The pull-back rule is

$$I_{m\circ T^{-1}}(\xi \mid \mathscr{A}) \circ T = I_m(T^{-1}\xi \mid T^{-1}\mathscr{A}). \tag{I.9.67}$$

The corresponding entropy identities follow by integration. If the sub-sigma-algebras $\mathscr{A}_n$ increase and $\mathscr{A}$ is generated by their union, modulo completion under m, then

$$H_m(\xi \mid \mathscr{A}_n) \longrightarrow H_m(\xi \mid \mathscr{A}). \tag{I.9.68}$$

Proof

For $C \in \xi$, define on C

$$q_D = \begin{cases} \mathbb{E}(\mathbf{1}_{C\cap D} \mid \mathscr{A})/p_C^{\mathscr{A}}, & p_C^{\mathscr{A}} > 0, \\ 0, & p_C^{\mathscr{A}} = 0. \end{cases} \tag{I.9.69}$$

Testing against sets of the form $C \cap A$, $A \in \mathscr{A}$, shows that the function obtained by putting these ratios on all cells C is a version of $E(1_D|\hat{\xi} \vee \mathscr{A})$. Thus, on $C \cap D$,

$$\mathbb{E}(\mathbf{1}_{C\cap D} \mid \mathscr{A}) = p_C^{\mathscr{A}}\,\mathbb{E}(\mathbf{1}_D \mid \hat{\xi} \vee \mathscr{A}). \tag{I.9.70}$$

Taking $-\log$ and summing over C, D proves (I.9.66). The defining integral identity after a change of variables proves (I.9.67).

For (I.9.68), conditional expectations in L^2 are orthogonal projections. For $p_n = E_m(1_C|\mathscr{A}_n)$ and $j \geq n$, the projection identity gives

$$\|p_j - p_n\|_2^2 = \|p_j\|_2^2 - \|p_n\|_2^2.$$

The squared norms increase and are bounded by $\|1_C\|_2^2$, so (p_n) is Cauchy. Its limit belongs to the closure of the union of these spaces, and $1_C - p_n$ is orthogonal to every fixed earlier space once n is large. Thus the limit is the orthogonal projection onto that closure. Indicators from the algebra $\cup_n \mathscr{A}_n$ are dense in $L^2(\mathscr{A})$, so the latter projection is $E(1_C|\mathscr{A})$. With $g(t) = -t\log t$,

$$H_m(\xi \mid \mathscr{A}_n) = \sum_{C\in\xi} \int g(p_C^{\mathscr{A}_n})\,dm. \tag{I.9.71}$$

The function g is bounded and uniformly continuous on $[0, 1]$; L^2 convergence therefore makes every integral in (I.9.71) converge. This proves (I.9.68). □

Lemma — 4.3 (coding sigma-algebra and block identity) [F]

One has

$$\widehat{\mathscr{P}} \vee \sigma^{-1}\mathscr{A}_\pi = \widehat{\mathscr{P}} \vee \mathscr{A}_\pi. \tag{I.9.72}$$

For every invariant probability m and every $n \geq 1$,

$$nh_\pi(\sigma, m) = H_m(\mathscr{P}_0^{n-1} \mid \sigma^{-n}\mathscr{A}_\pi) - H_m(\mathscr{P}_0^{n-1} \mid \mathscr{A}_\pi). \tag{I.9.73}$$

Proof

Every generator of the left side of (I.9.72) is assembled from sets

$$[i] \cap \sigma^{-1}\pi^{-1}D = [i] \cap \pi^{-1}(S_i(D \cap K)), \tag{I.9.74}$$

where D is Borel. Since $S_i : U \to S_i(U)$ is a diffeomorphism, $S_i(D \cap K)$ is Borel. Conversely,

$$[i] \cap \pi^{-1}D = [i] \cap \sigma^{-1}\pi^{-1}(S_i^{-1}(D \cap S_i(K))), \tag{I.9.75}$$

with the inverse image restricted to $S_i(U)$. Equations (I.9.74)–(I.9.75) prove (I.9.72).

Put $\mathscr{A}_j = \sigma^{-j}\mathscr{A}_\pi$ and

$$\mathscr{Q}_j = \bigvee_{\substack{0 \le i < n \\ i \ne j}} \sigma^{-i}\mathscr{P}.$$

For $0 \le j < n$, the chain rule gives

$$I_m(\mathscr{P}_0^{n-1} \mid \mathscr{A}_j) = I_m(\sigma^{-j}\mathscr{P} \mid \mathscr{A}_j) + I_m(\mathscr{Q}_j \mid \sigma^{-j}\widehat{\mathscr{P}} \vee \mathscr{A}_j),$$
$$I_m(\mathscr{P}_0^{n-1} \mid \mathscr{A}_{j+1}) = I_m(\sigma^{-j}\mathscr{P} \mid \mathscr{A}_{j+1}) + I_m(\mathscr{Q}_j \mid \sigma^{-j}\widehat{\mathscr{P}} \vee \mathscr{A}_{j+1}).$$

Since

$$\sigma^{-j}\widehat{\mathscr{P}} \vee \mathscr{A}_j = \sigma^{-j}\widehat{\mathscr{P}} \vee \mathscr{A}_{j+1}, \tag{I.9.76}$$

the two Q_j terms are identical. Subtraction and the pull-back rule (I.9.67) therefore give

$$\begin{aligned} &I_m(\mathscr{P}_0^{n-1} \mid \mathscr{A}_j) - I_m(\mathscr{P}_0^{n-1} \mid \mathscr{A}_{j+1}) \\ &\quad = I_{m\circ\sigma^{-j}}(\mathscr{P} \mid \mathscr{A}_\pi) \circ \sigma^j - I_{m\circ\sigma^{-j}}(\mathscr{P} \mid \sigma^{-1}\mathscr{A}_\pi) \circ \sigma^j. \end{aligned} \tag{I.9.77}$$

Integrate (I.9.77). Invariance makes its right side equal $-h_\pi(\sigma, m)$. Summing over $j = 0, \ldots, n-1$ telescopes and gives

$$H_m(\mathscr{P}_0^{n-1} \mid \mathscr{A}_\pi) - H_m(\mathscr{P}_0^{n-1} \mid \sigma^{-n}\mathscr{A}_\pi) = -nh_\pi(\sigma, m), \tag{I.9.78}$$

which is (I.9.73). □

Lemma — 4.4 (L-I.9.04: coding calculus) [F]

Lemmas 4.2–4.3 prove L-I.9.04. □

Definition — 4.5 (fiber-prefix multiplicity)

For $z \in K$, define

$$N_n(z) = \#\{u \in \Lambda^n : z \in K_u\}, \qquad M_n = \sup_{z \in K} N_n(z). \tag{I.9.79}$$

The coding has uniformly subexponential fiber-prefix multiplicity when

$$\lim_{n\to\infty} \frac{1}{n} \log M_n = 0. \tag{I.9.80}$$

Theorem — 4.6 (T-I.9.05: projection-entropy structure) [F]

Provenance: FH-01, Definition 2.1, Theorem 2.2(i),(iii), and Corollary 4.16. The formulation and proof below are local.
For every invariant probability m,

$$0 \le h_\pi(\sigma, m) \le h_m(\sigma), \tag{I.9.81}$$

and

$$\lim_{n\to\infty} \frac{1}{n} H_m(\mathscr{P}_0^{n-1} \mid \mathscr{A}_\pi) = h_m(\sigma) - h_\pi(\sigma, m). \tag{I.9.82}$$

If (I.9.80) holds, then

$$h_\pi(\sigma, m) = h_m(\sigma). \tag{I.9.83}$$

Proof

Let $\mathscr{D}_k$ be the half-open dyadic cubes of side 2^{-k} in $\mathbb{R}^d$. Only finitely many meet K. For a fixed n, invariance and the conditional entropy identity give

$$\begin{aligned} &H_m(\mathscr{P}_0^{n-1} \mid \sigma^{-n}\pi^{-1}\widehat{\mathscr{D}_k}) - H_m(\mathscr{P}_0^{n-1} \mid \pi^{-1}\widehat{\mathscr{D}_k}) \\ &\quad = H_m(\mathscr{P}_0^{n-1} \vee \sigma^{-n}\pi^{-1}\mathscr{D}_k) - H_m(\mathscr{P}_0^{n-1} \vee \pi^{-1}\mathscr{D}_k). \end{aligned} \tag{I.9.84}$$

An atom in the first partition is of the form

$$[u] \cap \sigma^{-n}\pi^{-1}Q = [u] \cap \pi^{-1}(S_u(Q \cap K)). \tag{I.9.85}$$

The set $S_u(Q \cap K)$ has diameter at most $\sqrt{d}2^{-k}$ because S_u is a contraction. In each coordinate its range has length at most this diameter, and hence meets at most $\lceil\sqrt{d}\rceil + 2$ half-open dyadic intervals of length 2^{-k}. Thus it meets at most $C_d = (\lceil\sqrt{d}\rceil + 2)^d$ cubes in $\mathscr{D}_k$. Each atom in the first partition of (I.9.84) meets at most C_d atoms in the second.

If every atom of a partition α meets at most C_d atoms of β, the finite Shannon inequality gives

$$H_m(\alpha \vee \beta) - H_m(\alpha) = H_m(\beta \mid \hat{\alpha}) \le \log C_d. \tag{I.9.86}$$

Since $H_m(\beta) \le H_m(\alpha \vee \beta)$, (I.9.84)–(I.9.86) imply that the left side of (I.9.84) is at least $-\log C_d$. Let $k \to \infty$. The dyadic sigma-algebras increase to $B(\mathbb{R}^d)$, so (I.9.68) and (I.9.73) yield

$$nh_\pi(\sigma, m) \ge -\log C_d. \tag{I.9.87}$$

Since n is arbitrary, $h_\pi \ge 0$.

Equation (I.9.73) and nonnegativity of conditional entropy give

$$nh_\pi(\sigma, m) \le H_m(\mathscr{P}_0^{n-1} \mid \sigma^{-n}\mathscr{A}_\pi) \le H_m(\mathscr{P}_0^{n-1}). \tag{I.9.88}$$

Divide by n and use the finite-generator entropy formula from A-DS to obtain the upper bound in (I.9.81).

We next determine the first term in (I.9.73). Because $\mathscr{A}_\pi \subset B(\Sigma)$, conditioning monotonicity gives

$$H_m(\mathscr{P}_0^{n-1} \mid \sigma^{-n}\mathscr{A}_\pi) \ge H_m(\mathscr{P}_0^{n-1} \mid \sigma^{-n}\mathscr{B}(\Sigma)). \tag{I.9.89}$$

Applying the chain rule from the last coordinate backward gives

$$\begin{aligned} H_m(\mathscr{P}_0^{n-1} \mid \sigma^{-n}\mathscr{B}(\Sigma)) &= \sum_{j=0}^{n-1} H_m(\sigma^{-j}\mathscr{P} \mid \sigma^{-(j+1)}\mathscr{B}(\Sigma)) \\ &= nH_m(\mathscr{P} \mid \sigma^{-1}\mathscr{B}(\Sigma)) = nh_m(\sigma). \end{aligned} \tag{I.9.90}$$

The opposite bound by $H_m(P_0^{n-1})$, divided by n, has the same limit $h_m(\sigma)$. Hence

$$\lim_{n\to\infty} \frac{1}{n} H_m(\mathscr{P}_0^{n-1} \mid \sigma^{-n}\mathscr{A}_\pi) = h_m(\sigma). \tag{I.9.91}$$

Subtract (I.9.73) to obtain (I.9.82).

It remains to prove the multiplicity criterion. Put $\mu = \pi_* m$ and, for $u \in \Lambda^n$,

$$\nu_u = \pi_*(\mathbf{1}_{[u]} m), \qquad p_u = \frac{d\nu_u}{d\mu}. \tag{I.9.92}$$

Then $\sum_u p_u = 1$ for μ-almost every z, and $p_u(z) = 0$ outside K_u. The conditional-expectation identity (I.9.55) gives

$$H_m(\mathscr{P}_0^{n-1} \mid \mathscr{A}_\pi) = \int_K \left(- \sum_{u\in\Lambda^n} p_u(z) \log p_u(z) \right) d\mu(z). \tag{I.9.93}$$

At a fixed z, at most $N_n(z)$ summands are nonzero. Gibbs' finite inequality, proved in L-I.8.02, therefore gives

$$-\sum_u p_u(z) \log p_u(z) \le \log N_n(z) \le \log M_n. \tag{I.9.94}$$

Divide (I.9.93)–(I.9.94) by n, use (I.9.80), and compare with (I.9.82). This proves (I.9.83).
□

I.9.5 Moving-scale ergodic averaging

Lemma — 5.1 (dominated moving-array theorem) [F]

Let T preserve a probability m. Suppose $F_k \to F$ almost everywhere and

$$\sup_{k\geq 1} |F_k| \in L^1(m). \tag{I.9.95}$$

Then

$$\frac{1}{n}\sum_{j=0}^{n-1} F_{n-j}(T^j x) \longrightarrow \mathbb{E}_m(F \mid \mathscr{I}_T)(x) \tag{I.9.96}$$

for m-almost every x.

Proof

Let

$$H_N = \sup_{k\geq N} |F_k - F|. \tag{I.9.97}$$

The functions H_N decrease to zero almost everywhere and are dominated by an L^1 function. Hence $\|H_N\|_1 \to 0$. Split the difference between the left side of (I.9.96) and the ordinary Birkhoff average of F into indices $n-j \geq N$ and $n-j < N$. The first part is at most

$$\frac{1}{n}\sum_{j=0}^{n-1} H_N(T^j x), \tag{I.9.98}$$

whose limsup is $E(H_N|I_T)(x)$ by Birkhoff.
If $G \in L^1(m)$, then

$$\sum_{n=1}^{\infty} m\{|G| > \epsilon n\} \leq \int \sum_{n=1}^{\infty} \mathbf{1}_{\{|G|>\epsilon n\}}\, dm \leq \frac{\|G\|_1}{\epsilon}. \tag{I.9.99}$$

Indeed, invariance makes the sum of the measures of the events $\{|G \circ T^n| > \varepsilon n\}$ finite by (I.9.99). For every M, the measure of their union over $n \geq M$ is bounded by the corresponding tail of that convergent series, which tends to zero. Hence $G(T^n x)/n \to 0$ almost everywhere. Thus the fewer than N exceptional terms $n-j < N$ contribute zero in the limit. As $N \to \infty$, order preservation of conditional expectation shows that $E(H_N|I_T)$ decreases to a nonnegative limit Y. Moreover, $\int Y dm = \lim_N \int H_N dm = 0$, so $Y = 0$ almost everywhere. The ordinary Birkhoff averages of F converge to $E(F|I_T)$, proving (I.9.96). ☐

Lemma — 5.2 (terminal-ball estimate) [F]

Let $\mu = \pi_* m$, where m is invariant, and fix $r_0 > 0$. Then

$$\frac{1}{n}\log\mu(B(\pi(\sigma^n\omega), r_0)) \longrightarrow 0 \tag{I.9.100}$$

for m-almost every ω.

Proof

Cover K by L balls $B(z_i, r_0/2)$. Put

$$A_n = \{\omega : \mu(B(\pi\omega, r_0)) \le e^{-\epsilon n}\}. \tag{I.9.101}$$

If $A_n \cap \pi^{-1}B(z_i, r_0/2)$ is nonempty, choose ω in this set. The triangle inequality gives

$$B(z_i, r_0/2) \subset B(\pi\omega, r_0), \tag{I.9.102}$$

and hence $\mu(B(z_i, r_0/2)) \le e^{-\epsilon n}$. It follows that

$$m(A_n) \le Le^{-\epsilon n}. \tag{I.9.103}$$

Invariance gives the same bound for $\{\omega : \sigma^n omega \in A_n\}$. The sum of these bounds is finite. As in the tail-union argument following (I.9.99), almost every ω belongs to only finitely many of these events. Therefore

$$\liminf_{n\to\infty}\frac{1}{n}\log\mu(B(\pi(\sigma^n\omega), r_0)) \ge -\epsilon \tag{I.9.104}$$

almost everywhere. The logarithm is at most zero. Intersect over rational $\varepsilon > 0$ to prove (I.9.100). ☐

Lemma — 5.3 (L-I.9.06: moving-scale package) [F]

Lemmas 5.1–5.2 prove L-I.9.06. ☐

I.9.6 Exact dimensionality without separation

Definition — 6.1 (conformal Lyapunov observable)

Put

$$\rho(\omega) = \|DS_{\omega_1}(\pi(\sigma\omega))\|, \qquad \ell_\Phi(\omega) = -\log\rho(\omega). \tag{I.9.105}$$

There are constants

$$0 < \rho_* \le \rho(\omega) \le \rho^* < 1. \tag{I.9.106}$$

For invariant m, define

$$\lambda_\Phi(m, \omega) = \mathbb{E}_m(\ell_\Phi \mid \mathscr{I})(\omega) > 0. \tag{I.9.107}$$

If m is ergodic, this is the constant

$$\chi_\Phi(m) = \int \ell_\Phi \, dm. \tag{I.9.108}$$

Lemma — 6.2 (uniform one-step conformal balls) [F]

For every $c > 1$ there is $r_0 \in (0,1)$ such that, for $i \in \Lambda$, $x \in K$, and $0 < r \leq r_0$,

$$B(S_i x, c^{-1}\|DS_i(x)\|r) \subset S_i(B(x,r)) \subset B(S_i x, c\|DS_i(x)\|r). \tag{I.9.109}$$

Consequently, for every $\omega \in \Sigma$,

$$\begin{aligned} B_\pi(\omega, c^{-1}\rho(\omega)r) \cap \mathscr{P}(\omega) &\subset \sigma^{-1}B_\pi(\sigma\omega, r) \cap \mathscr{P}(\omega) \\ &\subset B_\pi(\omega, c\rho(\omega)r) \cap \mathscr{P}(\omega), \end{aligned} \tag{I.9.110}$$

where $B_\pi(\omega, r) = \pi^{-1}B(\pi\omega, r)$.

Proof

Uniform differentiability on the compact set X gives, after using conformality,

$$|S_i(y) - S_i(x)| \leq c^{1/2}\|DS_i(x)\||y - x| \tag{I.9.111}$$

for $|x - y|$ below a uniform threshold. Apply the same estimate to S_i^{-1} near $S_i(X)$. Since

$$\|D(S_i^{-1})(S_i x)\| = \|DS_i(x)\|^{-1}, \tag{I.9.112}$$

the inverse estimate shows that every point of $B(S_i x, c^{-1}\|DS_i(x)\|r)$ has preimage in $B(x,r)$ after decreasing r_0; the two factors $c^{1/2}$ are absorbed by c. This proves (I.9.109).

On the cylinder $[\omega_1]$,

$$\pi(\tau) = S_{\omega_1}(\pi(\sigma\tau)). \tag{I.9.113}$$

Use injectivity in (I.9.109) to pull both inclusions back through (I.9.113). This gives (I.9.110). ☐

Lemma — 6.3 (geometrically spaced radii detect full local dimensions) [F]

Suppose $r_n \downarrow 0$ and constants $0 < a \leq b < 1$ satisfy

$$a \leq \frac{r_{n+1}}{r_n} \leq b. \tag{I.9.114}$$

For every finite Borel measure ν and $x \in \operatorname{spt}(\nu)$,

$$\liminf_{r\downarrow 0} \frac{\log \nu(B(x,r))}{\log r} = \liminf_{n\to\infty} \frac{\log \nu(B(x,r_n))}{\log r_n}, \tag{I.9.115}$$

and the analogous equality holds for lim sup.

Proof

Multiplying ν by a positive constant does not change either side of (I.9.115), because the added constant in the numerator is divided by $\log r \to -\infty$. We may therefore assume $\nu(\mathbb{R}^d) \le 1$.

If $r_{(n+1)} < r \le r_n$, monotonicity gives

$$\nu(B(x, r_{n+1})) \le \nu(B(x, r)) \le \nu(B(x, r_n)). \tag{I.9.116}$$

Moreover, (I.9.114) implies

$$\frac{\log r_{n+1}}{\log r_n} \longrightarrow 1. \tag{I.9.117}$$

All mass logarithms are now nonpositive. With $Q(r) = \log \nu(B(x, r))/\log r$, the two monotonicities in (I.9.116) give the exact squeeze

$$\frac{\log \nu(B(x, r_n))}{\log r_{n+1}} \le Q(r) \le \frac{\log \nu(B(x, r_{n+1}))}{\log r_n}. \tag{I.9.117a}$$

Each endpoint is the corresponding radius-sequence quotient multiplied by one of the radius-log ratios in (I.9.117), which tend to one. Taking the two liminf and limsup squeezes proves (I.9.115) and its lim sup counterpart. □

Theorem — 6.4 (T-I.9.07: two conformal scale inequalities) [F]

Provenance: FH-01, Theorem 2.6 and Sections 3, 5, and 6. This conformal specialization is proved locally.

Let Φ be a finite C^1 conformal IFS, let m be invariant, and put $\mu = \pi_* m$. For m-almost every ω,

$$\overline{d}_\mu(\pi\omega) \le \frac{h_\pi(\sigma, m, \omega)}{\lambda_\Phi(m, \omega)}, \tag{I.9.118}$$

$$\underline{d}_\mu(\pi\omega) \ge \frac{h_\pi(\sigma, m, \omega)}{\lambda_\Phi(m, \omega)}. \tag{I.9.119}$$

Proof

Fix $c > 1$ so small that

$$c\rho^* < 1. \tag{I.9.120}$$

Let r_0 be supplied by Lemma 6.2 and, after decreasing it, assume $r_0 < 1$. Define

$$\rho_n(\omega) = \prod_{j=0}^{n-1} \rho(\sigma^j \omega), \tag{I.9.121}$$

and the two moving radii

$$R_n^+(\omega) = r_0 c^n \rho_n(\omega), \qquad R_n^-(\omega) = r_0 c^{-n} \rho_n(\omega), \qquad R_0^\pm = r_0. \tag{I.9.122}$$

For $k \geq 1$, set

$$A_k^\pm(\omega) = \log \frac{m(B_\pi(\omega, R_k^\pm(\omega)) \cap \mathscr{P}(\omega))}{m(B_\pi(\omega, R_k^\pm(\omega)))}, \tag{I.9.123}$$

and

$$C_k^\pm(\omega) = \log \frac{m(\sigma^{-1} B_\pi(\sigma\omega, R_{k-1}^\pm(\sigma\omega)) \cap \mathscr{P}(\omega))}{m(\sigma^{-1} B_\pi(\sigma\omega, R_{k-1}^\pm(\sigma\omega)))}. \tag{I.9.124}$$

All the moving radii are positive Borel functions. Their ball masses are measurable by parameter integration of $1_(|\pi(\tau) - \pi(\omega)| \leq R_k^p m(\omega))$, and likewise with $\pi \circ \sigma$. Lemma 3.3 supplies a full-measure set on which both cell-specific envelopes are finite; every ratio above is then positive. Intersect this set and the sets for the subsequent almost-everywhere limits with all their inverse images under σ^j, $j \geq 0$. Invariance preserves full measure, and every orbit point used in the recursions lies in the resulting set. Lemma 3.3, first with $\theta = \pi$ and then with $\theta = \pi \circ \sigma$, gives

$$A_k^\pm \longrightarrow -I_m(\mathscr{P} \mid \mathscr{A}_\pi), \quad C_k^\pm \longrightarrow -I_m(\mathscr{P} \mid \sigma^{-1}\mathscr{A}_\pi) \tag{I.9.125}$$

almost everywhere. Moreover,

$$|C_k^\pm - A_k^\pm| \leq G_{\mathscr{P},\pi\circ\sigma} + G_{\mathscr{P},\pi} \in L^1(m). \tag{I.9.126}$$

Thus Lemma 5.1 applies to $F_k^p m = C_k^p m - A_k^p m$. Equations (I.9.64)–(I.9.65) yield

$$\frac{1}{n} \sum_{j=0}^{n-1} F_{n-j}^\pm(\sigma^j \omega) \longrightarrow -h_\pi(\sigma, m, \omega). \tag{I.9.127}$$

We first use the large radii R_n^+. Substituting $r = R_{(k-1)}^+(\sigma\omega)$ in the right inclusion of (I.9.110) gives

$$\sigma^{-1} B_\pi(\sigma\omega, R_{k-1}^+(\sigma\omega)) \cap \mathscr{P}(\omega) \subset B_\pi(\omega, R_k^+(\omega)) \cap \mathscr{P}(\omega). \tag{I.9.128}$$

After multiplying each whole-ball mass by its conditional cylinder fraction, (I.9.128) becomes

$$\begin{aligned} &\log m(B_\pi(\omega, R_k^+(\omega))) - \log m(B_\pi(\sigma\omega, R_{k-1}^+(\sigma\omega))) \\ &\qquad \geq C_k^+(\omega) - A_k^+(\omega) = F_k^+(\omega), \end{aligned} \tag{I.9.129}$$

where invariance identifies the denominator in (I.9.124) with the second whole-ball mass. Iterating (I.9.129) gives

$$\begin{aligned} \log \mu(B(\pi\omega, R_n^+(\omega))) \geq &\sum_{j=0}^{n-1} F_{n-j}^+(\sigma^j \omega) \\ &+ \log \mu(B(\pi(\sigma^n \omega), r_0)). \end{aligned} \tag{I.9.130}$$

Lemmas 5.2 and (I.9.127) imply

$$\liminf_{n\to\infty} \frac{1}{n} \log \mu(B(\pi\omega, R_n^+(\omega))) \geq -h_\pi(\sigma, m, \omega). \tag{I.9.131}$$

Birkhoff applied to $\log \rho$ gives

$$\lim_{n\to\infty} \frac{1}{n} \log R_n^+(\omega) = \log c - \lambda_\Phi(m, \omega) < 0. \tag{I.9.132}$$

The consecutive radius ratio is

$$\frac{R_{n+1}^+(\omega)}{R_n^+(\omega)} = c\rho(\sigma^n\omega) \in [c\rho_*, c\rho^*]. \tag{I.9.133}$$

Write $h = h_\pi(\sigma, m, \omega)$ and $D_+ = \lambda_\Phi(m, \omega) - \log c > 0$. For every $\varepsilon > 0$, eventually

$$\frac{1}{n} \log \mu(B(\pi\omega, R_n^+)) \geq -h - \epsilon, \qquad -\frac{1}{n} \log R_n^+ \geq D_+ - \epsilon.$$

Division by $\log R_n^+ < 0$ therefore bounds the quotient above by $(h+\varepsilon)/(D_+ - \varepsilon)$. Lemma 6.3 and $\varepsilon \downarrow 0$ give

$$\overline{d}_\mu(\pi\omega) \leq \frac{h_\pi(\sigma, m, \omega)}{\lambda_\Phi(m, \omega) - \log c}. \tag{I.9.134}$$

In particular, since a local dimension ratio is nonnegative, (I.9.134) shows that the local projection entropy is nonnegative on this full-measure set.

For the small radii, substitute $r = R_{(k-1)}^-(\sigma\omega)$ in the left inclusion of (I.9.110). Now

$$B_\pi(\omega, R_k^-(\omega)) \cap \mathscr{P}(\omega) \subset \sigma^{-1} B_\pi(\sigma\omega, R_{k-1}^-(\sigma\omega)) \cap \mathscr{P}(\omega). \tag{I.9.135}$$

The same algebra reverses (I.9.129):

$$\begin{aligned} &\log m(B_\pi(\omega, R_k^-(\omega))) - \log m(B_\pi(\sigma\omega, R_{k-1}^-(\sigma\omega))) \\ &\quad \leq F_k^-(\omega). \end{aligned} \tag{I.9.136}$$

After iteration,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mu(B(\pi\omega, R_n^-(\omega))) \leq -h_\pi(\sigma, m, \omega), \tag{I.9.137}$$

where Lemma 5.2 again removes the terminal ball. This time

$$\lim_{n\to\infty} \frac{1}{n} \log R_n^-(\omega) = -\log c - \lambda_\Phi(m, \omega), \tag{I.9.138}$$

and

$$\frac{R_{n+1}^-(\omega)}{R_n^-(\omega)} = c^{-1}\rho(\sigma^n\omega) \in [c^{-1}\rho_*, c^{-1}\rho^*]. \tag{I.9.139}$$

Put $D_- = \lambda_\Phi(m, \omega) + \log c$. From (I.9.137)–(I.9.138), eventually

$$\frac{1}{n} \log \mu(B(\pi\omega, R_n^-)) \leq -h + \epsilon, \qquad -\frac{1}{n} \log R_n^- \leq D_- + \epsilon.$$

If $h > 0$, take $0 < \varepsilon < h$; division by $\log R_n^- < 0$ bounds the quotient below by $(h-\varepsilon)/(D_- + \varepsilon)$. If $h = 0$, the required lower bound is already true because local dimension ratios for a probability are nonnegative. Lemma 6.3 and $\varepsilon \downarrow 0$ therefore yield

$$\underline{d}_\mu(\pi\omega) \geq \frac{h_\pi(\sigma, m, \omega)}{\lambda_\Phi(m, \omega) + \log c}. \tag{I.9.140}$$

Choose a countable sequence $c_l \downarrow 1$ satisfying (I.9.120) and intersect the corresponding full-measure sets. Letting $l \to \infty$ in (I.9.134) and (I.9.140) proves (I.9.118)–(I.9.119). □

Theorem — 6.5 (T-I.9.08: Feng–Hu conformal formula) [F]

Provenance: FH-01, Theorem 2.8. The theorem is not imported.
Under the hypotheses of Theorem 6.4,

$$d_\mu(\pi\omega) = \frac{h_\pi(\sigma, m, \omega)}{\lambda_\Phi(m, \omega)} \tag{I.9.141}$$

for m-almost every ω. If m is ergodic, μ is exact dimensional and

$$\dim_{\mathrm{H}} \mu = \dim_{\mathrm{P}} \mu = \frac{h_\pi(\sigma, m)}{\chi_\Phi(m)}. \tag{I.9.142}$$

Proof

The lower local dimension never exceeds the upper one. Equations (I.9.118)–(I.9.119) force equality with the displayed ratio; this proves (I.9.141). Under ergodicity, conditional expectation on $\mathscr{I}$ makes the numerator and denominator constant:

$$h_\pi(\sigma, m, \omega) = h_\pi(\sigma, m), \qquad \lambda_\Phi(m, \omega) = \chi_\Phi(m). \tag{I.9.143}$$

Let $s = h_\pi(\sigma, m)/\chi_\Phi(m)$. The set $E = \{z \in \mathrm{supp}(\mu) : \underline{d}_\mu(z) = \overline{d}_\mu(z) = s\}$ is Borel by the local-dimension measurability established in I.4. Equation (I.9.141) gives $m(\pi^{-1}E) = 1$, and hence $\mu(E) = 1$. Thus μ is exact dimensional, without needing to assert that an arbitrary image of a Borel set is Borel. The accepted local-to-global theorem D/R-I.4 gives (I.9.142). No separation condition was used. □

I.9.7 Open-set fibers and geometric Gibbs measures

Theorem — 7.1 (T-I.9.09: OSC prefix multiplicity) [F]

Let $\Phi = (S_i)_{(i \in \Lambda)}$ be a finite self-similar IFS with contraction ratios r_i and suppose OSC holds. There is $C < \infty$ such that

$$M_n \leq C(n+1) \quad (n \geq 1). \tag{I.9.144}$$

Consequently, for every invariant symbolic probability m,

$$h_\pi(\sigma, m) = h_m(\sigma). \tag{I.9.145}$$

For a strict Bernoulli vector p, the self-similar measure is exact dimensional and

$$\dim_{\mathrm{H}} \mu_{\Phi,p} = \dim_{\mathrm{P}} \mu_{\Phi,p} = \frac{-\sum_i p_i \log p_i}{-\sum_i p_i \log r_i}. \tag{I.9.146}$$

Proof

Let O be an OSC open set. Choose $y \in O$ and $a > 0$ such that

$$B(y, a) \subset O. \tag{I.9.147}$$

Put

$$D = \max_{x \in K} |x - y|, \quad r_{\min} = \min_i r_i, \quad r_{\max} = \max_i r_i. \tag{I.9.148}$$

If two words are incomparable, their images of O are disjoint. Indeed, remove their maximal common prefix; their next letters differ, so the claim follows from $S_i(O) \cap S_j(O) = \emptyset$ and injectivity of the common-prefix map.

Fix $z \in K$, $n \geq 1$, and an integer k. Let

$$\mathscr{W}_{n,k}(z) = \{u \in \Lambda^n : z \in K_u,\ 2^{-(k+1)} < r_u \leq 2^{-k}\}. \tag{I.9.149}$$

For each $u \in W_{(n,k)}(z)$, choose $x_u \in K$ with $S_u(x_u) = z$. The balls

$$S_u(B(y, a)) = B(S_u(y), a r_u) \tag{I.9.150}$$

are pairwise disjoint, their radii exceed $a2^{-(k+1)}$, and

$$|S_u(y) - z| = r_u |y - x_u| \leq D2^{-k}. \tag{I.9.151}$$

They all lie in $B(z, (D + a)2^{-k})$. Comparing d-dimensional Lebesgue volumes gives

$$\#\mathscr{W}_{n,k}(z) \leq \left(\frac{2(D + a)}{a}\right)^d =: C_1. \tag{I.9.152}$$

Every length-n ratio lies in

$$r_{\min}^n \leq r_u \leq r_{\max}^n. \tag{I.9.153}$$

The number of dyadic intervals in (I.9.149) meeting this range is at most

$$2 + n \log_2(r_{\max}/r_{\min}) \leq C_2(n + 1). \tag{I.9.154}$$

Sum (I.9.152) over those classes. This proves (I.9.144), uniformly in z. Equations (I.9.80) and (I.9.83) give (I.9.145).

For Bernoulli $m = \beta_p$, A-DS or L-I.8.01 gives

$$h_m(\sigma) = -\sum_i p_i \log p_i, \qquad \chi_\Phi(m) = -\sum_i p_i \log r_i. \tag{I.9.155}$$

Now use (I.9.145) in the fully proved formula (I.9.142). This yields (I.9.146) and discharges the former G-I.8.09. □

Corollary — 7.2 (C-I.9.10: geometric Gibbs dimension) [F]

Let Φ be a finite $C^{1+\alpha}$ conformal IFS and let m_t be the geometric Gibbs state from Theorem 2.2. Without any separation condition,

$$\dim_{\mathrm{H}}(\pi_* m_t) = \dim_{\mathrm{P}}(\pi_* m_t) = \frac{h_\pi(\sigma, m_t)}{\chi_\Phi(m_t)}. \tag{I.9.156}$$

If coding is injective, or more generally satisfies (I.9.80), then

$$\dim_{\mathrm{H}}(\pi_* m_t) = \dim_{\mathrm{P}}(\pi_* m_t) = \frac{h_{m_t}(\sigma)}{\chi_\Phi(m_t)} = t + \frac{P_\Phi(t)}{\chi_\Phi(m_t)}. \tag{I.9.157}$$

At $t = s_\Phi$, under the same multiplicity hypothesis,

$$\dim_{\mathrm{H}}(\pi_* m_{s_\Phi}) = \dim_{\mathrm{P}}(\pi_* m_{s_\Phi}) = s_\Phi. \tag{I.9.158}$$

Proof

The A-DS Gibbs theorem makes m_t mixing and hence ergodic. Equation (I.9.156) is Theorem 6.5. Under (I.9.80), Theorem 4.6 gives $h_\pi = h$. Equation (I.9.31) then gives

$$\frac{h_{m_t}(\sigma)}{\chi_\Phi(m_t)} = t + \frac{P_\Phi(t)}{\chi_\Phi(m_t)}, \tag{I.9.159}$$

which proves (I.9.157). Injective coding has $M_n = 1$, so it is a special case. Finally $P_\Phi(s_\Phi) = 0$, and (I.9.158) follows. □

I.9.8 Formal conclusion

The projection-entropy loss is exactly the exponential conditional ambiguity of symbolic prefixes over a geometric point, through (I.9.82). The moving conformal-ball argument converts the retained entropy rate into Euclidean mass scaling, giving (I.9.141) without any separation condition. OSC is used only afterward, to prove the independent prefix-multiplicity estimate (I.9.144).

Every theorem in the chapter is F. FH-01 is the point-of-statement provenance for Definitions 4.1 and Theorems 4.6, 6.4, and 6.5; its results are not proof premises.

I.9.9 Intuition and strategy

The following discussion explains the geometric ideas behind the proofs.

I.9.9.1 Why ordinary entropy is sometimes the wrong numerator

A symbolic measure records information in labels. The projected Euclidean measure records only information that survives the coding map. If two labels produce the same geometric branch, deciding which label occurred costs symbolic entropy but creates no geometric refinement.

Projection entropy subtracts precisely this invisible information. The two conditional entropies in its definition ask:

- how uncertain is the first symbol when the projected tail is known?
- how uncertain is the first symbol when the projected full sequence is known?

Their difference is the first-symbol information that actually moves the geometric point to a smaller scale.

I.9.9.2 Conditional prefix entropy as coding ambiguity

The identity

$$h_m(\sigma) - h_\pi(\sigma, m) = \lim_{n\to\infty} \frac{1}{n} H_m(\mathscr{P}_0^{n-1} \mid \mathscr{A}_\pi)$$

has a concrete interpretation. Condition on the final Euclidean point and ask how many length-n symbolic histories remain plausible. Their conditional entropy per digit is exactly the information destroyed by overlap.

This also explains the multiplicity criterion. If only $e^{o(n)}$ prefixes can reach any point, their entropy is $o(n)$, so no positive entropy rate is lost. In particular, the criterion permits fibers with infinitely many codes: it controls the uniform growth of possible prefixes, not only the cardinality of an individual fiber. No equivalence with pointwise finite fiber cardinality is asserted here.

I.9.9.3 What conformality contributes

At a point of a conformal branch, the derivative has one scalar contraction factor. Products of these factors give a single asymptotic Lyapunov scale. This is why the dimension formula has one denominator.

In a genuinely affine system, different directions contract at different rates. Then one scalar ball recursion is insufficient; conditional entropies must be assigned to a flag of subspaces. That is the later Feng/Ledrappier– Young pathway in Volume V.

I.9.9.4 Bounded distortion and exact dimensionality play different roles

Bounded distortion compares an entire conformal cylinder with one derivative scale. It is crucial for pressure, Gibbs cylinders, and geometric thermodynamic formalism.

The no-separation exact-dimensionality theorem needs less. It uses only a uniform one-step C^1 comparison between a small ball and its image. The argument follows a typical orbit and never compares all points in one long cylinder. Thus Theorem 6.5 is valid at C^1, even though the pressure package is stated at $C^{1+\alpha}$.

I.9.9.5 The two moving radii

There are two recursively chosen radii:

$$R_n^+ = r_0 c^n \rho_n, \qquad R_n^- = r_0 c^{-n} \rho_n.$$

The larger radius contains the image of the previous tail ball. This gives a lower bound for its logarithmic mass, which becomes an upper bound for local dimension after division by the negative number $\log R_n^+$.

The smaller radius is contained in the image of the previous tail ball. It gives an upper bound for logarithmic mass and hence a lower bound for local dimension. Letting c decrease to one makes the bounds meet.

This sign reversal is not cosmetic. It is the main algebraic hazard in the proof.

I.9.9.6 Why differentiation of arbitrary measures appears

The ratio

$$\frac{m([i] \cap \pi^{-1} B(\pi\omega, r))}{m(\pi^{-1} B(\pi\omega, r))}$$

is a conditional probability at geometric resolution r. As r tends to zero, differentiation of the pushed-forward measure turns it into the exact conditional probability of $[i]$ given the geometric point.

Lebesgue differentiation alone is not enough: the projected measure may be singular. The Besicovitch selection lemma supplies differentiation relative to any finite Borel measure on Euclidean space.

The logarithmic envelope is equally important. Pointwise convergence of those ratios does not justify averaging a moving triangular array. The weak density estimate gives one integrable function dominating all small-radius negative logarithms.

I.9.9.7 Why OSC restores classical entropy

Under OSC, the images of a feasible open set under incomparable words contain disjoint miniature copies of one fixed ball. These balls need not lie in the attractor cylinders themselves; the proof controls their distance from the cylinders. At a fixed contraction scale, only boundedly many such balls can cluster near a point. Length-n words can have different contraction sizes, but they occupy only $O(n)$ dyadic scale classes. Hence a point belongs to only $O(n)$ length-n cylinders.

The logarithm of $O(n)$ is negligible compared with n, so conditional prefix entropy has zero rate. This proves $h_\pi = h$ without requiring unique coding.

I.9.9.8 The pressure zero

The geometric potential is the logarithm of infinitesimal contraction. The Gibbs state at parameter t balances entropy against average contraction:

$$P_\Phi(t) = h_{m_t}(\sigma) - t\chi_\Phi(m_t).$$

At the pressure zero, the balance becomes $h = s\chi$. When coding loses no entropy rate, the projection-entropy dimension formula turns this identity into $\dim(\pi_* m_s) = s$.

With overlap, the same Gibbs state is still exact dimensional, but its dimension is h_π/χ; pressure alone does not calculate the entropy that survives projection.

I.9.9.9 Research lineage

The full no-separation mechanism is due to De-Jun Feng and Huyi Hu. The chapter proves the conformal specialization of their projection-entropy theory. The later affine theory keeps the same guiding question—how much symbolic information survives each geometric factor—but replaces the one conformal scale by several Lyapunov directions and conditional partitions.

I.9.10 Exercises

All exercises use only accepted incoming nodes and the F results of this chapter preceding the exercise's topic.

I.9.10.1 Exercise I.9.1 — Derivative cocycle

Prove (I.9.15) directly from the chain rule and conformality. Deduce (I.9.25).

I.9.10.2 Exercise I.9.2 — Pressure stability

For bounded continuous potentials φ, ψ, prove (I.9.36). Deduce that $t \mapsto \mathscr{P}(t\varphi)$ is convex.

I.9.10.3 Exercise I.9.3 — Similarity pressure

For a similarity IFS with ratios r_i, calculate $P_\Phi(t)$ explicitly and show that its zero satisfies $\sum_i r_i^s = 1$.

I.9.10.4 Exercise I.9.4 — Besicovitch angular separation

Fill in the cosine-rule calculation in (I.9.42) and derive a dimension-only cardinality bound by Euclidean volume packing.

I.9.10.5 Exercise I.9.5 — Conditional density

Let $\theta : \Omega \to \mathbb{R}^d$ and $A \in F$. Prove directly from the Radon–Nikodym identity that $d\theta_*(1_A m)/d(\theta_* m) \circ \theta$ is a version of $E(1_A|\theta^{-1}B(\mathbb{R}^d))$.

I.9.10.6 Exercise I.9.6 — Logarithmic envelope integral

For $0 < a \le 1$ and $B \ge 1$, verify

$$\int_0^\infty \min\{a, Be^{-t}\}\, dt = a\left(1 + \log \frac{B}{a}\right).$$

Use it to rederive (I.9.54).

I.9.10.7 Exercise I.9.7 — Coding sigma-algebra

Prove both inclusions in (I.9.72), taking care that Borel images under S_i are Borel.

I.9.10.8 Exercise I.9.8 — The two-symbol block identity

Write out (I.9.73) for $n = 2$ using the conditional chain rule. Identify the two shifted conditional-information differences before integrating, and explain why each integral equals h_π.

I.9.10.9 Exercise I.9.9 — Entropy under finite support

If a probability vector has at most M nonzero entries, prove that its entropy is at most $\log M$, with equality exactly when it has precisely M nonzero entries, all equal to $1/M$. Apply this to (I.9.93).

I.9.10.10 Exercise I.9.10 — A tail estimate for L^1

If $G \in L^1(m)$ and T preserves m, prove $G(T^n x)/n \to 0$ almost everywhere. Explain exactly where this enters the moving-array proof.

I.9.10.11 Exercise I.9.11 — Intermediate radii

Give the complete squeeze proving Lemma 6.3, including the reversal of inequalities caused by negative logarithms.

I.9.10.12 Exercise I.9.12 — Duplicate-label projection entropy

For

$$S_0(x) = S_1(x) = x/3, \qquad S_2(x) = x/3 + 2/3,$$

with nonnegative Bernoulli weights (p_0, p_1, p_2) summing to one, put $q = p_0 + p_1$. Compute $H(P|\sigma^{-1}A_\pi)$, $H(P|A_\pi)$, and h_π, including $q = 0$.

I.9.10.13 Exercise I.9.13 — OSC scale classes

Prove the exact integer bound behind (I.9.154) and show that equal contraction ratios improve (I.9.144) to $M_n \leq C$.

I.9.10.14 Exercise I.9.14 — Gibbs dimension at pressure zero

Assume uniformly subexponential fiber-prefix multiplicity. Starting only from (I.9.31), (I.9.83), and (I.9.142), prove (I.9.157)–(I.9.158).

Chapter I.10

L^q Dimensions and Multifractal Spectra

A single dimension cannot describe all the different rates at which mass accumulates. Moment sums emphasize different parts of a measure as the parameter varies, while level sets group points with the same local dimension. This chapter relates these two viewpoints through multifractal bounds and, under the stated additional assumptions, equalities. Particular care is needed for negative moments: a fixed grid and a family of centered balls are not interchangeable.

The definitions below fix this chapter's conventions. Geometric intuition and proof strategy are discussed separately after the main development; exercises and their solutions provide a second route through the material.

Definitions and notation

All logarithms are natural. The half-open dyadic cubes of generation n in $\mathbb{R}^d$ are

$$\mathscr{D}_n = \{2^{-n}(k + [0,1)^d) : k \in \mathbb{Z}^d\}.$$

Only finitely many such cubes meet the compact support of a measure. Terms already accepted in I.1–I.9 are reused with their established conventions.

Definition — I.10.01: Moment sums and L^q spectra

Let μ be a compactly supported Borel probability measure. For $q \in \mathbb{R}$ put

$$\mathscr{D}_n^+(\mu) = \{Q \in \mathscr{D}_n : \mu(Q) > 0\}, \qquad M_n(\mu, q) = \sum_{Q \in \mathscr{D}_n^+(\mu)} \mu(Q)^q.$$

Thus zero-mass cubes are omitted, including when $q = 0$, and $M_n(\mu, 0) = \#\mathscr{D}_n^+(\mu)$. The finite-scale exponent is

$$\tau_n^{\mathrm{dyad}}(\mu, q) = \frac{\log M_n(\mu, q)}{-n \log 2}.$$

The **lower** and **upper dyadic** L^q **spectra** are

$$\underline{\tau}_\mu^{\mathrm{dyad}}(q) = \liminf_{n\to\infty} \tau_n^{\mathrm{dyad}}(\mu, q), \qquad \overline{\tau}_\mu^{\mathrm{dyad}}(q) = \limsup_{n\to\infty} \tau_n^{\mathrm{dyad}}(\mu, q).$$

When they agree in $\mathbb{R}$, their common value is $\tau_\mu^{\mathrm{dyad}}(q)$. Fixed dyadic grids can be unstable for $q < 0$, so this notation is not silently identified with the centered definition below. Let

$$\mathscr{M}_\mu(q, r) = \sup \left\{ \sum_j \mu(B(x_j, r))^q \;:\; x_j \in \operatorname{supp} \mu,\; B(x_j, r) \text{ pairwise disjoint} \right\},$$

and define the **lower centered-packing** L^q **spectrum** by

$$\underline{\tau}_\mu^{\mathrm{pack}}(q) = \liminf_{r\downarrow 0} \frac{\log \mathscr{M}_\mu(q, r)}{\log r}.$$

The upper centered-packing spectrum uses lim sup; when the two agree, write $\tau_\mu^{\mathrm{pack}}(q)$. Every ball centered on the support has positive mass, so negative powers are defined. Dyadic and packing spectra are not identified for negative q without a separately proved comparison.

Definition — I.10.02: Renyi entropies and dimensions

For $q \in \mathbb{R} \setminus \{1\}$ the order-q dyadic **Renyi entropy** is

$$H_q(\mu, \mathscr{D}_n) = \frac{1}{1-q} \log M_n(\mu, q).$$

At $q = 1$, it is the Shannon entropy

$$H_1(\mu, \mathscr{D}_n) = -\sum_{Q \in \mathscr{D}_n^+(\mu)} \mu(Q) \log \mu(Q),$$

with $0 \log 0 = 0$. At $q = 0$, it is $\log \#\mathscr{D}_n^+(\mu)$. The lower and upper order-q **Renyi dimensions** are

$$\underline{D}_q^{\mathrm{dyad}}(\mu) = \liminf_{n\to\infty} \frac{H_q(\mu, \mathscr{D}_n)}{n \log 2}, \qquad \overline{D}_q^{\mathrm{dyad}}(\mu) = \limsup_{n\to\infty} \frac{H_q(\mu, \mathscr{D}_n)}{n \log 2}.$$

When equal, the common value is $D_q^{\mathrm{dyad}}(\mu)$. For $q \neq 1$ and an existing dyadic spectrum,

$$D_q^{\mathrm{dyad}}(\mu) = \frac{\tau_\mu^{\mathrm{dyad}}(q)}{q-1}.$$

The $q = 1$ value is the dyadic entropy dimension and $q = 0$ is the dyadic box dimension of the measure's positive-mass support at the registered grid. If $\tau_\mu^{\mathrm{pack}}(q)$ exists, define the **centered-packing generalized dimension**

$$D_q^{\mathrm{pack}}(\mu) = \frac{\tau_\mu^{\mathrm{pack}}(q)}{q-1} \quad (q \neq 1).$$

Definition — I.10.03: Level sets and Hausdorff spectrum

Using the accepted local dimension $d_\mu(x)$ from I.4, define

$$E_\mu(\alpha) = \{x \in \operatorname{supp}\mu : d_\mu(x) = \alpha\}, \qquad f_\mu(\alpha) = \begin{cases} \dim_{\mathrm{H}} E_\mu(\alpha), & E_\mu(\alpha) \neq \varnothing, \\ -\infty, & E_\mu(\alpha) = \varnothing. \end{cases}$$

Thus an empty level receives spectrum value $-\infty$. This does not change the standing geometric convention $\dim_{\mathrm{H}} \varnothing = 0$ from Chapter I.1. A theorem may instead say explicitly that the level set is empty.

Definition — I.10.04: Legendre formalism and slope terminology

For a finite concave function $\tau : J \to \mathbb{R}$ on an interval J, its **concave Legendre transform** is

$$\tau^*(\alpha) = \inf_{q \in J}\{q\alpha - \tau(q)\}.$$

A measure satisfies the **multifractal formalism at α relative to τ and J** if $f_\mu(\alpha) = \tau^*(\alpha)$. A number α is the **slope exposed by q** if

$$\tau(q') \leq \tau(q) + \alpha(q' - q) \quad (q' \in J).$$

If q is an interior point of J and τ is differentiable at q, this is equivalent to $\alpha = \tau'(q)$. A nondifferentiability point of τ is a **phase transition**. For a differentiable concave function on $\mathbb{R}$, its **extremal slopes** are

$$\alpha_{\min} = \lim_{q \to +\infty} \tau'(q), \qquad \alpha_{\max} = \lim_{q \to -\infty} \tau'(q),$$

and the open interval between them is the **interior slope range**.

Definition — I.10.05: Normalized Gibbs potential and pressure root

Let $\Sigma = \Lambda^{\mathbb{N}}$, let $\psi : \Sigma \to \mathbb{R}$ be Hölder, and let m_ψ be its equilibrium state. Its **normalized potential** is

$$\psi_0 = \psi - P(\psi), \qquad P(\psi_0) = 0.$$

Let Φ be a finite $C^{1+\alpha}$ conformal IFS and let $\phi < 0$ be the geometric potential from I.9. For $q \in \mathbb{R}$, the **geometric pressure root** $T_\psi(q)$ is the unique real number satisfying

$$P(q\psi_0 - T_\psi(q)\phi) = 0.$$

The sign is fixed by the contraction convention $\phi < 0$; in the self-similar Bernoulli case it becomes

$$\sum_{i\in\Lambda} p_i^q r_i^{-T_\psi(q)} = 1.$$

Definition — I.10.06: Quasi-Bernoulli measures

A probability m on Σ is **upper quasi-Bernoulli** if some $C_+ \geq 1$ satisfies

$$m([uv]) \leq C_+ m([u])m([v]) \quad (u, v \in \Lambda^*).$$

It is **lower quasi-Bernoulli** if some $C_{\rightarrow}0$ satisfies

$$m([uv]) \geq C_- m([u])m([v]).$$

It is **two-sided quasi-Bernoulli** if both conditions hold. A finite-shift Hölder Gibbs state is two-sided quasi-Bernoulli by the accepted Gibbs cylinder estimate; the deduction is proved locally.

Definition — I.10.07: Subexponential errors and q-Gibbs balls

A positive sequence (c_n) is **subexponential** if

$$\lim_{n\to\infty} \frac{\log c_n}{n} = 0.$$

A positive function $h : (0, r_0) \to (0, \infty)$ is **subpower at zero** if

$$\lim_{r\downarrow 0} \frac{\log h(r)}{\log(1/r)} = 0.$$

Given $q > 0$, an existing exponent $\tau(q)$, and $c > 1$, an auxiliary probability ν_q satisfies the **upper q-Gibbs ball inequality** if

$$\nu_q(B(x, r/c)) \leq h(r) r^{-\tau(q)} \mu(B(x, r))^q$$

for all sufficiently small r and every x, with h subpower at zero. It satisfies the **lower q-Gibbs ball inequality** if a positive subpower function h satisfies

$$\nu_q(B(x, cr)) \geq h(r) r^{-\tau(q)} \mu(B(x, r))^q$$

for all sufficiently small r and every x. The radius enlargement is part of the definition; the function h need not be the one in an upper inequality.

Definition — I.10.08: Dyadic neighbors and conformal stopping words

For $Q, Q' \in D_n$, write $Q \sim Q'$ if their closures intersect. For a finite conformal IFS with attractor K, define

$$\mathscr{W}_n = \{u \in \Lambda^*,\ |u| \geq 1 \,:\, \operatorname{diam} K_u \leq 2^{-n} \text{ and } \operatorname{diam} K_{u^-} > 2^{-n}\},$$

when $2^{-n} < \operatorname{diam} K$. If $2^{-n} \geq \operatorname{diam} K$, set $W_n = \{\varnothing\}$ with $S_\varnothing = id$ and $K_\varnothing = K$; the empty word has no parent. Here $\operatorname{diam} K > 0$, and the singleton case is handled separately in the results. For nonempty words, u^- is u with its last letter removed. This is the **generation-n conformal stopping family**. Its cylinders form a symbolic antichain cover; no separation property is included in the definition.

Terminal terminology (no outgoing proof dependency)

Question/Boundary 9.1 defines the asymptotically weak separation condition using the D-I.10.08 stopping family and a subexponential bound on the number of distinct maps meeting each dyadic cube. It also defines the ell-adic interval level sets used by the cited Testud examples. Both definitions are printed in full at that discussion, before use; neither is silently identified with an earlier stronger separation or ball-level convention.

All logarithms are natural and all Euclidean balls are closed. Let $\mathscr{D}_n$ be the half-open dyadic cubes of side 2^{-n} in $\mathbb{R}^d$. A sum over positive-mass cubes always omits the zero-mass cubes. The distinction between a fixed dyadic grid and centered packings is kept visible throughout, especially for negative q.

I.10.1 Moment sums, Rényi dimensions, and the universal upper bound

Definition — 1.1 (dyadic moment sums and spectra)

For a compactly supported Borel probability measure μ, define

$$\mathscr{D}_n^+(\mu) = \{Q \in \mathscr{D}_n \,:\, \mu(Q) > 0\}, \qquad M_n(\mu, q) = \sum_{Q \in \mathscr{D}_n^+(\mu)} \mu(Q)^q \tag{I.10.1}$$

for $q \in \mathbb{R}$. Thus $M_n(\mu, 0) = \#\mathscr{D}_n^+(\mu)$. Put

$$\tau_n^{\mathrm{dyad}}(\mu, q) = \frac{\log M_n(\mu, q)}{-n \log 2}, \tag{I.10.2}$$

and

$$\underline{\tau}_\mu^{\mathrm{dyad}}(q) = \liminf_{n\to\infty} \tau_n^{\mathrm{dyad}}(\mu, q), \qquad \overline{\tau}_\mu^{\mathrm{dyad}}(q) = \limsup_{n\to\infty} \tau_n^{\mathrm{dyad}}(\mu, q). \tag{I.10.3}$$

When the two finite values agree, their common value is $\tau_\mu^{\mathrm{dyad}}(q)$.

Definition — 1.2 (centered-packing spectrum)

For $r > 0$, let

$$\mathscr{M}_\mu(q,r) = \sup\left\{ \sum_{j=1}^{N} \mu(B(x_j,r))^q \ : \ \begin{array}{l} N < \infty,\ x_j \in \operatorname{supp}\mu, \\ B(x_1,r),\dots,B(x_N,r) \text{ are pairwise disjoint} \end{array} \right\}. \tag{I.10.4}$$

Every ball in (I.10.4) has positive mass, and for fixed r these masses have a positive uniform lower bound. Indeed, cover the compact support by finitely many balls $B(z_j,r/2)$ centered on it. For any support point x one of these centers satisfies $|x-z_j| \le r/2$, whence $B(z_j,r/4) \subset B(x,r)$. The minimum of the finitely many positive masses $\mu(B(z_j,r/4))$ is the required bound. The number of disjoint radius-r balls with centers in this bounded support is also uniformly finite by Euclidean volume packing. Thus $0 < \mathscr{M}_\mu(q,r) < \infty$ for every real q, including negative orders. Define

$$\underline{\tau}_\mu^{\text{pack}}(q) = \liminf_{r\downarrow 0} \frac{\log \mathscr{M}_\mu(q,r)}{\log r}, \qquad \overline{\tau}_\mu^{\text{pack}}(q) = \limsup_{r\downarrow 0} \frac{\log \mathscr{M}_\mu(q,r)}{\log r}. \tag{I.10.5}$$

When they agree finitely, write $\tau_\mu^{\text{pack}}(q)$. No negative-$q$ identification between (I.10.3) and (I.10.5) is part of the notation.

Definition — 1.3 (Rényi entropy and generalized dimensions)

For $q \neq 1$, the generation-n Rényi entropy is

$$H_q(\mu,\mathscr{D}_n) = \frac{1}{1-q}\log M_n(\mu,q). \tag{I.10.6}$$

At $q=1$ and $q=0$, respectively, put

$$H_1(\mu,\mathscr{D}_n) = -\sum_{Q\in\mathscr{D}_n^+(\mu)} \mu(Q)\log\mu(Q), \qquad H_0(\mu,\mathscr{D}_n) = \log \#\mathscr{D}_n^+(\mu). \tag{I.10.7}$$

The lower and upper dyadic Rényi dimensions are

$$\underline{D}_q^{\text{dyad}}(\mu) = \liminf_{n\to\infty} \frac{H_q(\mu,\mathscr{D}_n)}{n\log 2}, \qquad \overline{D}_q^{\text{dyad}}(\mu) = \limsup_{n\to\infty} \frac{H_q(\mu,\mathscr{D}_n)}{n\log 2}. \tag{I.10.8}$$

Their common value, when it exists, is $D_q^{\text{dyad}}(\mu)$. If $q \neq 1$ and the centered-packing spectrum exists, define

$$D_q^{\text{pack}}(\mu) = \frac{\tau_\mu^{\text{pack}}(q)}{q-1}. \tag{I.10.9}$$

Definition — 1.4 (level sets and Legendre formalism)

Using the local dimension from I.4, define

$$E_\mu(\alpha)=\{x\in \operatorname{supp}\mu\,:\,d_\mu(x)=\alpha\},\qquad f_\mu(\alpha)=\begin{cases}\dim_{\mathrm H}E_\mu(\alpha), & E_\mu(\alpha)\neq\varnothing,\\ -\infty, & E_\mu(\alpha)=\varnothing.\end{cases}\tag{I.10.10}$$

Only an empty spectrum level is assigned value $-\infty$; the geometric convention $\dim_{\mathrm H}\varnothing=0$ remains unchanged. If $\tau : J\to\mathbb{R}$ is finite and concave on an interval J, its concave Legendre transform is

$$\tau^*(\alpha)=\inf_{q\in J}\{q\alpha-\tau(q)\}.\tag{I.10.11}$$

The multifractal formalism holds at α, relative to τ and J, when

$$f_\mu(\alpha)=\tau^*(\alpha).\tag{I.10.12}$$

A number α is a slope exposed by $q\in J$ if

$$\tau(s)\le\tau(q)+\alpha(s-q)\quad(s\in J).\tag{I.10.13}$$

At an interior differentiability point of J, this says $\alpha=\tau'(q)$. A point of nondifferentiability is called a phase transition. For a differentiable concave function on $\mathbb{R}$, its extremal slopes and interior slope range are

$$\alpha_{\min}=\lim_{q\to+\infty}\tau'(q),\qquad \alpha_{\max}=\lim_{q\to-\infty}\tau'(q),\qquad (\alpha_{\min},\alpha_{\max}).\tag{I.10.13a}$$

Lemma — 1.5 (L-I.10.01: finite moment and entropy calculus) [F]

Let $a=(a_1,\dots,a_N)$ be a probability vector after its zero entries have been deleted, and put $Z_a(q)=\sum_j a_j^q$. Then

$$q\longmapsto \log Z_a(q)\quad\text{is convex,}\qquad q\longmapsto\frac{\log Z_a(q)}{1-q}\quad\text{is nonincreasing.}\tag{I.10.14}$$

The second function is understood with its continuous extension at $q=1$.
Moreover,

$$\lim_{q\to1}\frac{\log Z_a(q)}{1-q}=-\sum_j a_j\log a_j.\tag{I.10.15}$$

Consequently $q\mapsto\tau_n^{\mathrm{dyad}}(\mu,q)$ is concave. If $\{a_j : j\in I\}$ is divided into groups $I_1,\dots,I_L$, each of cardinality at most A, then, for $b_\ell=\sum_{j\in I_\ell}a_j$,

$$\sum_j a_j^q \le \sum_\ell b_\ell^q \le A^{q-1} \sum_j a_j^q \quad (q \ge 1), \tag{I.10.16}$$

and

$$A^{q-1} \sum_j a_j^q \le \sum_\ell b_\ell^q \le \sum_j a_j^q \quad (0 < q \le 1). \tag{I.10.17}$$

Proof

For $q_0, q_1 \in \mathbb{R}$ and $0 < \lambda < 1$, Hölder's inequality gives

$$\begin{aligned} Z_a((1-\lambda)q_0 + \lambda q_1) &= \sum_j (a_j^{q_0})^{1-\lambda} (a_j^{q_1})^\lambda \\ &\le Z_a(q_0)^{1-\lambda} Z_a(q_1)^\lambda. \end{aligned} \tag{I.10.18}$$

This proves the first assertion in (I.10.14). Let J be a random index with $\mathbb{P}(J = j) = a_j$, let $Y = \log a_J$, and set

$$K(t) = \log \mathbb{E}(e^{tY}) = \log \sum_j a_j^{1+t}. \tag{I.10.19}$$

The function K is convex and $K(0) = 0$. Therefore its secant slope $K(t)/t$ is nondecreasing in t, including across $t = 0$. Since

$$\frac{\log Z_a(q)}{1-q} = -\frac{K(q-1)}{q-1}, \tag{I.10.20}$$

the second assertion in (I.10.14) follows. Differentiating the finite sum at zero gives $K'(0) = \sum_j a_j \log a_j$, and (I.10.15) follows from (I.10.20). Division of the convex function $\log Z_a(q)$ by $-n \log 2 < 0$ proves the concavity of τ_n^{dyad}.

For nonnegative $x_1, \dots, x_k$, $k \le A$, convexity gives

$$\sum_{j=1}^k x_j^q \le \left(\sum_{j=1}^k x_j \right)^q \le k^{q-1} \sum_{j=1}^k x_j^q \quad (q \ge 1), \tag{I.10.21}$$

while concavity gives both reversed comparisons, with the same factor, for $0 < q \le 1$. Sum (I.10.21) over the groups. This proves (I.10.16)–(I.10.17). □

Theorem — 1.6 (T-I.10.02: universal multifractal upper bound) [F]

For every compactly supported probability μ, every $\alpha \in \mathbb{R}$ with $E_\mu(\alpha) \ne \emptyset$, and every $q \in \mathbb{R}$ for which the lower centered-packing spectrum is finite,

$$\dim_{\mathrm{H}} E_\mu(\alpha) \le q\alpha - \underline{\tau}_\mu^{\mathrm{pack}}(q). \tag{I.10.22}$$

If $q\alpha - \underline{\tau}_\mu^{\mathrm{pack}}(q) < 0$, then the level set is empty. Consequently, for all real α, on every interval J on which the spectrum is finite,

$$f_\mu(\alpha) \le \inf_{q\in J}\{q\alpha - \underline{\tau}_\mu^{\mathrm{pack}}(q)\}. \tag{I.10.23}$$

Proof

At a support point x, one has $0 < \mu(B(x,r)) \le 1$ for small r, so every finite local dimension is nonnegative. Hence $E_\mu(\alpha) = \emptyset$ for $\alpha < 0$. The assertion about $f_\mu(\alpha)$ is then immediate, because its value is $-\infty$. There is no assertion that the geometric dimension of the empty set is bounded by a negative number. Assume $\alpha \ge 0$.

Fix $\varepsilon > 0$. For $k \ge 1$, let F_k be the set of $x \in E_\mu(\alpha)$ such that, whenever $0 < r < 1/k$,

$$r^{\alpha+\varepsilon} \le \mu(B(x,r)) \le r^{\alpha-\varepsilon}. \tag{I.10.24}$$

Then $E_\mu(\alpha) = \bigcup_k F_k$. Fix k, take $r < 1/(3k)$, and choose a maximal disjoint family

$$\{B(x_j,r) : 1 \le j \le N_r\}, \qquad x_j \in F_k. \tag{I.10.25}$$

Compactness makes the family finite, and maximality gives

$$F_k \subset \bigcup_{j=1}^{N_r} B(x_j, 3r). \tag{I.10.26}$$

If $q \ge 0$, the left inequality in (I.10.24) gives

$$N_r r^{q(\alpha+\varepsilon)} \le \sum_{j=1}^{N_r} \mu(B(x_j,r))^q \le \mathscr{M}_\mu(q,r). \tag{I.10.27}$$

If $q < 0$, the right inequality in (I.10.24), with the reversal caused by the negative power, gives

$$N_r r^{q(\alpha-\varepsilon)} \le \sum_{j=1}^{N_r} \mu(B(x_j,r))^q \le \mathscr{M}_\mu(q,r). \tag{I.10.28}$$

Write $\tau = \underline{\tau}_\mu^{\mathrm{pack}}(q)$. From the definition of a liminf, for every $\delta > 0$ and all sufficiently small r,

$$\frac{\log \mathscr{M}_\mu(q,r)}{\log r} \ge \tau - \delta. \tag{I.10.29}$$

Because $\log r < 0$, (I.10.29) is equivalent to

$$\mathscr{M}_\mu(q,r) \le r^{\tau-\delta}. \tag{I.10.30}$$

Equations (I.10.27)–(I.10.30) imply in both cases

$$N_r \le r^{-q\alpha-|q|\varepsilon+\tau-\delta}. \tag{I.10.31}$$

If $q\alpha - \tau < 0$, choose $\varepsilon, \delta > 0$ so small that $q\alpha - \tau + |q|\varepsilon + \delta < 0$. Equation (I.10.31) would force $N_r < 1$ for all sufficiently small r, impossible when F_k is nonempty. Thus every F_k, and hence the level set, is empty in this case. Otherwise, take any positive s satisfying

$$s > q\alpha - \tau + |q|\varepsilon + \delta, \tag{I.10.32}$$

then the s-cost of the cover (I.10.26) tends to zero:

$$\sum_{j=1}^{N_r} (6r)^s \le 6^s r^{s-q\alpha+|q|\varepsilon+\tau-\delta} \longrightarrow 0. \tag{I.10.33}$$

Thus

$$\dim_{\mathrm{H}} F_k \le q\alpha - \tau + |q|\varepsilon + \delta. \tag{I.10.34}$$

Countable stability, followed by $\varepsilon, \delta \downarrow 0$, proves (I.10.22). Taking the infimum in q proves (I.10.23). □

I.10.2 Strictly separated self-similar measures

Let $\Phi = (S_i)_{i\in\Lambda}$ be a finite similarity IFS on $\mathbb{R}^d$, with contraction ratios $0 < r_i < 1$, attractor K, and SSC. Let $p_i > 0$, $\sum_i p_i = 1$, and $\mu = \mu_{\Phi,p}$. For a word $u = u_1 \cdots u_n$, write

$$r_u = \prod_{j=1}^{n} r_{u_j}, \qquad p_u = \prod_{j=1}^{n} p_{u_j}. \tag{I.10.35}$$

For $0 < r < 1$, use the stopping antichain

$$\mathscr{W}(r) = \{u \,:\, r_u \le r < r_{u^-}\}. \tag{I.10.36}$$

Here the empty word has ratio 1. If the alphabet has just one symbol, $p_1 = 1$ and K is its fixed point. Every moment sum equals one, every spectrum and generalized dimension is zero, and the only nonempty local-dimension level is $\alpha = 0$. This verifies all the assertions in Sections 2–3 for that case. In their geometric proofs below we may therefore assume at least two symbols, so SSC gives a positive separation gap and $\operatorname{diam} K > 0$.

Lemma — 2.1 (stopping geometry and moment comparison) [F]

There are constants $c, C, A > 0$, depending only on Φ, with the following properties for every sufficiently small r:

(1) if $u, v \in \mathscr{W}(r)$, $u \ne v$, then

$$cr \le r_u \le r, \qquad \operatorname{dist}(K_u, K_v) \ge cr, \qquad \operatorname{diam} K_u \le Cr; \tag{I.10.37}$$

(2) every ball of radius Cr meets at most A stopping cylinders, and every stopping cylinder meets at most A cubes of side comparable to r;

(3) for every real q,

$$\mathscr{M}_\mu(q, ar) \asymp_q Z(r, q) := \sum_{u \in \mathscr{W}(r)} p_u^q \tag{I.10.38}$$

for one fixed $a > 0$, with analogous two-sided comparisons after replacing ar by any fixed constant multiple of r; and

(4) for $q \geq 0$, if $2^{-n-1} < r \leq 2^{-n}$, then

$$M_n(\mu, q) \asymp_q Z(r, q). \tag{I.10.39}$$

The constants implicit in (I.10.38)–(I.10.39) may depend on q, but not on r or n.

Proof

Set $r_* = \min_i r_i$. From (I.10.36),

$$r_* r < r_u \leq r. \tag{I.10.40}$$

Let

$$\Delta = \min_{i \neq j} \operatorname{dist}(S_i K, S_j K) > 0. \tag{I.10.41}$$

Distinct stopping words are incomparable. If w is the common prefix of u and v, their next letters differ and similarity gives

$$\operatorname{dist}(K_u, K_v) \geq r_w \Delta. \tag{I.10.42}$$

Since w is a proper prefix of u, $r_w > r$, so (I.10.42) is at least Δr. Also $\operatorname{diam} K_u = r_u \operatorname{diam} K$. This proves (I.10.37).

A Euclidean packing estimate now bounds the number of cr-separated sets K_u that can meet a ball of radius Cr. The same estimate bounds the number of cubes of side comparable to r meeting a fixed K_u. This proves item 2.

Choose $x_u \in K_u$. By decreasing $a > 0$, the balls $B(x_u, ar)$ are pairwise disjoint and meet no stopping cylinder other than K_u. Hence

$$\mu(B(x_u, ar)) \leq p_u. \tag{I.10.43}$$

For $q < 0$, (I.10.43) gives

$$\mathscr{M}_\mu(q, ar) \geq \sum_u p_u^q = Z(r, q). \tag{I.10.44}$$

For $q \geq 0$, enlarge the radius by a fixed factor so that the ball about x_u contains K_u. The resulting bounded-overlap family can be split into a bounded number of disjoint subfamilies, and

$$\mu(B(x_u, Cr)) \geq p_u. \tag{I.10.45}$$

It follows that $\mathscr{M}_\mu(q, Cr) \gtrsim_q Z(r, q)$.

Conversely, take a disjoint family of balls $B(x_j, \rho)$, centered on K. For $q < 0$, use the stopping family at scale $r = \rho/(2 \operatorname{diam} K + 2)$. The stopping cylinder $K_{u(j)}$ containing x_j then lies inside $B(x_j, \rho)$, and hence

$$\mu(B(x_j, \rho)) \geq p_{u(j)}. \tag{I.10.46a}$$

The separation and packing estimate show that a fixed u can be assigned to only a bounded number of disjoint balls. Raising (I.10.46a) to a negative power therefore bounds the ball sum above by a constant times $Z(r, q)$. For $q > 0$, use stopping scale comparable to ρ: each ball meets at most a bounded number of stopping cylinders, so its mass is at most the sum of their p_u, and each stopping cylinder is charged only a bounded number of times. Applying (I.10.16)–(I.10.17) gives the required upper comparison. The already constructed separated balls give the reverse comparison. To justify the fixed scale changes, fix $\lambda > 0$. Every word in $\mathscr{W}(\lambda r)$ extends, or is extended by, a word in $\mathscr{W}(r)$, and the difference of their lengths is bounded by an integer N_λ depending only on $\lambda, r_*, \max_i r_i$. Passing between the two antichains therefore replaces each summand by a sum over at most $|\Lambda|^{N_\lambda}$ descendants and multiplies every weight by one of finitely many factors p_v^q, $|v| \leq N_\lambda$. Consequently

$$Z(\lambda r, q) \asymp_{q,\lambda} Z(r, q).$$

This proves (I.10.38) at one fixed radius normalization and hence at every fixed multiple of that radius.

For a dyadic cube at scale $2^{-n} \asymp r$, item 2 says that its mass is a sum of pieces from at most A stopping cylinders and that each cylinder is split among at most A cubes. Equations (I.10.16)–(I.10.17), with the zero-order counting version, give (I.10.39). This argument is unavailable for $q < 0$: an arbitrarily small positive piece created by a grid boundary can dominate a negative moment. □

Theorem — 2.2 (T-I.10.03: self-similar pressure equation) [F]

For every $q \in \mathbb{R}$, the centered-packing spectrum exists and is the unique real number $\tau(q)$ satisfying

$$\sum_{i \in \Lambda} p_i^q r_i^{-\tau(q)} = 1. \tag{I.10.46}$$

For $q \geq 0$, the dyadic spectrum also exists and

$$\tau_\mu^{\text{dyad}}(q) = \tau_\mu^{\text{pack}}(q) = \tau(q). \tag{I.10.47}$$

No equality with a fixed dyadic grid is asserted for $q < 0$. The function $\tau : \mathbb{R} \to \mathbb{R}$ is concave and real analytic. With

$$\theta_i(q) = p_i^q r_i^{-\tau(q)}, \tag{I.10.48}$$

one has $\sum_i \theta_i(q) = 1$ and

$$\tau'(q) = \frac{\sum_i \theta_i(q) \log p_i}{\sum_i \theta_i(q) \log r_i}. \tag{I.10.49}$$

Proof

For fixed q, the left side of (I.10.46) is continuous and strictly increasing in τ, tends to 0 as $\tau \to -\infty$, and tends to ∞ as $\tau \to +\infty$. Thus there is a unique root.

For $u \in \mathscr{W}(r)$, define $\theta_u(q) = \prod_{j=1}^{|u|} \theta_{u_j}(q)$. The stopping cylinders $[u]$, $u \in \mathscr{W}(r)$, partition the full shift, so the Bernoulli probability with one-step weights $\theta(q)$ gives

$$\sum_{u\in\mathscr{W}(r)} \theta_u(q) = 1. \tag{I.10.50}$$

By (I.10.40),

$$p_u^q = \theta_u(q) r_u^{\tau(q)} \asymp_q \theta_u(q) r^{\tau(q)}. \tag{I.10.51}$$

Summing (I.10.51) and using (I.10.50) gives

$$Z(r,q) \asymp_q r^{\tau(q)}. \tag{I.10.52}$$

Lemma 2.1 and division by $\log r$ prove the centered-packing assertion and (I.10.47).

Let

$$F(q,t) = \sum_i \exp(q \log p_i - t \log r_i) - 1. \tag{I.10.53}$$

At $t = \tau(q)$,

$$\partial_t F(q, \tau(q)) = -\sum_i \theta_i(q) \log r_i > 0. \tag{I.10.54}$$

Here is a power-series proof of analyticity, so no analytic implicit-function theorem is left as an undeclared input. Fix q_0, put $t_0 = \tau(q_0)$, $a_i = \log p_i$, $b_i = -\log r_i > 0$, and $w_i = p_i^{q_0} r_i^{-t_0}$. Then $\sum_i w_i = 1$ and $c = \sum_i w_i b_i > 0$. For $\rho > 0$, let $\mathscr{A}_\rho$ be the space of real coefficient sequences identified with series $u(z) = \sum_{j\geq 0} u_j z^j$, with norm

$$\|u\|_\rho = \sum_{j\geq 0} |u_j| \rho^j < \infty. \tag{I.10.54a}$$

It is complete by completeness of ℓ^1. The usual coefficient convolution defines multiplication, and absolute summation gives $\|uv\|_\rho \leq \|u\|_\rho \|v\|_\rho$. Thus $e^u = \sum_{k\geq 0} u^k/k!$ converges in this norm and $\|e^u - 1\|_\rho \leq e^{\|u\|_\rho} - 1$. The series z has norm ρ. Define

$$\mathscr{G}(u) = u - c^{-1}\left(\sum_i w_i e^{a_i z + b_i u} - 1\right). \tag{I.10.54b}$$

Choose $R > 0$, then $\rho > 0$, small enough that

$$c^{-1} \sum_i w_i b_i (e^{|a_i|\rho + b_i R} - 1) \leq \tfrac{1}{2}, \qquad c^{-1} \sum_i w_i (e^{|a_i|\rho} - 1) \leq R/2. \tag{I.10.54c}$$

For $\|u\|_\rho, \|v\|_\rho \leq R$, differentiate the absolutely convergent exponential series along the segment from v to u and integrate on $[0,1]$. Since $c = \sum_i w_i b_i$, (I.10.54c) gives $\|\mathscr{G}(u) - \mathscr{G}(v)\|_\rho \leq \frac{1}{2}\|u - v\|_\rho$. Also $\|\mathscr{G}(0)\|_\rho \leq R/2$, so this closed ball is invariant. Iterating from zero gives successive differences at most $2^{-k}\|\mathscr{G}(0)\|_\rho$; completeness gives a fixed series $u \in \mathscr{A}_\rho$. Evaluation at any real $|z| < \rho$ preserves sums and products by absolute convergence, so

$$\sum_i w_i e^{a_i z + b_i u(z)} = 1, \qquad \tau(q_0 + z) = t_0 + u(z). \tag{I.10.54d}$$

The second identity is the already proved uniqueness of the real pressure root. Thus τ equals a convergent real power series near every q_0, which is the meaning of real analytic. Differentiating (I.10.46) gives

$$0 = \sum_i \theta_i(q)\big(\log p_i - \tau'(q)\log r_i\big), \tag{I.10.55}$$

which is (I.10.49). A second differentiation yields

$$\tau''(q) = \frac{\sum_i \theta_i(q)\big(\log p_i - \tau'(q)\log r_i\big)^2}{\sum_i \theta_i(q)\log r_i} \le 0. \tag{I.10.56}$$

Thus τ is concave. □

I.10.3 The complete self-similar spectrum

For a probability vector $\theta = (\theta_i)_{i\in\Lambda}$, set

$$H(\theta) = -\sum_i \theta_i \log\theta_i, \quad \chi(\theta) = -\sum_i \theta_i \log r_i, \quad \alpha(\theta) = \frac{\sum_i \theta_i \log p_i}{\sum_i \theta_i \log r_i}, \quad s(\theta) = \frac{H(\theta)}{\chi(\theta)}. \tag{I.10.57}$$

As usual $0\log 0 = 0$.

Lemma — 3.1 (symbolic frequency level sets) [F]

Let

$$\Sigma(\theta) = \Big\{\omega \,:\, \lim_{n\to\infty}\frac{1}{n}\#\{1 \le j \le n \,:\, \omega_j = i\} = \theta_i \text{ for every } i\Big\}. \tag{I.10.58}$$

Then

$$\pi(\Sigma(\theta)) \subset E_\mu(\alpha(\theta)), \qquad \dim_{\mathrm{H}} \pi(\Sigma(\theta)) \ge s(\theta). \tag{I.10.59}$$

Proof

If $\omega \in \Sigma(\theta)$, then

$$\frac{\log p_{\omega|_n}}{\log r_{\omega|_n}} = \frac{\sum_i N_i(\omega|_n)\log p_i}{\sum_i N_i(\omega|_n)\log r_i} \longrightarrow \alpha(\theta). \tag{I.10.60}$$

The SSC ball–cylinder comparison from I.8, together with $r_* \le r_{\omega|_{n+1}}/r_{\omega|_n} \le \max_i r_i < 1$, converts (I.10.60) into

$$d_\mu(\pi\omega) = \alpha(\theta). \tag{I.10.61}$$

Let $I = \{i : \theta_i > 0\}$ and take the Bernoulli measure with weights $(\theta_i)_{i\in I}$ on the subshift $I^{\mathbb{N}}$. Birkhoff's theorem gives full measure to $\Sigma(\theta)$. Its projection is exact dimensional by T-I.8.05 and has dimension

$$\frac{H(\theta)}{\chi(\theta)} = s(\theta). \tag{I.10.62}$$

Since that projection is carried by $\pi(\Sigma(\theta))$, the local-to-global dimension theorem of I.4 proves the second part of (I.10.59). □

Lemma — 3.2 (finite-dimensional Legendre duality) [F]

Let

$$a_i = \frac{\log p_i}{\log r_i}, \qquad \alpha_{\min} = \min_i a_i, \qquad \alpha_{\max} = \max_i a_i. \tag{I.10.63}$$

For every $\alpha \in [\alpha_{\min}, \alpha_{\max}]$,

$$\max_{\theta : \alpha(\theta)=\alpha} s(\theta) = \inf_{q\in\mathbb{R}}\{q\alpha - \tau(q)\}. \tag{I.10.64}$$

The maximum exists. Formula (I.10.64) includes both endpoints.

Proof

The constraint set is nonempty: choose $w_i \geq 0$ with $\sum_i w_i = 1$ and $\sum_i w_i a_i = \alpha$, and put

$$\theta_i = \frac{w_i/(-\log r_i)}{\sum_j w_j/(-\log r_j)}.$$

Then $\alpha(\theta) = \alpha$. The constraint set is a closed subset of the probability simplex, and $\chi(\theta) \geq -\log(\max_i r_i) > 0$. Hence s is continuous and the maximum exists.

For fixed q, the vector $\theta(q)$ from (I.10.48) is strict. Relative entropy gives, for every probability vector θ,

$$\begin{aligned} 0 \leq D(\theta\|\theta(q)) &= \sum_i \theta_i \log \frac{\theta_i}{\theta_i(q)} \\ &= -H(\theta) - q\sum_i \theta_i \log p_i + \tau(q)\sum_i \theta_i \log r_i. \end{aligned} \tag{I.10.65}$$

If $\alpha(\theta) = \alpha$, division by $\chi(\theta) = -\sum_i \theta_i \log r_i$ yields

$$s(\theta) \leq q\alpha - \tau(q). \tag{I.10.66}$$

Thus the left side of (I.10.64) is at most the right side.

The derivative τ' is continuous and nonincreasing. From (I.10.49), $\tau'(q) = \alpha(\theta(q))$. We next show

$$\lim_{q\to+\infty} \tau'(q) = \alpha_{\min}, \qquad \lim_{q\to-\infty} \tau'(q) = \alpha_{\max}. \tag{I.10.67}$$

For the first limit, put $I_- = \{i : a_i = \alpha_{\min}\}$ and $\beta(q) = \tau(q) - q\alpha_{\min}$. Equation (I.10.46) becomes

$$\sum_i \exp\big(q(a_i - \alpha_{\min}) \log r_i\big) r_i^{-\beta(q)} = 1. \tag{I.10.68}$$

The factors outside I_- tend to zero as $q \to +\infty$. Monotonicity in β, with two fixed comparison values, shows that

$$\beta(q) \longrightarrow \beta_-, \qquad \sum_{i\in I_-} r_i^{-\beta_-} = 1. \tag{I.10.69}$$

Concavity then forces the derivative limit in (I.10.67); equivalently one may take weighted averages in (I.10.49). The second limit follows identically with $I_+ = \{i : a_i = \alpha_{\max}\}$ and $q \to -\infty$.

If α is in the open interval, (I.10.67) and continuity give a q such that $\tau'(q) = \alpha$. Taking $\theta = \theta(q)$ makes (I.10.65) an equality, so (I.10.64) follows.

At $\alpha = \alpha_{\min}$, the identity

$$\alpha(\theta) = \sum_i \frac{\theta_i(-\log r_i)}{\chi(\theta)} a_i \tag{I.10.70}$$

shows that the constraint is equivalent to support on I_-. Put $s_- = -\beta_-$, so

$$\sum_{i\in I_-} r_i^{s_-} = 1. \tag{I.10.71}$$

The same relative-entropy calculation on the face I_-, now with weights $r_i^{s_-}$, gives

$$\max_{\operatorname{supp}\theta\subset I_-} s(\theta) = s_-. \tag{I.10.72}$$

On the other hand, (I.10.69) gives

$$\lim_{q\to+\infty} \{q\alpha_{\min} - \tau(q)\} = -\beta_- = s_-. \tag{I.10.73}$$

Together with (I.10.66), this proves (I.10.64) at the left endpoint. The right endpoint is identical, using I_+ and $q \to -\infty$. □

Theorem — 3.3 (T-I.10.04: full self-similar spectrum) [F]

The level set $E_\mu(\alpha)$ is empty when $\alpha \notin [\alpha_{\min}, \alpha_{\max}]$. For every α in that closed interval,

$$\dim_{\mathrm{H}} E_\mu(\alpha) = \max_{\theta:\alpha(\theta)=\alpha} \frac{H(\theta)}{\chi(\theta)} = \inf_{q\in\mathbb{R}}\{q\alpha - \tau_\mu^{\mathrm{pack}}(q)\}. \tag{I.10.74}$$

Proof

For every word u,

$$\frac{\log p_u}{\log r_u} = \sum_i \frac{N_i(u)(-\log r_i)}{-\log r_u} a_i \in [\alpha_{\min}, \alpha_{\max}]. \tag{I.10.75}$$

The SSC ball–cylinder comparison used in Lemma 3.1 implies that whenever a local dimension exists, it is a limit of the ratios in (I.10.75). This proves the emptiness assertion.

For an admissible α, Lemma 3.1 gives

$$\dim_{\mathrm{H}} E_\mu(\alpha) \geq \max_{\theta : \alpha(\theta)=\alpha} s(\theta). \tag{I.10.76}$$

The universal upper bound, Theorem 1.6, and Theorem 2.2 give

$$\dim_{\mathrm{H}} E_\mu(\alpha) \leq \inf_{q\in\mathbb{R}} \{q\alpha - \tau(q)\}. \tag{I.10.77}$$

Lemma 3.2 identifies the right sides of (I.10.76)–(I.10.77), proving (I.10.74), including at both endpoints. □

Point-of-statement provenance. The pressure/Legendre architecture is PW97-MF, Theorem 5, local PDF p. 33 / printed p. 265. Theorem 3.3 is the fully proved strict Bernoulli self-similar specialization, with the endpoint faces and negative-order normalization made explicit.

Corollary — 3.4 (C-I.10.05: generalized dimensions) [F]

For $q > 0$, $q \neq 1$,

$$D_q^{\mathrm{dyad}}(\mu) = D_q^{\mathrm{pack}}(\mu) = \frac{\tau(q)}{q-1}. \tag{I.10.78}$$

At $q = 1$,

$$D_1^{\mathrm{dyad}}(\mu) = \tau'(1) = \frac{-\sum_i p_i \log p_i}{-\sum_i p_i \log r_i}. \tag{I.10.79}$$

At $q = 0$, both zero-order dimensions equal the similarity dimension s determined by

$$\sum_i r_i^s = 1. \tag{I.10.80}$$

For $q < 0$, (I.10.9) and (I.10.46) give the centered-packing generalized dimension; no dyadic equality is asserted. Finally, the following are equivalent:

(1) $\log p_i / \log r_i$ is independent of i;

(2) τ is affine;

(3) all centered-packing generalized dimensions, together with D_1^{dyad}, have one common value;

(4) every point of K has the same local dimension.

Proof

Equation (I.10.78) is (I.10.6)–(I.10.9) combined with (I.10.47). Since $\tau(1) = 0$, the entropy-dimension limit follows from a squeeze, not an interchange of the limits in n and q. For $q_- < 1 < q_+$ with $q_{\to}0$, Lemma 1.5 gives

$$\frac{H_{q_+}(\mu, \mathscr{D}_n)}{n \log 2} \le \frac{H_1(\mu, \mathscr{D}_n)}{n \log 2} \le \frac{H_{q_-}(\mu, \mathscr{D}_n)}{n \log 2}.$$

The outside limits exist by (I.10.47). Consequently

$$\frac{\tau(q_+)}{q_+ - 1} \le \underline{D}_1^{\text{dyad}}(\mu) \le \overline{D}_1^{\text{dyad}}(\mu) \le \frac{\tau(q_-)}{q_- - 1}.$$

Let q_+ and q_- tend to one. Differentiability of τ and $\tau(1) = 0$ make both endpoints tend to $\tau'(1)$. Formula (I.10.49), with $\theta_i(1) = p_i$, now gives (I.10.79). At $q = 0$, equation (I.10.46) is $\sum_i r_i^{-\tau(0)} = 1$, so $-\tau(0) = s$; Lemma 2.1 gives the dyadic zero-order value.

If all $a_i = \log p_i / \log r_i$ equal c, then $p_i = r_i^c$, (I.10.46) gives

$$\tau(q) = c(q - 1), \tag{I.10.81}$$

and T-I.8.06 says that μ is c-Ahlfors regular. Thus all generalized dimensions and all pointwise local dimensions equal c.

Conversely, (I.10.56) shows that an affine τ forces

$$\log p_i - \tau'(q) \log r_i = 0 \quad \text{for every } i, \tag{I.10.82}$$

so all a_i are equal. If all generalized dimensions are equal, then $\tau(q) = c(q - 1)$ for all $q \ne 1$, hence τ is affine. If every point has the same local dimension, apply (I.10.61) to each constant sequence i^∞: its local dimension is a_i, so all a_i agree. This closes the cycle of implications. □

I.10.4 Hölder-Gibbs measures under strong separation

Let $\Sigma = \Lambda^{\mathbb{N}}$, let σ be the left shift, and let Φ be a finite $C^{1+\gamma}$ conformal IFS with SSC, where $0 < \gamma \le 1$. Its geometric potential from I.9 is

$$\phi(\omega) = \log \|DS_{\omega_1}(\pi(\sigma\omega))\| < 0. \tag{I.10.83}$$

Let $\psi : \Sigma \to \mathbb{R}$ be Hölder and let m_ψ be its equilibrium/Gibbs state.

Definition — 4.1 (normalized potential and geometric pressure root)

Put

$$\psi_0 = \psi - P(\psi), \qquad P(\psi_0) = 0. \tag{I.10.84}$$

For $q \in \mathbb{R}$, the geometric pressure root $T_\psi(q)$ is the unique real t such that

$$P(q\psi_0 - t\phi) = 0. \tag{I.10.85}$$

For this section write $T = T_\psi$, $\mu = \pi_* m_\psi$, and let m_q denote the equilibrium state of

$$\eta_q = q\psi_0 - T(q)\phi. \tag{I.10.86}$$

When T is differentiable, define the extremal slopes

$$\alpha_{\min} = \lim_{q\to+\infty} T'(q), \qquad \alpha_{\max} = \lim_{q\to-\infty} T'(q). \tag{I.10.87}$$

The interval $(\alpha_{\min}, \alpha_{\max})$, possibly empty, is the interior slope range.

Lemma — 4.2 (pressure-root regularity without analytic perturbation) [F]

The root in (I.10.85) exists uniquely for every q. The function $T : \mathbb{R} \to \mathbb{R}$ is finite, concave, continuously differentiable, and

$$T'(q) = \frac{\int \psi_0 \, dm_q}{\int \phi \, dm_q}. \tag{I.10.88}$$

Proof

There are constants $0 < c_\phi \le C_\phi < \infty$ such that

$$-C_\phi \le \phi \le -c_\phi < 0. \tag{I.10.89}$$

Pressure is monotone in the potential and changes by at most the uniform norm of a perturbation. Hence, for fixed q, the function $t \mapsto P(q\psi_0 - t\phi)$ is continuous and satisfies

$$P(q\psi_0 - t\phi) - P(q\psi_0 - s\phi) \in [c_\phi(t-s), C_\phi(t-s)] \quad (t > s). \tag{I.10.90}$$

It tends to opposite infinities as $t \to \pm\infty$, which proves existence and uniqueness.

For an invariant probability ν, abbreviate

$$A_\nu = \int \psi_0 \, d\nu, \qquad B_\nu = \int \phi \, d\nu < 0. \tag{I.10.91}$$

The variational principle and (I.10.85) say

$$h_\nu(\sigma) + qA_\nu - T(q)B_\nu \le 0, \tag{I.10.92}$$

with equality exactly at $\nu = m_q$. Dividing by $B_\nu < 0$ gives

$$T(q) = \inf_{\nu \in \mathcal{M}_\sigma(\Sigma)} \frac{h_\nu(\sigma) + qA_\nu}{B_\nu}. \tag{I.10.93}$$

An infimum of affine functions of q is concave. Also $|A_\nu/B_\nu| \le \|\psi_0\|_\infty / c_\phi$, so (I.10.93) makes T globally Lipschitz.

We record the compactness argument needed below. From any sequence of invariant probabilities, diagonal extraction on the countable cylinder algebra produces limiting cylinder masses. Their finite additivity and consistency define an invariant probability ν

on Σ, and the subsequence converges weakly to ν. For the first-coordinate partition $\mathscr{P}$, the accepted finite-generator entropy identity is

$$h_\nu(\sigma) = \inf_{n\geq 1} \frac{1}{n} H_\nu(\mathscr{P}_0^{n-1}). \tag{I.10.94}$$

Each finite-block entropy on the right is continuous in the cylinder masses. An infimum of continuous functions is upper semicontinuous; hence

$$\nu_j \longrightarrow \nu \quad \Longrightarrow \quad \limsup_j h_{\nu_j}(\sigma) \leq h_\nu(\sigma). \tag{I.10.95}$$

Suppose $q_j \to q$ and a subsequence of m_{q_j} converges to m. The continuity of T, the continuity of the two potentials, and (I.10.95) give

$$\begin{aligned} h_m(\sigma) + qA_m - T(q)B_m &\geq \limsup_j \{h_{m_{q_j}}(\sigma) + q_j A_{m_{q_j}} - T(q_j)B_{m_{q_j}}\} \\ &= 0. \end{aligned} \tag{I.10.96}$$

The variational principle makes the left side at most zero. Thus m is an equilibrium state for η_q. Uniqueness in A-DS gives $m = m_q$. Every subsequential limit is the same, so

$$m_{q_j} \longrightarrow m_q. \tag{I.10.97}$$

Let $h > 0$. Evaluating the pressure at $q + h$ on m_q, and at q on m_{q+h}, gives

$$\frac{A_{m_{q+h}}}{B_{m_{q+h}}} \leq \frac{T(q+h) - T(q)}{h} \leq \frac{A_{m_q}}{B_{m_q}}. \tag{I.10.98}$$

Indeed, both inequalities follow directly from (I.10.92), with the reversal when dividing by the negative B. By (I.10.97), the two outer terms tend to the same value as $h \downarrow 0$. The corresponding comparison for $h < 0$ proves differentiability and (I.10.88). Continuity of the quotient in (I.10.88), together with (I.10.97), proves continuity of T'. □

Lemma — 4.3 (packing and dyadic pressure formulas) [F]

For every $q \in \mathbb{R}$,

$$\tau_\mu^{\text{pack}}(q) = T(q). \tag{I.10.99}$$

For every $q > 0$,

$$\tau_\mu^{\text{dyad}}(q) = T(q), \tag{I.10.100}$$

and the zero-order dyadic exponent agrees with the centered-packing exponent of K. No negative-order fixed-grid conclusion is asserted.

Proof

The Gibbs property for m_ψ, normalization (I.10.84), the Gibbs property for m_q, and the Bowen bounds for ψ_0, ϕ give, uniformly for $\omega \in [u]$,

$$m_\psi([u]) \asymp e^{S_{|u|}\psi_0(\omega)}, \qquad m_q([u]) \asymp e^{qS_{|u|}\psi_0(\omega) - T(q)S_{|u|}\phi(\omega)}. \tag{I.10.101}$$

I.9 bounded distortion also gives

$$\operatorname{diam} K_u \asymp e^{S_{|u|}\phi(\omega)}. \tag{I.10.102}$$

Let

$$\mathscr{W}(r) = \{u \,:\, \operatorname{diam} K_u \le r < \operatorname{diam} K_{u^-}\}. \tag{I.10.103}$$

Restrict (I.10.103) to nonempty words and $0 < r < \operatorname{diam} K$. The case of a one-symbol alphabet has a singleton attractor, $\psi_0 = 0$, $T = 0$ and all spectra zero; its interior slope range is empty, so it requires no stopping construction. For at least two symbols, SSC gives $\operatorname{diam} K > 0$.

Here bounded distortion, not just the one-step conorm, gives the child-to-parent comparison. Write $u = vi$, and let C_{bd} be a cylinder distortion constant from I.9. Choosing two points in $S_i(K)$ and comparing their image separation under S_v with $\operatorname{diam} K_v$ gives

$$\frac{\operatorname{diam} K_{vi}}{\operatorname{diam} K_v} \ge C_{\mathrm{bd}}^{-2} \frac{\operatorname{diam} S_i(K)}{\operatorname{diam} K}.$$

The right side has a positive minimum over the finite alphabet. Thus

$$cr \le \operatorname{diam} K_u \le r \quad (u \in \mathscr{W}(r)). \tag{I.10.104}$$

SSC and bounded distortion give the same uniform separation and bounded incidence properties as Lemma 2.1. Consequently, for all real q, the centered packing moment at a fixed multiple of r is comparable to

$$Z_\psi(r, q) = \sum_{u \in \mathscr{W}(r)} m_\psi([u])^q. \tag{I.10.105}$$

For $q > 0$, the dyadic moment at side comparable to r is comparable to the same sum. The negative-q exclusion is exactly the grid-boundary obstruction identified after (I.10.46a).

Equations (I.10.101)–(I.10.104) yield

$$m_\psi([u])^q \asymp_q m_q([u])\, r^{T(q)} \quad (u \in \mathscr{W}(r)). \tag{I.10.106}$$

The stopping cylinders partition Σ, so $\sum_{u \in \mathscr{W}(r)} m_q([u]) = 1$. Summing (I.10.106) gives

$$Z_\psi(r, q) \asymp_q r^{T(q)}. \tag{I.10.107}$$

Taking logarithmic exponents proves (I.10.99)–(I.10.100). At $q = 0$, both counts have the box exponent of the SSC attractor, by the same bounded incidence argument. ☐

Theorem — 4.4 (T-I.10.06: Hölder-Gibbs interior formalism) [F]

For every $\alpha \in (\alpha_{\min}, \alpha_{\max})$, there is $q \in \mathbb{R}$ with $\alpha = T'(q)$, and

$$\dim_{\mathrm{H}} E_\mu(\alpha) = q\alpha - T(q) = \inf_{s\in\mathbb{R}}\{s\alpha - T(s)\}. \tag{I.10.108}$$

The probability $\nu_q = \pi_* m_q$ is carried by $E_\mu(\alpha)$, is exact dimensional, and

$$\dim_{\mathrm{H}} \nu_q = q\alpha - T(q). \tag{I.10.109}$$

At $\alpha_{\min}$ or $\alpha_{\max}$, Theorem 1.6 supplies the upper bound, but this theorem asserts no endpoint equality without an additional proved endpoint hypothesis.

Proof

The derivative T' is continuous and nonincreasing. Its image contains the open interval between its two limits in (I.10.87), so an appropriate q exists.

For m_q-almost every ω, Birkhoff's theorem and (I.10.88) give

$$\frac{S_n\psi_0(\omega)}{S_n\phi(\omega)} \longrightarrow \frac{\int \psi_0\, dm_q}{\int \phi\, dm_q} = T'(q) = \alpha. \tag{I.10.110}$$

The first Gibbs estimate in (I.10.101), (I.10.102), SSC, and the bounded ratio between consecutive cylinder scales convert (I.10.110) into

$$d_\mu(\pi\omega) = \alpha. \tag{I.10.111}$$

Thus $\nu_q(E_\mu(\alpha)) = 1$.

The measure m_q is an ergodic invariant Gibbs state by the allowed dynamics theorem. T-I.9.08 applies to every such invariant measure, and T-I.9.05 identifies its projection entropy with ordinary entropy because SSC makes coding injective. Therefore ν_q is exact dimensional and

$$\dim_{\mathrm{H}} \nu_q = \frac{h_{m_q}(\sigma)}{-\int \phi\, dm_q}. \tag{I.10.112}$$

The pressure-zero identity for m_q is

$$h_{m_q}(\sigma) + q\int \psi_0\, dm_q - T(q)\int \phi\, dm_q = 0. \tag{I.10.113}$$

Use (I.10.88) in (I.10.112)–(I.10.113) to obtain

$$\dim_{\mathrm{H}} \nu_q = qT'(q) - T(q) = q\alpha - T(q). \tag{I.10.114}$$

Since ν_q is carried by the level set, (I.10.114) gives the lower bound in (I.10.108). Theorem 1.6 and Lemma 4.3 give

$$\dim_{\mathrm{H}} E_\mu(\alpha) \le \inf_s\{s\alpha - T(s)\} \le q\alpha - T(q). \tag{I.10.115}$$

Finally, concavity and $\alpha = T'(q)$ imply $T(s) \le T(q) + \alpha(s-q)$, so the middle infimum in (I.10.115) equals $q\alpha - T(q)$. This proves all assertions. □

Point-of-statement provenance. Theorem 4.4 is the finite-full-shift IFS specialization of PW97-MF, Theorems 1 and 4, local PDF pp. 14–15 and 27 / printed pp. 246–247 and 259. Lemma 4.2 replaces analytic perturbation by a local variational differentiability proof, and the endpoint restriction is part of the printed theorem here.

I.10.5 Subpower Gibbs balls without separation

Definition — 5.1 (quasi-Bernoulli source)

A Borel probability m on Σ is upper quasi-Bernoulli if some $C_+ \geq 1$ satisfies

$$m([uv]) \leq C_+ m([u])m([v]) \quad (u, v \in \Lambda^*). \tag{I.10.116}$$

It is lower quasi-Bernoulli if some $C_- > 0$ satisfies

$$m([uv]) \geq C_- m([u])m([v]). \tag{I.10.117}$$

It is two-sided quasi-Bernoulli if both hold. Zero-probability letters are deleted from the alphabet. Every finite-full-shift Hölder Gibbs state is two-sided quasi-Bernoulli. To see this, let its Gibbs constant be C_G, choose $\omega \in [uv]$, and use the exact identity

$$S_{|uv|}\psi(\omega) = S_{|u|}\psi(\omega) + S_{|v|}\psi(\sigma^{|u|}\omega).$$

Apply the Gibbs estimate separately to the three cylinders at these representatives. The pressure terms add in the same way, so

$$C_G^{-3} m([u])m([v]) \leq m([uv]) \leq C_G^3 m([u])m([v]).$$

Definition — 5.2 (subexponential and subpower errors)

A positive sequence (c_n) is subexponential if

$$\lim_{n\to\infty} \frac{\log c_n}{n} = 0. \tag{I.10.118}$$

A positive function $h : (0, r_0) \to (0, \infty)$ is subpower at zero if

$$\lim_{r\downarrow 0} \frac{\log h(r)}{\log(1/r)} = 0. \tag{I.10.119}$$

Given $q > 0$, an exponent $\tau(q)$, and $c > 1$, an auxiliary probability ν_q satisfies the upper q-Gibbs ball inequality if

$$\nu_q(B(x, c^{-1}r)) \leq h(r) r^{-\tau(q)} \mu(B(x, r))^q \tag{I.10.120}$$

for every x and all sufficiently small r, with h subpower. It satisfies the lower q-Gibbs ball inequality if

$$\nu_q(B(x, cr)) \geq h(r) r^{-\tau(q)} \mu(B(x, r))^q \tag{I.10.120a}$$

for every x and all sufficiently small r, again with h subpower. The radius enlargement and the direction of the inequality are part of the definition.

Lemma — 5.3 (disjoint enlarged-ball moment bound) [F]

Fix $q > 0$ and $c > 1$. If $\{B(x_j, c^{-1}2^{-n})\}_j$ is pairwise disjoint, then

$$\sum_j \mu(B(x_j, 2^{-n}))^q \leq C(d, c, q) M_n(\mu, q). \tag{I.10.121}$$

More generally, if these smaller balls have multiplicity at most L, the same inequality holds with $LC(d, c, q)$ in place of $C(d, c, q)$.

Proof

Each ball $B(x_j, 2^{-n})$ meets at most 3^d cubes in $\mathcal{D}_n$. Conversely, if a cube $Q \in \mathcal{D}_n$ meets $B(x_j, 2^{-n})$, then the disjoint ball $B(x_j, c^{-1}2^{-n})$ lies in a fixed enlargement of Q. Volume comparison shows that at most $C_0(d, c)$ indices j can be associated with one Q. The finite-vector inequality

$$\left(\sum_{\ell=1}^{k} a_\ell\right)^q \leq \max(1, k^{q-1}) \sum_{\ell=1}^{k} a_\ell^q \tag{I.10.122}$$

first over cubes meeting a ball and then over balls meeting a cube proves (I.10.121). For the multiplicity variant, sum the Euclidean volumes of the smaller balls associated with one cube. They lie in the same fixed enlarged cube, and their pointwise multiplicity is at most L; their total volume is therefore at most L times the volume of that enlargement. This changes the association bound by the factor L and proves the stated variant. □

Lemma — 5.4 (L-I.10.07: abstract Gibbs-ball criterion) [F]

Let μ be compactly supported. Suppose that its lower centered-packing spectrum is finite and concave on $\mathbb{R}$, and that on an open interval

$$(a, b) \subset (0, \infty) \tag{I.10.123}$$

it agrees with an existing dyadic spectrum; denote the common function by τ. Fix one $q \in (a, b)$, and suppose an auxiliary probability ν_q satisfies (I.10.120). No auxiliary probabilities at neighboring parameters are required. Then, for ν_q-almost every x,

$$\underline{d}_\mu(x) \geq \tau'(q+), \qquad \overline{d}_\mu(x) \leq \tau'(q-). \tag{I.10.124}$$

If τ is differentiable at q and $\alpha = \tau'(q)$, then

$$\nu_q(E_\mu(\alpha)) = 1, \qquad \underline{d}_{\nu_q}(x) \geq q\alpha - \tau(q) \quad \text{for } \nu_q\text{-almost every } x, \tag{I.10.125}$$

and

$$\dim_{\mathrm{H}} E_\mu(\alpha) = q\alpha - \tau(q) = \inf_{s\in\mathbb{R}}\{s\alpha - \underline{\tau}_\mu^{\mathrm{pack}}(s)\}. \tag{I.10.126}$$

Proof

First ν_q is supported on $\mathrm{supp}(\mu)$. At a point outside that closed set, take r small enough that $\mu(B(x,r)) = 0$; (I.10.120), since $q > 0$, gives $\nu_q(B(x,r/c)) = 0$. A countable subcover of the open complement proves the support assertion. Fix $q \in (a,b)$, let c be the constant in (I.10.120), and fix $\gamma > 0$. Put

$$\beta_+ = \tau'(q+), \qquad F_n = \{x : \mu(B(x,2^{-n})) \ge 2^{-n(\beta_+-\gamma)}\}. \tag{I.10.127}$$

Apply the Besicovitch covering theorem L-I.9.03 to the balls $B(x,c^{-1}2^{-n})$, $x \in F_n$. It produces a covering family of multiplicity at most B_d; a decomposition into disjoint families is not asserted by the cited version. For $u > 0$, (I.10.120), (I.10.127), and the bounded-multiplicity version of Lemma 5.3 give

$$\begin{aligned}\nu_q(F_n) &\le Ch(2^{-n})2^{n\tau(q)}\sum_j \mu(B(x_j,2^{-n}))^q \\ &\le Ch(2^{-n})2^{n\{\tau(q)+u(\beta_+-\gamma)\}}M_n(\mu,q+u).\end{aligned} \tag{I.10.128}$$

Choose $u > 0$ so small that $q+u \in (a,b)$ and

$$\frac{\tau(q+u)-\tau(q)}{u} \ge \beta_+ - \frac{\gamma}{4}. \tag{I.10.129}$$

Since the dyadic spectrum exists at $q+u$, and h is subpower, for all large n,

$$M_n(\mu,q+u) \le 2^{-n\{\tau(q+u)-u\gamma/4\}}, \qquad h(2^{-n}) \le 2^{nu\gamma/4}. \tag{I.10.130}$$

Substitution into (I.10.128) gives

$$\nu_q(F_n) \le C2^{-nu\gamma/4}. \tag{I.10.131}$$

Thus $\sum_n \nu_q(F_n) < \infty$. Borel–Cantelli gives $\nu_q(\limsup_n F_n) = 0$.

The lower local dimension can be computed along 2^{-n}. Indeed, when $2^{-(n+1)} < r \le 2^{-n}$, ball monotonicity sandwiches the mass between the two adjacent dyadic radii, and

$$\frac{\log 2^{-(n+1)}}{\log 2^{-n}} \longrightarrow 1.$$

Hence

$$\{x : \underline{d}_\mu(x) < \beta_+ - \gamma\} \subset \limsup_n F_n. \tag{I.10.132}$$

It follows that $\underline{d}_\mu(x) \ge \beta_+ - \gamma$ almost everywhere. Letting $\gamma \downarrow 0$ proves the first inequality in (I.10.124).

For the second inequality put $\beta_- = \tau'(q-)$ and

$$G_n = \{x \in \operatorname{supp}\mu \,:\, \mu(B(x,2^{-n})) \le 2^{-n(\beta_-+\gamma)}\}. \tag{I.10.133}$$

Choose $0 < u < q$, $q - u \in (a,b)$. The same Besicovitch argument, now using

$$\mu(B)^q = \mu(B)^{q-u}\mu(B)^u \le \mu(B)^{q-u}2^{-nu(\beta_-+\gamma)} \quad \text{on } G_n, \tag{I.10.134}$$

gives

$$\nu_q(G_n) \le Ch(2^{-n})2^{n\{\tau(q)-u(\beta_-+\gamma)\}}M_n(\mu, q-u). \tag{I.10.135}$$

Choose u so small that

$$\frac{\tau(q) - \tau(q-u)}{u} \le \beta_- + \frac{\gamma}{4}. \tag{I.10.136}$$

The analogues of (I.10.130) again yield

$$\nu_q(G_n) \le C2^{-nu\gamma/4}. \tag{I.10.137}$$

Borel–Cantelli and dyadic-radius interpolation show that $\overline{d}_\mu(x) \le \beta_- + \gamma$ almost everywhere. Let $\gamma \downarrow 0$, proving (I.10.124).

At a differentiability point, (I.10.124) proves the first assertion of (I.10.125). Apply (I.10.120) with r replaced by cr:

$$\nu_q(B(x,r)) \le h(cr)(cr)^{-\tau(q)}\mu(B(x,cr))^q. \tag{I.10.138}$$

Take logarithms, divide by $\log r < 0$, and use the subpower property and $\log(cr)/\log r \to 1$. At every point where $d_\mu(x) = \alpha$,

$$\underline{d}_{\nu_q}(x) \ge q\alpha - \tau(q). \tag{I.10.139}$$

The local-to-global theorem of I.4, applied to the restriction of ν_q to its full-measure level set, gives

$$\dim_{\mathrm{H}} E_\mu(\alpha) \ge q\alpha - \tau(q). \tag{I.10.140}$$

Theorem 1.6 gives the reverse inequality at q. Finally, concavity and $\alpha = \tau'(q)$ imply

$$\underline{\tau}_\mu^{\mathrm{pack}}(s) \le \tau(q) + \alpha(s-q) \quad (s \in \mathbb{R}). \tag{I.10.141}$$

Thus every term in the infimum in (I.10.126) is at least $q\alpha - \tau(q)$, while equality holds at $s = q$. This proves (I.10.126). □

Point-of-statement provenance. Lemma 5.4 expands FENG07-MF, Proposition 2.1, local PDF pp. 6–10. The generic upper bound invoked but not proved in that paper is Theorem 1.6 here.

I.10.6 Subpower geometry for C^1 conformal systems

Let $\Phi = (S_i)_{i\in\Lambda}$ be a finite C^1 conformal IFS on the compact set $X \subset \mathbb{R}^d$, in the exact sense of I.9. Let K be its attractor and assume in this section that $\operatorname{diam} K > 0$. The singleton case is handled directly in Theorems 8.1–8.2 below.

Definition — 6.1 (neighbors and conformal stopping family)

For $Q, Q' \in \mathscr{D}_n$, write

$$Q \sim Q' \quad \Longleftrightarrow \quad \overline{Q} \cap \overline{Q'} \neq \varnothing. \tag{I.10.142}$$

For $n \geq 1$ with $2^{-n} < \operatorname{diam} K$, define

$$\mathscr{W}_n = \{u \in \Lambda^*,\ |u| \geq 1 \ :\ \operatorname{diam} K_u \leq 2^{-n} < \operatorname{diam} K_{u^-}\}. \tag{I.10.143}$$

Here u^- removes the last letter. If instead $2^{-n} \geq \operatorname{diam} K$, set $\mathscr{W}_n = \{\varnothing\}$, where the empty word has $S_\varnothing = id$ and $K_\varnothing = K$; no parent is assigned to it. There are only finitely many such coarse generations. The cylinders $[u]$, $u \in \mathscr{W}_n$, form a finite antichain partition of Σ. No separation condition is included.

Lemma — 6.2 (uniform C^1 subpower quasi-similarity) [F]

Choose $\eta > 0$ so small that the closed η-neighborhood U_0 of X is contained in the common open domain. This compact neighborhood is forward invariant: contraction and $S_i(X) \subset X$ give $\operatorname{dist}(S_i y, X) \leq \vartheta \operatorname{dist}(y, X) \leq \vartheta\eta$ for $y \in U_0$. For every $\varepsilon > 0$, there is $C_\varepsilon \geq 1$ such that, for every word u, $x \in X$, and $y \in U_0$,

$$C_\varepsilon^{-1} e^{-\varepsilon|u|} \|DS_u(x)\|\,|x-y| \leq |S_u(x) - S_u(y)| \leq C_\varepsilon e^{\varepsilon|u|} \|DS_u(x)\|\,|x-y|. \tag{I.10.144}$$

Consequently,

$$C_\varepsilon^{-1} e^{-\varepsilon|u|} \|DS_u(x)\| \operatorname{diam} K \leq \operatorname{diam} K_u \leq C_\varepsilon e^{\varepsilon|u|} \|DS_u(x)\| \operatorname{diam} K. \tag{I.10.145}$$

Proof

Put $a_i(z) = \|DS_i(z)\|$. Compactness, conformality, and invertibility give

$$a_* := \min_{i, z \in U_0} a_i(z) > 0. \tag{I.10.146}$$

Uniform differentiability gives, for every $\delta > 0$, a radius $\rho_\delta > 0$ such that

$$|S_i(w) - S_i(z) - DS_i(z)(w-z)| \leq a_*(1 - e^{-\delta})|z-w| \tag{I.10.147}$$

whenever $z, w \in U_0$, $|z-w| \leq \rho_\delta$, and $i \in \Lambda$. Conformality and (I.10.147) imply the following bounds, using $2 - e^{-\delta} \leq e^\delta$ for the upper estimate:

$$e^{-\delta} a_i(z)|z-w| \leq |S_i(z) - S_i(w)| \leq e^\delta a_i(z)|z-w|. \tag{I.10.148}$$

Let $0 < \vartheta < 1$ be a common contraction constant. Choose N_δ so that

$$\vartheta^{N_\delta} \operatorname{diam} U_0 < \rho_\delta. \tag{I.10.149}$$

For the finitely many words v with $|v| \le N_\delta$, the quotient

$$\frac{|S_v(z) - S_v(w)|}{\|DS_v(z)\|\,|z-w|} \tag{I.10.150}$$

has a positive lower bound and a finite upper bound on $(U_0 \times U_0) \setminus \{z = w\}$, after extending to the diagonal by uniform differentiability. Positivity away from the diagonal follows from injectivity and compactness.

Now write $u = u_1 \cdots u_k$. In the chain from the inner letters toward the outer letters, all but at most N_δ comparisons occur after a tail has contracted z, w to within ρ_δ. Apply (I.10.148) at those comparisons and absorb the remaining finite block using (I.10.150). Conformality makes the derivative norms multiply exactly, so

$$C_\delta^{-1} e^{-\delta k} \|DS_u(z)\| |z-w| \le |S_u(z) - S_u(w)| \le C_\delta e^{\delta k} \|DS_u(z)\| |z-w|. \tag{I.10.151}$$

This is (I.10.144), after renaming $\delta = \varepsilon$. To compare cylinder diameter with the derivative at an arbitrary $x \in X$, we also need to compare derivatives at different base points. Let

$$\omega(t) = \max_i \sup\{|\log a_i(z) - \log a_i(w)| \,:\, z, w \in U_0,\ |z-w| \le t\}.$$

Uniform continuity gives $\omega(t) \to 0$ as $t \to 0$. The chain rule, conformality and forward invariance of U_0 imply

$$|\log \|DS_u(z)\| - \log \|DS_u(w)\|| \le \sum_{j=0}^{|u|-1} \omega(\vartheta^j \operatorname{diam} U_0) \le C'_\varepsilon + \varepsilon|u|.$$

For the last inequality, choose a fixed index beyond which each summand is at most ε; the finitely many earlier summands have bounded sum. Apply (I.10.144) to two points of K, and replace the derivative at the first point by that at x using this comparison. Using $\varepsilon/2$ in each estimate and then enlarging the constant gives (I.10.145), by taking the supremum of the pairwise distances. □

Lemma — 6.3 (stopping length and pull-back geometry) [F]

There are $A_0, A_1 > 0$ such that $|u| \le A_0 n + A_1$ for $u \in \mathscr{W}_n$. For every $\delta > 0$ there is $C_\delta \ge 1$ such that

$$C_\delta^{-1} e^{-\delta n} 2^{-n} \le \operatorname{diam} K_u \le 2^{-n} \quad (u \in \mathscr{W}_n). \tag{I.10.152}$$

For every $\delta > 0$, there is $C_\delta \ge 1$ such that, whenever $m, n \ge 1$, $u \in \mathscr{W}_n$, and $Q \in \mathscr{D}_{m+n}$ satisfies $S_u^{-1}(Q) \cap K \ne \emptyset$,

$$\operatorname{diam}\big(S_u^{-1}(Q) \cap U_0\big) \le C_\delta e^{\delta n} 2^{-m}. \tag{I.10.153}$$

If Q^* is the cube concentric with Q and of twice its side length, then $S_u^{-1}(Q^*) \cap U_0$ contains a Euclidean ball of radius

$$C_\delta^{-1} e^{-\delta n} 2^{-m}. \tag{I.10.154}$$

Proof

For a nonempty stopping word, the upper contraction estimate gives

$$2^{-n} < \operatorname{diam} K_{u^-} \le \vartheta^{|u|-1} \operatorname{diam} K, \tag{I.10.155}$$

which proves the length bound. The function

$$(i, x, y) \longmapsto \frac{|S_i(x) - S_i(y)|}{|x - y|} \tag{I.10.156}$$

has a positive lower bound $b > 0$: near the diagonal use conformality and uniform differentiability, and away from it use injectivity and compactness. Write $u = vi$, where $v = u^-$. The composition convention is $S_u = S_v \circ S_i$, so a lower Lipschitz bound for S_i cannot be applied to K_v to compare K_u with K_v. Instead choose $z, w \in S_i(K)$ with $|z - w| \ge \operatorname{diam}(S_i(K))/2 \ge b \operatorname{diam}(K)/2$. Lemma 6.2 gives

$$\begin{aligned}
\operatorname{diam} K_u &\ge C_\varepsilon^{-1} e^{-\varepsilon|v|} \|DS_v(z)\|\, b \operatorname{diam} K/2, \\
\operatorname{diam} K_v &\le C_\varepsilon e^{\varepsilon|v|} \|DS_v(z)\| \operatorname{diam} K, \\
\operatorname{diam} K_u &\ge \frac{b}{2C_\varepsilon^2} e^{-2\varepsilon|v|} \operatorname{diam} K_v > \frac{b}{2C_\varepsilon^2} e^{-2\varepsilon|v|} 2^{-n}.
\end{aligned} \tag{I.10.157}$$

Take ε small relative to $\delta/(A_0 + 1)$ and use the length bound to prove (I.10.152). For the finitely many coarse generations with the empty stopping word, the length bound is immediate and (I.10.152) follows by increasing C_δ, since $\operatorname{diam} K > 0$ is fixed. Only this subpower lower bound, not a uniform constant-ratio bound, is needed below.

Fix $x \in S_u^{-1}(Q) \cap K$. Equations (I.10.145), (I.10.152), and the length bound show, after choosing the ε in Lemma 6.2 small relative to δ/A_0, that

$$C_\delta^{-1} e^{-\delta n} 2^{-n} \le \|DS_u(x)\| \le C_\delta e^{\delta n} 2^{-n}. \tag{I.10.158}$$

If $y \in S_u^{-1}(Q) \cap U_0$, then

$$|S_u(x) - S_u(y)| \le \sqrt{d}\, 2^{-(m+n)}. \tag{I.10.159}$$

Use the lower bound in (I.10.144) and (I.10.158), once more absorbing $|u| \le A_0 n + A_1$ into δn. This gives (I.10.153).

Every point of Q lies at distance at least $2^{-(m+n+1)}$ from the complement of Q^*. If

$$|y - x| \le C_\delta^{-1} e^{-\delta n} 2^{-m}, \tag{I.10.160}$$

then the upper bounds in (I.10.144) and (I.10.158), with the constant chosen small enough, give

$$|S_u(y) - S_u(x)| < 2^{-(m+n+1)}. \tag{I.10.161}$$

After also decreasing the constant so that the ball remains in U_0, (I.10.161) proves (I.10.154). ☐

Lemma — 6.4 (finite-incidence moment comparison) [F]

Let ρ be a compactly supported probability and $\mathscr{E} = \{E_j\}$ a finite Borel cover of its support. Suppose every E_j meets at most N cubes in $\mathscr{D}_m$, and every cube in $\mathscr{D}_m$ meets at most N members of $\mathscr{E}$. Then, for every $q > 0$,

$$N^{-(q+1)} M_m(\rho, q) \leq \sum_j \rho(E_j)^q \leq N^{q+1} M_m(\rho, q). \tag{I.10.162}$$

Proof

For each j,

$$\rho(E_j) \leq \sum_{Q \in \mathscr{D}_m : Q \cap E_j \neq \varnothing} \rho(Q). \tag{I.10.163}$$

Apply (I.10.122) to (I.10.163), sum in j, and use the second incidence bound. The rough factor N^{q+1} dominates the factors for both $q \geq 1$ and $0 < q < 1$, proving the upper inequality. Since the E_j cover the support,

$$\rho(Q) \leq \sum_{j : E_j \cap Q \neq \varnothing} \rho(E_j). \tag{I.10.164}$$

The same argument with cubes and sets interchanged gives $M_m(\rho, q) \leq N^{q+1} \sum_j \rho(E_j)^q$. ☐

Lemma — 6.5 (L-I.10.08: uniform pull-back moment comparison) [F]

For every $q > 0$ there is a positive sequence $a_n(q)$ satisfying

$$\lim_{n \to \infty} \frac{\log a_n(q)}{n} = 0 \tag{I.10.165}$$

such that, for every compactly supported probability ρ on K, every $m, n \geq 1$, and every $u \in \mathscr{W}_n$,

$$a_n(q)^{-1} M_m(\rho, q) \leq \sum_{Q \in \mathscr{D}_{m+n}} \rho(S_u^{-1} Q)^q \leq a_n(q) M_m(\rho, q). \tag{I.10.166}$$

Here $\rho(S_u^{-1} Q)$ means $\rho(S_u^{-1}(Q) \cap K)$.

Proof

Fix $\delta > 0$ and consider

$$\mathscr{E}_u = \{S_u^{-1}(Q) \cap K \ : \ Q \in \mathscr{D}_{m+n},\ S_u^{-1}(Q) \cap K \neq \varnothing\}. \tag{I.10.167}$$

It covers K. By (I.10.153), each member meets at most

$$Ce^{d\delta n} \tag{I.10.168}$$

cubes of $\mathscr{D}_m$. Conversely, fix a generation-m cube P. For every member of $\mathscr{E}_u$ meeting P, choose its center in that intersection. Equation (I.10.154) assigns a ball of radius $ce^{-\delta n}2^{-m}$ at this center. The doubled image cubes Q^* have uniformly bounded overlap; injectivity of S_u therefore makes the assigned balls uniformly bounded-overlap. A volume comparison inside a fixed enlargement of a generation-m cube shows that such a cube meets at most

$$Ce^{d\delta n} \tag{I.10.169}$$

members of $\mathscr{E}_u$.

Lemma 6.4 now gives (I.10.166) with a bound $C_\delta e^{C(q+1)\delta n}$, with $C_\delta \geq 1$. Choose $\delta_k \downarrow 0$ and strictly increasing integers N_k so large that $\log C_{\delta_k}/N_k \leq \delta_k$. On $N_k \leq n < N_{k+1}$, set $a_n(q) = C_{\delta_k}e^{C(q+1)\delta_k n}$; use any one of the valid bounds for the finitely many earlier n. Then $0 \leq n^{-1}\log a_n(q) \leq (1 + C(q+1))\delta_k$ on the k-th block. This proves (I.10.165) with the required uniformity in m, u, ρ. □

Proof concordance. Lemmas 6.2–6.5 expand FENG07-MF, Lemmas 3.4–3.7, Proposition 3.8, Corollary 3.9, and Proposition 3.3, local PDF pp. 11–15. No bounded-distortion hypothesis is inserted.

I.10.7 Quasi-product moments and auxiliary measures

Let m be a probability on Σ, let $\mu = \pi_* m$, and retain the C^1 conformal IFS of Section 6. No separation condition is assumed.

Lemma — 7.1 (quasi-Bernoulli tail decomposition) [F]

If m is upper quasi-Bernoulli, then for every word u and every Borel set $A \subset \Sigma$,

$$m([u] \cap \sigma^{-|u|}A) \leq C_+ m([u])m(A). \tag{I.10.170}$$

The reverse inequality with C_- holds for a lower quasi-Bernoulli measure. Consequently, if $\mathscr{G}$ is any finite antichain partition of Σ, then for every Borel $E \subset \mathbb{R}^d$,

$$\mu(E) \leq C_+ \sum_{u\in\mathscr{G}} m([u])\mu(S_u^{-1}E) \tag{I.10.171}$$

in the upper case, and

$$\mu(E) \geq C_- \sum_{u\in\mathscr{G}} m([u])\mu(S_u^{-1}E) \tag{I.10.172}$$

in the lower case.

Proof

For a tail cylinder $A = [v]$, (I.10.170) is exactly (I.10.116), since

$$[u] \cap \sigma^{-|u|}[v] = [uv]. \tag{I.10.173}$$

Additivity first extends the inequality from cylinders to their finite disjoint unions, hence to the cylinder algebra. The collection of sets A for which (I.10.170) holds is closed under increasing unions and decreasing intersections. The monotone-class theorem therefore extends the inequality from the algebra to all Borel sets. The lower case is identical.

For $\omega \in [u]$,

$$\pi(\omega) = S_u(\pi(\sigma^{|u|}\omega)). \tag{I.10.174}$$

Decompose $\pi^{-1}E$ over the disjoint cylinders $[u]$, $u \in \mathscr{G}$, apply (I.10.170) to $A = \pi^{-1}(S_u^{-1}E)$, and sum. This proves (I.10.171); (I.10.172) follows in the same way. □

Lemma — 7.2 (L-I.10.09: dyadic quasi-product inequalities) [F]

Fix $q \geq 1$. If m is upper quasi-Bernoulli, there is a positive subexponential sequence $c_n^+(q)$ such that, for $m_0, n \geq 1$ and $\widetilde{Q} \in \mathscr{D}_n$,

$$\sum_{\substack{Q \in \mathscr{D}_{m_0+n} \\ Q \subset \widetilde{Q}}} \mu(Q)^q \leq c_n^+(q) M_{m_0}(\mu, q) \sum_{\substack{B \in \mathscr{D}_n \\ B \sim \widetilde{Q}}} \mu(B)^q. \tag{I.10.175}$$

Fix $0 < q < 1$. If m is lower quasi-Bernoulli, there is a positive subexponential sequence $c_n^-(q)$ such that

$$\sum_{\substack{B \in \mathscr{D}_n \\ B \sim \widetilde{Q}}} \sum_{\substack{Q \in \mathscr{D}_{m_0+n} \\ Q \subset B}} \mu(Q)^q \geq c_n^-(q) M_{m_0}(\mu, q) \mu(\widetilde{Q})^q. \tag{I.10.176}$$

Proof

Let

$$\mathscr{G}(\widetilde{Q}) = \{u \in \mathscr{W}_n \;:\; K_u \cap \widetilde{Q} \neq \varnothing\}, \qquad W = \sum_{u \in \mathscr{G}(\widetilde{Q})} m([u]). \tag{I.10.177}$$

If $Q \subset \widetilde{Q}$, only the words in $\mathscr{G}(\widetilde{Q})$ contribute to (I.10.171). For $W > 0$, convexity gives

$$\begin{aligned} \mu(Q)^q &\leq C_+^q \left(\sum_{u \in \mathscr{G}(\widetilde{Q})} m([u]) \mu(S_u^{-1}Q) \right)^q \\ &\leq C_+^q W^q \sum_{u \in \mathscr{G}(\widetilde{Q})} \frac{m([u])}{W} \mu(S_u^{-1}Q)^q. \end{aligned} \tag{I.10.178}$$

Every K_u, $u \in \mathscr{G}(\widetilde{Q})$, has diameter at most 2^{-n} and meets $\widetilde{Q}$, so

$$K_u \subset U(\widetilde{Q}) := \bigcup_{\substack{B \in \mathscr{D}_n \\ B \sim \widetilde{Q}}} B. \tag{I.10.179}$$

The union of the disjoint symbolic cylinders $[u]$ is therefore contained in $\pi^{-1}U(\widetilde{Q})$. Hence

$$W \le \mu(U(\widetilde{Q})) \le \sum_{B\sim\widetilde{Q}} \mu(B). \tag{I.10.180}$$

There are at most 3^d neighbors. Sum (I.10.178) over $Q \subset \widetilde{Q}$, use (I.10.180), (I.10.21), and the upper half of Lemma 6.5:

$$\begin{aligned}\sum_{Q\subset\widetilde{Q}} \mu(Q)^q &\le C\left(\sum_{B\sim\widetilde{Q}} \mu(B)\right)^q \max_{u\in\mathscr{W}_n} \sum_{Q\in\mathscr{D}_{m_0+n}} \mu(S_u^{-1}Q)^q \\ &\le C a_n(q) M_{m_0}(\mu,q) \sum_{B\sim\widetilde{Q}} \mu(B)^q.\end{aligned} \tag{I.10.181}$$

This is (I.10.175), after absorbing the constant into a subexponential sequence. The case $W = 0$ is trivial.

Now let $0 < q < 1$ and use (I.10.172). Concavity gives

$$\mu(Q)^q \ge C_-^q W^q \sum_{u\in\mathscr{G}(\widetilde{Q})} \frac{m([u])}{W} \mu(S_u^{-1}Q)^q. \tag{I.10.182}$$

Every sequence coding a point of $\widetilde{Q}$ has its $\mathscr{W}_n$-prefix in $\mathscr{G}(\widetilde{Q})$, and therefore

$$\mu(\widetilde{Q}) \le W. \tag{I.10.183}$$

For a fixed $u \in \mathscr{G}(\widetilde{Q})$, (I.10.179) implies that $\mu(S_u^{-1}Q) = 0$ for every fine cube Q outside the neighboring union. Sum (I.10.182) over all fine cubes in that union, use (I.10.183), and take the minimum in u:

$$\begin{aligned}&\sum_{B\sim\widetilde{Q}} \sum_{\substack{Q\in\mathscr{D}_{m_0+n}\\ Q\subset B}} \mu(Q)^q \\ &\quad\ge C_-^q \mu(\widetilde{Q})^q \min_{u\in\mathscr{W}_n} \sum_{Q\in\mathscr{D}_{m_0+n}} \mu(S_u^{-1}Q)^q \\ &\quad\ge C_-^q a_n(q)^{-1} M_{m_0}(\mu,q)\mu(\widetilde{Q})^q.\end{aligned} \tag{I.10.184}$$

This proves (I.10.176) with a subexponential $c_n^-(q)$. □

Lemma — 7.3 (perturbed Fekete lemma) [F]

Let $A_n > 0$ and let $c_n > 0$ satisfy $\log c_n/n \to 0$.

(1) If

$$A_{m+n} \le c_n A_m A_n \quad (m,n \ge 1), \tag{I.10.185}$$

then $\lim_n n^{-1} \log A_n$ exists in $[-\infty, \infty)$.

(2) If the inequality in (I.10.185) is reversed, then the limit exists in $(-\infty, \infty]$.

Proof

We prove the first assertion. Fix ℓ, write $n = k\ell + r$, $0 \le r < \ell$, and iterate (I.10.185) with the second block of length ℓ. If $r \ge 1$, this starts from A_r; if $r = 0$, it starts from A_ℓ and uses only $k - 1$ iterations. Thus a constant $C_\ell > 0$, independent of k, satisfies

$$A_{k\ell+r} \le (c_\ell A_\ell)^k C_\ell. \tag{I.10.186}$$

Thus

$$\limsup_{n\to\infty} \frac{\log A_n}{n} \le \frac{\log c_\ell + \log A_\ell}{\ell}. \tag{I.10.187}$$

Let $\ell \to \infty$. The right side has liminf equal to $\liminf_\ell \ell^{-1} \log A_\ell$, proving existence. Reversing every inequality proves the second assertion. □

Lemma — 7.4 (tilting sequence) [F]

Suppose $A_n > 0$ and $n^{-1} \log A_n \to a \in \mathbb{R}$. For every $\varepsilon > 0$, there is a decreasing sequence $t_n > 0$ such that

$$t_1 = \varepsilon, \qquad t_n \downarrow 0, \tag{I.10.188}$$

and, with

$$b_n = \exp(t_1 + \cdots + t_n), \tag{I.10.189}$$

one has

$$b_{m+n} \le b_m b_n, \qquad b_n \le e^{\varepsilon n}, \qquad \sum_{n=1}^{\infty} A_n b_n e^{-an} = \infty. \tag{I.10.190}$$

Proof

Set $t_1 = \varepsilon$. Having chosen $t_1, \dots, t_n$, set

$$t_{n+1} = \begin{cases} t_n/2, & A_n b_n e^{-an} \ge 1, \\ t_n, & A_n b_n e^{-an} < 1. \end{cases} \tag{I.10.191}$$

The sequence is decreasing. If it had a positive limit t, then $\log(A_n e^{-an}) = o(n)$ and $\log b_n \ge tn$ for all large n, so $A_n b_n e^{-an} \ge 1$ eventually. Formula (I.10.191) would then halve t_n repeatedly, a contradiction. Thus $t_n \downarrow 0$.

The halving alternative must occur infinitely often; at each preceding index the summand in (I.10.190) is at least one. Hence the series diverges. Since t_n is decreasing,

$$t_{m+1} + \cdots + t_{m+n} \le t_1 + \cdots + t_n, \tag{I.10.192}$$

which proves $b_{m+n} \le b_m b_n$. The bound $b_n \le e^{\varepsilon n}$ is immediate. □

Lemma — 7.5 (L-I.10.10: moment limits and auxiliary q-measures) [F]

Under the upper quasi-Bernoulli hypothesis, for every $q \geq 1$ the dyadic spectrum exists finitely and there are a probability ν_q and a subpower function h_q such that

$$\nu_q(B(x, r/(16\sqrt{d}))) \leq h_q(r) r^{-\tau_\mu^{\mathrm{dyad}}(q)} \mu(B(x,r))^q \tag{I.10.193}$$

for every x and $0 < r < 1$. Under the lower quasi-Bernoulli hypothesis, for every $0 < q < 1$ the dyadic spectrum exists finitely and there are ν_q, h_q such that

$$\nu_q(B(x, 16\sqrt{d}\, r)) \geq h_q(r) r^{-\tau_\mu^{\mathrm{dyad}}(q)} \mu(B(x,r))^q \tag{I.10.194}$$

for every x and $0 < r < 1/4$.

Proof: existence of the limits

Write $A_n = M_n(\mu, q)$. Sum (I.10.175) over $\widetilde{Q} \in \mathscr{D}_n$. Every $B \in \mathscr{D}_n$ is a neighbor of at most 3^d cubes, so

$$A_{m+n} \leq 3^d c_n^+(q) A_m A_n \quad (q \geq 1). \tag{I.10.195}$$

For $0 < q < 1$, sum (I.10.176). Its left side counts every fine cube at most 3^d times, and hence

$$A_{m+n} \geq 3^{-d} c_n^-(q) A_m A_n. \tag{I.10.196}$$

Lemma 7.3 gives the limit of $n^{-1} \log A_n$. To see that it is finite, place the compact support in a fixed ball. At most $C2^{dn}$ dyadic cubes have positive mass, and their masses sum to one. The finite-vector inequality (I.10.122), in both directions, gives

$$C_q^{-1} 2^{-dqn} \leq A_n \leq C_q 2^{dn}. \tag{I.10.197}$$

Thus

$$\tau = \tau_\mu^{\mathrm{dyad}}(q) = -\frac{1}{\log 2} \lim_{n\to\infty} \frac{\log A_n}{n} \in \mathbb{R}. \tag{I.10.198}$$

Proof: the upper auxiliary measure for $q \geq 1$

Apply Lemma 7.4 with $a = -\tau \log 2$, and define, for $s < \tau$,

$$Z_\varepsilon(s) = \sum_{k=1}^{\infty} b_k 2^{ks} A_k. \tag{I.10.199}$$

Because $b_k = e^{o(k)}$ and $A_k = 2^{-k\tau + o(k)}$, the series is finite for $s < \tau$. Lemma 7.4 gives

$$Z_\varepsilon(s) \longrightarrow \infty \quad \text{as } s \uparrow \tau. \tag{I.10.200}$$

Let $Q_k(x)$ be the generation-k half-open dyadic cube containing x, and define

$$\varphi_{\varepsilon,s}(x) = \sum_{k=1}^{\infty} b_k 2^{k(s+d)} \mu(Q_k(x))^q. \tag{I.10.201}$$

Termwise integration gives

$$\int_{\mathbb{R}^d} \varphi_{\varepsilon,s}(x)\, dx = Z_\varepsilon(s). \tag{I.10.202}$$

Hence

$$\nu_{\varepsilon,s}(E) = Z_\varepsilon(s)^{-1} \int_E \varphi_{\varepsilon,s}(x)\, dx \tag{I.10.203}$$

is a probability. All these probabilities are supported in one fixed compact neighborhood of $\operatorname{supp}\mu$.

Fix $\widetilde{Q} \in \mathscr{D}_n$. Split the series in $\nu_{\varepsilon,s}(\widetilde{Q})$ at $k = n$. The contribution of $k \le n$, divided by $Z_\varepsilon(s)$, tends to zero as $s \uparrow \tau$. For $k = n+m$, integration over the fine cubes in $\widetilde{Q}$, followed by (I.10.175), gives

$$\begin{aligned} &Z_\varepsilon(s)^{-1} \sum_{m=1}^{\infty} b_{n+m} 2^{(n+m)s} \sum_{\substack{Q \in \mathscr{D}_{n+m} \\ Q \subset \widetilde{Q}}} \mu(Q)^q \\ &\le b_n c_n^+(q) 2^{ns} \sum_{B \sim \widetilde{Q}} \mu(B)^q. \end{aligned} \tag{I.10.204}$$

Here $b_{n+m} \le b_n b_m$, and the remaining sum is $Z_\varepsilon(s)$.

Here is the compactness step in full. Let L be a compact set containing the supports of all $\nu_{\varepsilon,s}$. Choose a countable uniformly dense family $(f_j)_{j\ge1}$ in $C(L)$. Successive subsequence extraction makes

$$\lim_{k\to\infty} \int f_j \, d\nu_{\varepsilon,s_k} \tag{I.10.204a}$$

exist for every j, along some $s_k \uparrow \tau$. The resulting functional on the span of the f_j is bounded by the uniform norm, positive, and takes the value one on the constant function. Uniform-density extension and the A-RA Riesz representation theorem therefore give a Borel probability ν_ε on L, and the same density argument gives

$$\int f\, d\nu_{\varepsilon,s_k} \longrightarrow \int f\, d\nu_\varepsilon \quad (f \in C(L)). \tag{I.10.204b}$$

For later use, we also prove the needed open/closed inequalities. If $G \subset L$ is a proper open set, the continuous functions

$$g_j(x) = \min\{1, j\operatorname{dist}(x, L\setminus G)\} \tag{I.10.204c}$$

increase to $\mathbf{1}_G$. Equation (I.10.204b) and monotone convergence give

$$\nu_\varepsilon(G) \le \liminf_k \nu_{\varepsilon,s_k}(G). \tag{I.10.204d}$$

Taking complements gives $\limsup_k \nu_{\varepsilon,s_k}(F) \le \nu_\varepsilon(F)$ for closed $F \subset L$. Applying the open inequality to interiors in (I.10.204) gives

$$\nu_\varepsilon(\operatorname{int}\widetilde{Q}) \le b_n c_n^+(q) 2^{n\tau} \sum_{B\sim\widetilde{Q}} \mu(B)^q. \tag{I.10.205}$$

Take a weak limit as $\varepsilon \downarrow 0$. Since $b_n \le e^{\varepsilon n}$, the resulting probability ν_q satisfies

$$\nu_q(\operatorname{int}\widetilde{Q}) \le c_n^+(q) 2^{n\tau} \sum_{B\sim\widetilde{Q}} \mu(B)^q. \tag{I.10.206}$$

One must not sum bounds for separate cube interiors to control their internal faces: a weak limit may charge those faces. Instead, for any finite family $\mathscr{C} \subset \mathscr{D}_n$, set $G = \operatorname{int}(\bigcup_{Q\in\mathscr{C}} Q)$. Before either weak limit, $\nu_{\varepsilon,s}(G) \le \sum_{Q\in\mathscr{C}} \nu_{\varepsilon,s}(Q)$. Sum (I.10.204) over this finite family; the finitely many initial-series contributions still tend to zero. Apply the open-set inequality first as $s \uparrow \tau$ and then as $\varepsilon \downarrow 0$. This proves the stronger statement

$$\nu_q\left(\operatorname{int}\bigcup_{Q\in\mathscr{C}} Q\right) \le c_n^+(q) 2^{n\tau} \sum_{Q\in\mathscr{C}} \sum_{B\sim Q} \mu(B)^q. \tag{I.10.206a}$$

Given $0 < r < 1$, put $\tilde r = r/(12\sqrt{d})$ and choose n so that

$$2\tilde r < 2^{-n} \le 4\tilde r. \tag{I.10.207}$$

The ball $B(x, r/(16\sqrt{d}))$ lies in the interior of the union of the at most 2^d generation-n cubes meeting $B(x,\tilde r)$. Every neighbor of one of those cubes lies in $B(x,r)$, and there are at most 6^d resulting incidences. Equation (I.10.206a), which includes internal grid faces, therefore gives

$$\nu_q(B(x, r/(16\sqrt{d}))) \le C_d c_n^+(q) 2^{n\tau} \mu(B(x,r))^q. \tag{I.10.208}$$

Since $2^{-n} \asymp_d r$, define

$$h_q(r) = C_{d,q,\tau} \sup\{c_k^+(q) \,:\, |k - \log_2(1/r)| \le C_d\}. \tag{I.10.209}$$

The sequence is subexponential, so h_q is subpower. Equations (I.10.208)–(I.10.209) prove (I.10.193).

Proof: the lower auxiliary measure for $0 < q < 1$

We repeat the construction, not merely its conclusion. Apply Lemma 7.4 together with (I.10.199)–(I.10.203). For $\widetilde{Q} \in \mathscr{D}_n$, sum $\nu_{\varepsilon,s}(B)$ over $B \sim \widetilde{Q}$ and retain only the terms $k = n+m$. Because $b_{n+m} \ge b_m$, (I.10.176) gives

$$\begin{aligned}\sum_{B\sim\widetilde{Q}} \nu_{\varepsilon,s}(B) &\ge Z_\varepsilon(s)^{-1} \sum_{m=1}^{\infty} b_m 2^{(n+m)s} c_n^-(q) A_m \mu(\widetilde{Q})^q \\ &= c_n^-(q) 2^{ns} \mu(\widetilde{Q})^q.\end{aligned} \tag{I.10.210}$$

Take a weak limit as $s \uparrow \tau$. Portmanteau for the finite union of closed neighboring cubes gives

$$\sum_{B \sim \widetilde{Q}} \nu_q(\overline{B}) \geq c_n^-(q) 2^{n\tau} \mu(\widetilde{Q})^q. \tag{I.10.211}$$

Now choose n so that $2r < 2^{-n} \leq 4r$. The ball $B(x,r)$ is contained in the union of at most 2^d generation-n cubes. By the finite-vector inequality,

$$\mu(B(x,r))^q \leq 2^{dq} \sum_{\widetilde{Q} \cap B(x,r) \neq \emptyset} \mu(\widetilde{Q})^q. \tag{I.10.212}$$

Every neighbor B of one of these $\widetilde{Q}$ is contained in $B(x, 16\sqrt{d}\, r)$. Each closed cube occurs at most 3^d times, and a point belongs to at most 2^d closed generation-n cubes. Thus the sum of their indicator functions, counting repetitions, is at most 6^d times the indicator of that enlarged ball. Sum (I.10.211) and use (I.10.212):

$$\nu_q(B(x, 16\sqrt{d}\, r)) \geq C_{d,q}^{-1} c_n^-(q) 2^{n\tau} \mu(B(x,r))^q. \tag{I.10.213}$$

With

$$h_q(r) = C_{d,q,\tau}^{-1} \inf\{c_k^-(q) : |k - \log_2(1/r)| \leq C_d\}, \tag{I.10.214}$$

subexponentiality makes h_q subpower. This proves (I.10.194) and completes the independent lower construction. □

Proof concordance. Lemmas 7.1–7.5 expand FENG07-MF, Proposition 3.1, Remark 3.10, Lemmas 4.1–4.2, and Propositions 4.3–4.4, local PDF pp. 10–20. In particular, the proof after (I.10.210) fills the source's omitted lower-q calculation.

I.10.8 Feng's no-separation formalism

Lemma — 8.1 (positive-order grid/packing concordance) [F]

For every compactly supported probability μ and every $q > 0$, there are constants $a, A, C > 0$, independent of n, such that

$$\mathscr{M}_\mu(q, a2^{-n}) \leq C M_n(\mu, q), \qquad M_n(\mu, q) \leq C \mathscr{M}_\mu(q, A2^{-n}). \tag{I.10.215}$$

The same comparisons hold for an arbitrary r after choosing $2^{-n} \asymp r$. Hence, if the dyadic spectrum exists at $q > 0$, then the centered-packing spectrum exists and

$$\tau_\mu^{\mathrm{pack}}(q) = \tau_\mu^{\mathrm{dyad}}(q). \tag{I.10.216}$$

Proof

The first comparison is Lemma 5.3, with dyadic scale comparable to the ball radius. For the other direction, choose $x_Q \in Q \cap \operatorname{supp} \mu$ for every $Q \in \mathscr{D}_n^+(\mu)$. A ball

$$B(x_Q, 2\sqrt{d}\,2^{-n}) \tag{I.10.217}$$

contains Q. Color the dyadic cubes by their coordinate indices modulo an integer $L > 8\sqrt{d} + 4$. Within each of the L^d color classes the balls in (I.10.217) are pairwise disjoint. Since $q > 0$,

$$\sum_{Q \text{ in one color}} \mu(Q)^q \le \mathscr{M}_\mu(q, 2\sqrt{d}\,2^{-n}). \tag{I.10.218}$$

Sum over the colors. This proves the second comparison, with a harmless reciprocal renaming of the fixed scale constant. Repeating the incidence argument whenever $2^{-n-1} < r \le 2^{-n}$ gives the arbitrary-radius comparison. Constant factors disappear after division by $\log r$, which proves (I.10.216). ☐

Lemma — 8.2 (all-real lower packing spectrum for two-sided sources) [F]

Let m be two-sided quasi-Bernoulli on a finite alphabet and $\mu = \pi_* m$ for a finite C^1 conformal IFS. Then $\underline{\tau}_\mu^{\text{pack}}(q)$ is finite for every $q \in \mathbb{R}$ and is concave on $\mathbb{R}$.

Proof

For a fixed disjoint family of r-balls, Hölder's inequality gives

$$\sum_j \mu(B_j)^{(1-\lambda)q_0+\lambda q_1} \le \left(\sum_j \mu(B_j)^{q_0}\right)^{1-\lambda} \left(\sum_j \mu(B_j)^{q_1}\right)^{\lambda}. \tag{I.10.219}$$

Taking suprema preserves the inequality. After logarithms and division by $\log r < 0$, its direction reverses. Taking a liminf gives

$$\underline{\tau}_\mu^{\text{pack}}((1-\lambda)q_0 + \lambda q_1) \ge (1-\lambda)\underline{\tau}_\mu^{\text{pack}}(q_0) + \lambda \underline{\tau}_\mu^{\text{pack}}(q_1). \tag{I.10.220}$$

It remains to prove finiteness. Positive q is bounded by the compact support cube count as in (I.10.197) and Lemma 8.1. At $q = 0$, the maximal number of disjoint r-balls centered on a fixed compact set lies between one and Cr^{-d}, so the lower spectrum is finite. Let $q < 0$. Put

$$p_* = \min_{i \in \Lambda} m([i]) > 0. \tag{I.10.221}$$

Iteration of the lower quasi-Bernoulli inequality gives

$$m([\omega|_N]) \ge C_-^{N-1} \prod_{j=1}^{N} m([\omega_j]) \ge C_-^{N-1} p_*^N. \tag{I.10.222}$$

Choose $N = N(r) \le C\log(1/r) + C$ so that uniform contraction makes $\operatorname{diam} K_{\omega|_N} \le r$. Since $\pi\omega \in K_{\omega|_N}$,

$$\mu(B(\pi\omega, r)) \ge \mu(K_{\omega|_N}) \ge m([\omega|_N]) \ge cr^L \tag{I.10.223}$$

for fixed $c, L > 0$. A disjoint family of r-balls in a fixed bounded set has at most Cr^{-d} members. Because $q < 0$, (I.10.223) yields

$$1 \le \mathscr{M}_\mu(q, r) \le Cr^{-d+Lq}. \tag{I.10.224}$$

Division by $\log r$ bounds the lower spectrum between two finite numbers. Together with (I.10.220), this proves the lemma. □

Theorem — 8.3 (T-I.10.11: positive-q spectrum without separation) [F]

Let Φ be a finite C^1 conformal IFS, with no separation condition. Let m be Bernoulli or, more generally, two-sided quasi-Bernoulli, and put $\mu = \pi_* m$. Then, for every $q > 0$,

$$\tau_\mu^{\mathrm{dyad}}(q) = \lim_{n\to\infty} \frac{\log M_n(\mu, q)}{-n \log 2} \in \mathbb{R} \tag{I.10.225}$$

exists, the centered-packing spectrum exists, and

$$\tau_\mu^{\mathrm{pack}}(q) = \tau_\mu^{\mathrm{dyad}}(q). \tag{I.10.226}$$

Proof

If K is a singleton, every moment sum equals one and both spectra vanish. Assume $\operatorname{diam} K > 0$. For $q \ge 1$, upper quasi-Bernoulli and Lemma 7.5 prove existence and finiteness. For $0 < q < 1$, lower quasi-Bernoulli and the other half of Lemma 7.5 do the same. Lemma 8.1 gives (I.10.226). □

Point-of-statement provenance. Theorem 8.3 is the Bernoulli/quasi-Bernoulli specialization and expansion of FENG07-MF, Corollary 4.5 and Remark 3.10, local PDF pp. 15–20. Lemmas 6.2–7.5 supply the complete local proof.

Theorem — 8.4 (T-I.10.12: differentiable-parameter formalism) [F]

Under the hypotheses of Theorem 8.3, let $t \ge 1$, assume that τ_μ^{pack} is differentiable at t, and put

$$\alpha = (\tau_\mu^{\mathrm{pack}})'(t). \tag{I.10.227}$$

Then

$$\dim_{\mathrm{H}} E_\mu(\alpha) = t\alpha - \tau_\mu^{\mathrm{pack}}(t) = \inf_{q\in\mathbb{R}} \{q\alpha - \underline{\tau}_\mu^{\mathrm{pack}}(q)\}. \tag{I.10.228}$$

For a fixed $t \ge 1$, the same conclusion holds for an upper quasi-Bernoulli source if the lower packing spectrum is finite on $\mathbb{R}$, the dyadic spectrum exists and agrees with it on an open interval containing t, and the upper quasi-product construction applies on that interval.

Proof

Lemma 8.2 supplies the finite concave all-real lower packing spectrum. Theorems 8.3 and Lemma 8.1 identify it with the dyadic spectrum throughout $(0, \infty)$. Since $t \geq 1$, the upper half of Lemma 7.5 supplies the upper t-Gibbs ball inequality (I.10.193). All hypotheses of Lemma 5.4 therefore hold at t, and (I.10.228) is exactly (I.10.126). This includes $t = 1$: neighboring moment limits exist on both sides of one, but the lemma requires an auxiliary upper-Gibbs measure only at t, not below it.

Under the hypotheses in the last paragraph of the statement, those same three inputs are assumed or follow from the upper half of Lemma 7.5, so the same application of Lemma 5.4 proves the extension. □

Point-of-statement provenance. Theorem 8.4 is FENG07-MF, Theorem 1.1 and Theorem 4.6, local PDF pp. 4 and 20, with every invoked source lemma proved in Sections 5–7. The factor in the upper auxiliary ball is $r/(16\sqrt{d})$, as in FENG07-MF (6), not a prose approximation.

I.10.9 Exact boundary of the chapter

Research boundary — Question/Boundary 9.1 (Q-I.10.13) [Q]

The preceding theorems intentionally make no stronger claim in the following regimes.
(1) For a general overlapping self-conformal measure, Theorem 8.4 does not extend the equality to $0 < t < 1$. The **asymptotically weak separation condition** means that there are integers $t_n \geq 1$ with $\log t_n / n \to 0$ such that, for every $Q \in \mathscr{D}_n$,

$$\#\{S_u \,:\, u \in \mathscr{W}_n,\ K_u \cap Q \neq \emptyset\} \leq t_n.$$

Here distinct maps, not words, are counted; maps are equal when they agree on the common IFS domain. This is FENG07-MF, Definition 5.1, local PDF p. 21. For a Bernoulli self-conformal measure whose IFS satisfies this condition, its Theorem 5.7 (local PDF p. 25) proves the Legendre equality at every differentiability parameter $t > 0$. That additional theorem is referenced, not proved or imported as a premise here [Fen07, Definition 5.1 and Theorem 5.7].
(2) An all-real centered-packing formula does not imply a fixed-dyadic negative-q formula. Tiny positive pieces cut by grid boundaries can dominate the latter.
(3) At a phase transition, the interval between the right and left derivatives consists of supporting slopes, but differentiability-based construction of one exact level measure is unavailable.
(4) Theorem 4.4 does not assert equality at either extremal Gibbs slope.
(5) Testud gives self-similar examples with a nonconcave **grid-defined** multifractal spectrum. To specify this convention, fix an integer $\ell \geq 2$, let $I_n(x)$ be the unique half-open interval $[k\ell^{-n}, (k+1)\ell^{-n})$ containing $x \in [0, 1)$, and set

$$E_\mu^{(\ell)}(\alpha) = \left\{ x \,:\, \lim_{n\to\infty} \frac{\log \mu(I_n(x))}{-n \log \ell} = \alpha \right\}.$$

The cited spectrum is $\alpha \mapsto \dim_{\mathrm{H}} E_\mu^{(\ell)}(\alpha)$, with empty levels assigned $-\infty$. Precise examples are in Testud [Tes06, Sections 6.2–6.3, local PDF pp. 15–17]. The source uses the opposite sign for its moment exponent. No equivalence with our ball-defined levels, especially at negative orders, is asserted without an additional proof. This cited boundary result has no outgoing premise edge.
These are scope statements and research warnings, not proof gaps in any F node.

I.10.10 Intuition and strategy

The following discussion explains the geometric ideas behind the proofs.

I.10.10.1 What the parameter q sees

At a fixed scale, the moment sum

$$\sum_Q \mu(Q)^q$$

changes its preference as q changes. Large positive q emphasizes heavy cubes. Orders between zero and one reward the large number of small positive masses. Negative q magnifies the smallest positive masses so aggressively that the way a fixed grid cuts the support can change the answer. This is why the all-real theory uses balls centered on the support, while the fixed dyadic grid is retained only in the stable nonnegative range.

The sign convention can initially look reversed. Since the scale r tends to zero, $\log r < 0$. A uniform s-dimensional measure has roughly r^{-s} occupied cells of mass r^s, so

$$\sum_Q \mu(Q)^q \approx r^{s(q-1)}.$$

Thus $\tau(q) = s(q-1)$, $\tau(0) = -s$, and $\tau(q)/(q-1) = s$.

I.10.10.2 Why the Legendre transform appears

Suppose a collection of scale-r cells has mass about r^α and cardinality about $r^{-f(\alpha)}$. Its contribution to the q-moment is heuristically

$$r^{-f(\alpha)} r^{q\alpha} = r^{q\alpha - f(\alpha)}.$$

The dominant exponent suggests

$$\tau(q) = \inf_\alpha \{q\alpha - f(\alpha)\},$$

and hence

$$f(\alpha) \le \inf_q \{q\alpha - \tau(q)\}.$$

The upper inequality is universal and comes from a cover. Equality is much harder: one must construct a measure carried by the chosen local-dimension level set and prove that this measure has the predicted dimension.

I.10.10.3 The separated Bernoulli model

For a self-similar Bernoulli measure, a word u has mass p_u and diameter comparable to r_u. At a prescribed geometric scale, the stopping words do not all have the same symbolic length, but their contraction ratios differ only by one final-letter factor. The equation

$$\sum_i p_i^q r_i^{-\tau(q)} = 1$$

creates a new probability vector $\theta_i(q) = p_i^q r_i^{-\tau(q)}$. Summing its word probabilities over a stopping antichain gives exactly one; after removing the geometric factor, the q-moment is therefore $r^{\tau(q)}$.

The endpoint level sets need not be singletons. If several symbols have the same smallest value of $\log p_i / \log r_i$, every sequence using only those symbols lies at the same extremal local dimension. Their sub-IFS can have positive Hausdorff dimension. The endpoint proof works on this whole symbol face.

I.10.10.4 Tilting a Gibbs measure

For a Hölder Gibbs measure, the cylinder mass is controlled by a potential ψ_0, while the cylinder diameter is controlled by the negative geometric potential ϕ. The tilted potential

$$q\psi_0 - T(q)\phi$$

is normalized to pressure zero. Its equilibrium state favors precisely the cylinders that dominate the order-q moment. Under that tilted state, the ratio of Birkhoff averages is

$$\frac{\int \psi_0 \, dm_q}{\int \phi \, dm_q} = T'(q),$$

which becomes the local dimension of the original Gibbs measure. The pressure-zero identity then turns the dimension of the tilted state into $qT'(q) - T(q)$.

No analytic spectral perturbation is conceptually necessary here. Uniqueness of equilibrium states lets nearby tilted measures converge to the current one, and two elementary variational comparisons squeeze the difference quotient of T.

I.10.10.5 What overlaps destroy

With overlaps, a geometric cube does not correspond to one symbolic cylinder. Many words can feed mass into the same region, and a geometric restriction of the measure is a mixture of pulled-back copies. Exact bounded distortion is also unavailable for a merely C^1 system.

Two weaker facts survive:

1. continuity of the derivative makes the distortion accumulated over $O(n)$ steps equal to $e^{o(n)}$;
2. a quasi-Bernoulli symbolic source makes a prefix and its tail independent up to a constant.

Together these yield moment products up to $e^{o(n)}$. Such errors vanish after division by the scale exponent n, so the positive-order spectrum still has a limit.

I.10.10.6 The auxiliary measure

An exact Gibbs measure at order q would assign a scale-n cube a mass comparable to

$$2^{n\tau(q)}\mu(Q)^q.$$

Overlaps prevent these desired masses from being exactly consistent between scales. Feng's construction averages the desired densities over all scales with a slowly increasing weight. A carefully chosen subexponential tilt makes the normalizing series diverge exactly at the spectrum exponent. Weak limits then satisfy a ball version of the Gibbs estimate, with only a subpower error.

For $q \geq 1$, convexity produces an upper auxiliary estimate. For $0 < q < 1$, concavity produces a lower estimate. Only the upper estimate concentrates the auxiliary measure on a differentiable local-dimension level set; this is the structural reason the no-separation theorem begins at $t = 1$.

I.10.10.7 Reading the failure boundary

A differentiability point selects one slope and one tilted level measure. At a phase transition there is a whole interval of supporting slopes, and this selection mechanism no longer supplies all corresponding level sets. Negative orders create an additional problem: very small masses are unstable under fixed-grid cuts. These are distinct obstructions, and neither should be hidden by writing a Legendre formula without its parameter range and normalization.

I.10.11 Exercises

These exercises use the preceding chapters and the results of this chapter. References such as "Theorem 3.3" refer to Chapter I.10.

I.10.11.1 Exercise 1 — Rényi monotonicity from convexity

Let $a_1, \dots, a_N > 0$, $\sum_i a_i = 1$. Starting only from convexity of

$$K(t) = \log \sum_i a_i^{1+t},$$

prove that $H_q(a) = (1-q)^{-1} \log \sum_i a_i^q$, initially defined for $q \neq 1$, has a continuous extension to $q = 1$, compute its value there, and prove that the extended function is nonincreasing on $\mathbb{R}$.

I.10.11.2 Exercise 2 — The universal upper bound at two signs

Assume that $E_\mu(\alpha) \neq \emptyset$ and that $\underline{\tau}_\mu^{\mathrm{pack}}$ is finite at $q_+ > 0$ and $q_- < 0$. Write the two separate mass inequalities used to prove

$$\dim_{\mathrm{H}} E_\mu(\alpha) \le \min\{q_+\alpha - \underline{\tau}_\mu^{\mathrm{pack}}(q_+), q_-\alpha - \underline{\tau}_\mu^{\mathrm{pack}}(q_-)\}.$$

Explain exactly where the inequality reverses in the negative-order case.

I.10.11.3 Exercise 3 — Equal-ratio Bernoulli spectrum

Suppose $r_i = r \in (0,1)$ for $1 \le i \le N$, the IFS has SSC, and $p_i > 0$. Compute $\tau_\mu^{\mathrm{pack}}(q)$, its derivative, and $D_q^{\mathrm{dyad}}(\mu)$ for $q > 0$, $q \neq 1$. Compute D_1 by taking the limit.

I.10.11.4 Exercise 4 — A nontrivial endpoint face

Consider an SSC self-similar measure with three maps and data

$$(r_1, r_2, r_3) = (1/4, 1/4, 1/5), \qquad (p_1, p_2, p_3) = (1/8, 1/8, 3/4).$$

Show that symbols $1, 2$ share one local-dimension ratio. Determine whether that ratio is $\alpha_{\min}$ or $\alpha_{\max}$, and compute the Hausdorff dimension supplied by the corresponding two-symbol endpoint face. You need not simplify logarithmic quotients involving symbol 3 beyond the comparison required.

I.10.11.5 Exercise 5 — When is the pressure spectrum affine?

For the separated Bernoulli pressure root, derive

$$\tau''(q) = \frac{\sum_i \theta_i(q)(\log p_i - \tau'(q)\log r_i)^2}{\sum_i \theta_i(q)\log r_i}.$$

Use it to prove that τ is affine on a nonempty open interval if and only if $p_i = r_i^s$ for every i, where $\sum_i r_i^s = 1$.

I.10.11.6 Exercise 6 — Difference quotients of the Gibbs pressure root

Let $T(q)$ be defined by $P(q\psi_0 - T(q)\phi) = 0$, and let m_q be the tilted equilibrium state. For $h > 0$, prove directly from the two variational inequalities that

$$\frac{\int \psi_0 \, dm_{q+h}}{\int \phi \, dm_{q+h}} \le \frac{T(q+h) - T(q)}{h} \le \frac{\int \psi_0 \, dm_q}{\int \phi \, dm_q}.$$

Identify the sign that forces both reversals.

I.10.11.7 Exercise 7 — Dimension of the tilted state

Assume the hypotheses of Theorem 4.4 and $\alpha = T'(q)$. Starting from pressure zero and the conformal entropy/Lyapunov dimension formula, derive

$$\dim_{\mathrm{H}}(\pi_* m_q) = q\alpha - T(q).$$

I.10.11.8 Exercise 8 — Why negative dyadic moments are not merged with packings

Let a positive mass a be split by a grid boundary into εa and $(1-\varepsilon)a$. For $q<0$, compute the ratio

$$\frac{(\varepsilon a)^q + ((1-\varepsilon)a)^q}{a^q}$$

and its limit as $\varepsilon \downarrow 0$. Explain why a bounded-incidence argument alone cannot compare negative-order dyadic and cylinder moments.

I.10.11.9 Exercise 9 — Quasi-Bernoulli tails

Prove (I.10.170) first for finite disjoint unions of tail cylinders and then for all Borel tail events. Deduce (I.10.171) for a finite stopping antichain.

I.10.11.10 Exercise 10 — Perturbed Fekete limits

Let $A_n, c_n > 0$, $\log c_n/n \to 0$, and $A_{m+n} \le c_n A_m A_n$. Fill in the remainder term in (I.10.186) and prove that $\lim_n n^{-1}\log A_n$ exists. State the exact reversed argument for a supermultiplicative inequality.

I.10.11.11 Exercise 11 — Constructing the slow tilt

Verify every assertion of Lemma 7.4 for the recursive choice

$$t_{n+1} = \begin{cases} t_n/2, & A_n b_n e^{-an} \ge 1, \\ t_n, & A_n b_n e^{-an} < 1. \end{cases}$$

In particular, prove both $t_n \to 0$ and divergence of the tilted series.

I.10.11.12 Exercise 12 — From C^1 continuity to $e^{o(n)}$

Let $\omega(\rho)$ be a common modulus of continuity of the logarithmic derivative scales of a finite conformal IFS. Show that

$$\sum_{k=0}^{n-1} \omega(\vartheta^k) = o(n).$$

Explain how exponentiating this estimate leads to a subexponential distortion factor.

I.10.11.13 Exercise 13 — Positive-order grid/packing comparison

Supply the coloring argument in Lemma 8.1 in detail. Choose an explicit integer $L = L(d)$ that makes the balls centered in cubes of the same color disjoint, and show that the number of colors is independent of the scale.

I.10.11.14 Exercise 14 — Dependency chain for Feng's theorem

Assume m is two-sided quasi-Bernoulli and $t \geq 1$ is a differentiability point of τ_μ^{pack}. List, in logical order, the exact internal results that establish:

1. existence of the positive-order dyadic spectrum;
2. equality with the centered-packing spectrum;
3. the upper t-Gibbs ball estimate;
4. concentration on $E_\mu(\tau'(t))$; and
5. equality in the Legendre upper bound.

For each arrow, name the hypothesis it consumes.

Chapter I.11

Random Recursive and Percolation Fractals

Random construction introduces a second issue beyond geometric scaling: the construction may die out. We first establish the probability and martingale tools needed to distinguish extinction from survival. We then study random recursive sets, fractal percolation, and multiplicative cascades. Dimension statements and limiting measures must always be read with their survival and nondegeneracy hypotheses.

The definitions below fix this chapter's conventions. Geometric intuition and proof strategy are discussed separately after the main development; exercises and their solutions provide a second route through the material.

Definitions and notation

All random objects live on a declared probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Euclidean similarity and dimension terminology already accepted in I.1–I.8 is reused with its established meaning.

Definition — I.11.01: Probability, filtration, and conditioning conventions

Random variables $(X_j)_{j\in J}$ are **independent** if for every finite set of distinct indices $j_1, \dots, j_k$ and Borel sets $B_1, \dots, B_k$,

$$\mathbb{P}\left(\bigcap_{r=1}^{k}\{X_{j_r} \in B_r\}\right) = \prod_{r=1}^{k} \mathbb{P}(X_{j_r} \in B_r).$$

They are **i.i.d.** if they are independent and have one common law. A **filtration** is an increasing sequence $\mathscr{F}_0 \subset \mathscr{F}_1 \subset \cdots$ of sub-sigma-algebras of $\mathscr{F}$. A sequence (Y_n) is **adapted** if each Y_n is $\mathscr{F}_n$-measurable. An integrable adapted sequence (Y_n) is a **martingale** if

$$\mathbb{E}(Y_{n+1} \mid \mathscr{F}_n) = Y_n \quad \text{almost surely for every } n.$$

An event holds **almost surely** if its complement has probability zero. If $S \in \mathscr{F}$ and $\mathbb{P}(S) > 0$, then a statement holds **almost surely conditioned on** S when its exceptional set N satisfies $\mathbb{P}(N \cap S) = 0$; the conditional probability is

$$\mathbb{P}(A \mid S) = \frac{\mathbb{P}(A \cap S)}{\mathbb{P}(S)}.$$

Definition — I.11.02: Rooted word trees and survival

Let Λ be a finite alphabet, let $\varnothing$ be the empty word, and let $\Lambda^* = \bigcup_{n\geq 0} \Lambda^n$. A **rooted subtree** is a prefix-closed set $T \subset \Lambda^*$ containing $\varnothing$. Its generation and boundary are

$$T_n = T \cap \Lambda^n, \qquad \partial T = \{\omega \in \Lambda^{\mathbb{N}} : \omega|_n \in T \text{ for every } n\}.$$

Write $Z_n = \#T_n$. **Extinction** is the event $\mathrm{Ext} = \{Z_n = 0 \text{ for some } n\}$; **survival** is $\mathrm{Surv} = \mathrm{Ext}^c$. For a finite-branching tree, survival is equivalent to $\partial T \neq \varnothing$.

Definition — I.11.03: Galton–Watson law

In this chapter let ξ be a bounded nonnegative integer-valued random variable, with $\xi \leq L$ almost surely. Use an alphabet with at least L symbols and label the children $1, \dots, \xi$. A **Galton–Watson tree with offspring law** ξ is a rooted random tree in which the numbers of children of all vertices are independent copies of ξ. Its probability-generating function and mean are

$$f(z) = \mathbb{E}(z^\xi) = \sum_{k\geq 0} \mathbb{P}(\xi = k) z^k, \qquad m = \mathbb{E}\xi.$$

The process is **subcritical**, **critical**, or **supercritical** when $m < 1$, $m = 1$, or $m > 1$, respectively. Unbounded offspring laws are not part of this finite-alphabet kernel. When $m > 0$, its **normalized population process** is

$$W_n = \frac{Z_n}{m^n}.$$

The natural filtration $\mathscr{F}_n$ is generated by every offspring variable attached to a vertex of generation less than n.

Definition — I.11.04: M-adic fractal percolation

Fix integers $M \geq 2$ and $d \geq 1$, choose $p \in [0,1]$, and put $b = M^d$. Divide $[0,1]^d$ into its b congruent closed M-adic subcubes. Retain each child independently with probability p, and repeat the same independent experiment inside every retained cube. If $\mathcal{Q}_n(\omega)$ is the family of retained generation-n cubes, define

$$K_n(\omega) = \bigcup_{Q \in \mathcal{Q}_n(\omega)} Q, \qquad K(\omega) = \bigcap_{n=1}^{\infty} K_n(\omega).$$

This is **homogeneous M-adic fractal percolation with parameter p**. Its survival event is $\{K \neq \emptyset\}$, equivalently survival of the Galton–Watson tree with offspring law $\mathrm{Bin}(b, p)$. Its **natural filtration** $(\mathscr{F}_n)$ is generated by all retain/delete decisions for cubes of generations at most n.

Definition — I.11.05: Finite random recursive similarity construction

Let $\mathcal{S} = (S_i)_{i\in\Lambda}$ be a finite similarity IFS, with ratios $0 < r_i < 1$. Let $(A_u)_{u\in\Lambda^*}$ be i.i.d. random subsets of Λ. The **active tree** is

$$T = \{u = u_1 \cdots u_n \;:\; u_j \in A_{u|_{j-1}} \text{ for } 1 \le j \le n\}.$$

With $K_{\mathcal{S}}$ the deterministic attractor and $S_u = S_{u_1} \circ \cdots \circ S_{u_n}$, the **random recursive limit set** is

$$K_T = \bigcap_{n=1}^{\infty} \bigcup_{u\in T_n} S_u(K_{\mathcal{S}}).$$

The model is **supercritical** when $\mathbb{E}\#A_\emptyset > 1$. The symbolic branching results do not require separation. The geometric dimension theorem T-I.11.03 additionally assumes SSC for the deterministic IFS; the cube-percolation argument treats its shared boundaries separately.

Definition — I.11.06: Mean pressure and weighted branching measure

For the model of D-I.11.05 define

$$\Phi(t) = \mathbb{E}\left(\sum_{i\in A_\emptyset} r_i^t\right) \quad (t \ge 0).$$

In the supercritical case the unique $s > 0$ satisfying $\Phi(s) = 1$ is the **Malthusian exponent**. For $u = u_1 \cdots u_n$, put $r_u = \prod_{j=1}^n r_{u_j}$ and

$$X_n = \sum_{u\in T_n} r_u^s.$$

For every $u \in T$, let $T^{(u)}$ be the descendant tree rooted at u, with u relabelled as the empty word, and let X_u denote the limit of the analogous weighted martingale in $T^{(u)}$; this limit may be zero. On $X_\emptyset > 0$, the **natural branching measure** on ∂T is determined by

$$\nu_T([u]) = \frac{r_u^s X_u}{X_\emptyset} \quad (u \in T).$$

Its geometric projection is $(\pi_{\mathcal{S}})_*\nu_T$.

Definition — I.11.07: Scalar Mandelbrot cascade

Fix M, d as in D-I.11.04 and $b = M^d$. Let $W \geq 0$ be bounded with $\mathbb{E}W = 1$, and attach independent copies W_u to the non-root vertices of the full b-ary cube tree. Use the standard half-open M-adic partition, closing only the outer faces of $[0,1]^d$, so every point receives one cube label at every generation. For a generation-n cube Q_u, define

$$\mu_n(Q_u) = b^{-n} \prod_{j=1}^{n} W_{u|_j}.$$

Equivalently, μ_n has constant density $\prod_{j=1}^{n} W_{u|_j}$ on Q_u. The total mass is $Y_n = \mu_n([0,1]^d)$. For every word u, Y_u denotes the limiting total mass of the independently rescaled descendant cascade rooted at u. A weak limit μ is a **non-degenerate scalar Mandelbrot cascade** if $\mathbb{P}(Y > 0) > 0$, where $Y = \mu([0,1]^d)$. On $\{Y > 0\}$, its **normalized cascade probability** is

$$\overline{\mu} = Y^{-1}\mu.$$

The **positive-weight tree** retains exactly the children for which $W_u > 0$.

Definition — I.11.08: Peyrière probability and size-biased law

Assume $\mathbb{E}Y = 1$. The **Peyrière probability** on $\Omega \times [0,1]^d$ is

$$\mathbb{Q}(B) = \mathbb{E}\Big(\int \mathbf{1}_B(\omega, x)\, d\mu_\omega(x)\Big).$$

If x has successive cube labels $I_1(x), I_2(x), \ldots$, then the weights $W_{I_1(x)}, W_{I_1(x)I_2(x)}, \ldots$ form the **spine weights**. The **size-biased law** of W is the probability law

$$\mathbb{P}^*(W \in B) = \mathbb{E}\big(W \mathbf{1}_{\{W \in B\}}\big).$$

The conventions $0 \log 0 = 0$ and $\log 0 = -\infty$ are used. The size-biased law assigns mass zero to $W = 0$; thus logarithms of spine weights are finite almost surely.

Definition — I.11.09: Projection and independent-sum preview notation

For a line $\ell \subset \mathbb{R}^2$, let P_ℓ be the orthogonal projection onto ℓ. If $E^{(1)}, \ldots, E^{(k)} \subset \mathbb{R}$ and $a_1, \ldots, a_k \in \mathbb{R}$, their **algebraic sum with coefficients** a is

$$a_1 E^{(1)} + \cdots + a_k E^{(k)} = \left\{ \sum_{j=1}^{k} a_j x_j \;:\; x_j \in E^{(j)} \right\}.$$

The word **simultaneously** in a random preview means that one probability-one event works for every parameter named in the theorem, rather than a separate exceptional event for each parameter.

Terminal terminology

Q-I.11.11 prints definitions, before use, of its crossing threshold, projection Hölder densities, intersection averages, Salem-measure convention, the four spatial-independence conditions, and random affine code trees. These boundary definitions add no outgoing F premise and import no theorem.

Every statement labelled F is proved below. The two statements labelled G are quoted with exact source locators and are never used as premises. Intuition and historical commentary follow the main development.

I.11.1 Probability interface and the local martingale theorem

All random objects are defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Random variables $(U_j)_{j\in J}$ are independent when, for distinct $j_1, \dots, j_r$ and Borel sets $B_1, \dots, B_r$,

$$\mathbb{P}\left(\bigcap_{\ell=1}^{r}\{U_{j_\ell} \in B_\ell\}\right) = \prod_{\ell=1}^{r} \mathbb{P}(U_{j_\ell} \in B_\ell). \tag{I.11.1}$$

They are identically distributed when their push-forward probability laws coincide, and i.i.d. when both properties hold. A filtration is an increasing sequence

$$\mathscr{F}_0 \subset \mathscr{F}_1 \subset \cdots \subset \mathscr{F}. \tag{I.11.2}$$

An integrable sequence (M_n) is adapted when M_n is $\mathscr{F}_n$-measurable. An integrable adapted sequence is a martingale when

$$\mathbb{E}(M_{n+1} \mid \mathscr{F}_n) = M_n \quad \text{almost surely.} \tag{I.11.3}$$

An event holds almost surely when its complement has probability zero. If $S \in \mathscr{F}$ and $\mathbb{P}(S) > 0$, a statement holds almost surely conditioned on S when its exceptional event N satisfies $\mathbb{P}(N \cap S) = 0$, equivalently when

$$\mathbb{P}(N \mid S) = \frac{\mathbb{P}(N \cap S)}{\mathbb{P}(S)} = 0. \tag{I.11.3a}$$

Conditional expectation, including existence, uniqueness, the tower property, and the pull-out property, is imported from ARA-KB1, Problems 23–26, printed pp. 439–440 / local PDF pp. 460–461. The strong law for i.i.d. integrable variables is imported from ARA-KA1, Theorem 9.8, printed pp. 394–398 / local PDF pp. 417–421. The summable-event Borel–Cantelli lemma is imported from ARA-KA1, Lemma 9.9, printed p. 394 / local PDF p. 417. These are the only probability results used without local proof.

We shall repeatedly use the following consequences of integration. If $V \geq 0$, $t > 0$, and $V \in L^2$, then

$$\mathbb{P}(V > t) \leq t^{-1}\mathbb{E}V, \qquad \mathbb{P}(|V - \mathbb{E}V| > t) \leq t^{-2} \operatorname{Var} V. \tag{I.11.4}$$

Indeed, $t\mathbf{1}_{\{V>t\}} \leq V$ and $t^2\mathbf{1}_{\{|V-\mathbb{E}V|>t\}} \leq |V - \mathbb{E}V|^2$; integration proves both inequalities.

Lemma — 1.1 — The exact L^2-martingale convergence tool

Let $(M_n, \mathscr{F}_n)$ be a real martingale with

$$\sup_{n\geq 0} \mathbb{E}|M_n|^2 < \infty. \tag{I.11.5}$$

Then there is $M_\infty \in L^2$ such that

$$M_n \longrightarrow M_\infty \quad \text{almost surely and in } L^2. \tag{I.11.6}$$

For every finite square-integrable real martingale $(N_k)_{k=0}^r$,

$$\mathbb{E} \max_{0\leq k\leq r} |N_k|^2 \leq 4\mathbb{E}|N_r|^2. \tag{I.11.7}$$

Proof

Put $N^* = \max_{0\leq k\leq r} |N_k|$. For $\lambda > 0$, let A_k be the event that k is the first index with $|N_k| \geq \lambda$. The A_k are disjoint, $A_k \in \mathscr{F}_k$, and their union is $\{N^* \geq \lambda\}$. Positivity of conditional expectation applied to $-|N_r| \leq N_r \leq |N_r|$, and the martingale property, give

$$|N_k| = |\mathbb{E}(N_r \mid \mathscr{F}_k)| \leq \mathbb{E}(|N_r| \mid \mathscr{F}_k). \tag{I.11.8}$$

Consequently,

$$\lambda \mathbb{P}(N^* \geq \lambda) \leq \sum_{k=0}^r \int_{A_k} |N_k| \, d\mathbb{P} \leq \mathbb{E}\big(|N_r| \mathbf{1}_{\{N^*\geq\lambda\}}\big). \tag{I.11.9}$$

Integrating the tail identity for $(N^*)^2$, using Tonelli, and then Cauchy–Schwarz yields

$$\begin{aligned} \mathbb{E}(N^*)^2 &= 2\int_0^\infty \lambda \mathbb{P}(N^* \geq \lambda)\, d\lambda \\ &\leq 2\mathbb{E}\left(|N_r| \int_0^{N^*} d\lambda\right) = 2\mathbb{E}(|N_r| N^*) \\ &\leq 2\|N_r\|_2 \|N^*\|_2. \end{aligned} \tag{I.11.10}$$

This proves (I.11.7), including the case $\|N^*\|_2 = 0$.
Martingale differences are orthogonal in L^2: if $j < k$, then

$$\mathbb{E}\big((M_{j+1} - M_j)(M_{k+1} - M_k)\big) = \mathbb{E}\big((M_{j+1} - M_j)\mathbb{E}(M_{k+1} - M_k \mid \mathscr{F}_k)\big) = 0. \tag{I.11.11}$$

Hence $\mathbb{E}|M_n|^2$ is nondecreasing and bounded, and, for $n > N$,

$$\mathbb{E}|M_n - M_N|^2 = \mathbb{E}|M_n|^2 - \mathbb{E}|M_N|^2 \longrightarrow 0 \quad \text{uniformly in } n > N \text{ as } N \to \infty. \tag{I.11.12}$$

Thus (M_n) is Cauchy in the complete space L^2; write its L^2 limit as M_∞. Apply (I.11.7) to the finite martingale $M_{N+k} - M_N$, $0 \leq k \leq r$, and then let $r \to \infty$. From (I.11.4) and monotone convergence,

$$\mathbb{P}\Big(\sup_{n\geq N}|M_n - M_N| > \varepsilon\Big) \leq \frac{4}{\varepsilon^2}\sup_{n\geq N}\mathbb{E}|M_n - M_N|^2. \tag{I.11.13}$$

Choose $N_j \uparrow \infty$ so that the right side of (I.11.13), with $\varepsilon = 2^{-j}$, is at most 2^{-j}. Borel–Cantelli implies that almost surely

$$\sup_{n\geq N_j}|M_n - M_{N_j}| \leq 2^{-j} \quad \text{for all sufficiently large } j. \tag{I.11.14}$$

The entire sequence is therefore almost surely Cauchy. Its almost-sure limit agrees with its L^2 limit, because L^2 convergence implies convergence in probability and probability limits are unique. This proves (I.11.6). □

I.11.2 Rooted trees and Galton–Watson growth

Let Λ be a finite alphabet, $\Lambda^* = \bigcup_{n\geq 0}\Lambda^n$, and let $\varnothing$ be the empty word. For $u = u_1 \cdots u_n$, write $|u| = n$ and $u|_j = u_1 \cdots u_j$. A rooted subtree $T \subset \Lambda^*$ contains $\varnothing$ and is prefix closed. Its generation, population, and boundary are

$$T_n = T \cap \Lambda^n, \qquad Z_n = \#T_n, \qquad \partial T = \{\omega \in \Lambda^{\mathbb{N}} : \omega|_n \in T \text{ for all } n\}. \tag{I.11.15}$$

Extinction and survival are the events

$$\mathsf{Ext} = \{Z_n = 0 \text{ for some } n\}, \qquad \mathsf{Surv} = \mathsf{Ext}^c. \tag{I.11.16}$$

For a finitely branching tree,

$$\mathsf{Surv} = \{\partial T \neq \varnothing\}. \tag{I.11.17}$$

To prove the nontrivial inclusion, assume $T_n \neq \varnothing$ for every n. Some child of the root has descendants at arbitrarily large depths; choose the first such child. Iterating this finite-pigeonhole argument constructs an infinite branch.

Let ξ be a bounded nonnegative integer-valued variable, and choose $\Lambda = \{1, \dots, L\}$ with $\xi \leq L$ almost surely. Label the children of a vertex $1, \dots, \xi$. A Galton–Watson tree with offspring law ξ gives all vertices independent copies of ξ children. Put

$$f(z) = \mathbb{E}z^{\xi}, \qquad m = \mathbb{E}\xi, \qquad \sigma^2 = \operatorname{Var}\xi. \tag{I.11.18}$$

The process is called subcritical, critical, or supercritical according as $m < 1$, $m = 1$, or $m > 1$. For a square-integrable variable V, the conditional variance notation used below means $\operatorname{Var}(V \mid \mathscr{F}_n) = \mathbb{E}((V - \mathbb{E}(V \mid \mathscr{F}_n))^2 \mid \mathscr{F}_n)$.

The natural $\mathscr{F}_n$ is generated by offspring variables attached to vertices of generations less than n.

Lemma — L-I.11.01 — Galton–Watson extinction and population martingale [F]

Suppose $Z_0 = 1$ and $m > 0$. The extinction probability q is the smallest fixed point of f in $[0,1]$. If $m < 1$, or if $m = 1$ and $\mathbb{P}(\xi = 1) < 1$, then $q = 1$. If $m > 1$, then $q < 1$. In the supercritical case,

$$W_n = \frac{Z_n}{m^n} \tag{I.11.19}$$

is an L^2-bounded martingale satisfying

$$\mathbb{E}W_n^2 = 1 + \frac{\sigma^2}{m(m-1)}(1 - m^{-n}). \tag{I.11.20}$$

Its limit W satisfies

$$\mathbb{P}(W > 0) = 1 - q, \qquad \lim_{n\to\infty} \frac{\log Z_n}{n} = \log m \quad \text{on Surv.} \tag{I.11.21}$$

Proof

Let $q_n = \mathbb{P}(Z_n = 0)$. By conditioning on the root offspring and using independence of its descendant trees,

$$q_{n+1} = \mathbb{E}(q_n^{\xi}) = f(q_n), \qquad q_0 = 0. \tag{I.11.22}$$

The events $\{Z_n = 0\}$ increase to Ext, so $q_n \uparrow q = \mathbb{P}(\mathrm{Ext})$. Continuity of the polynomial f gives $f(q) = q$. If $a \in [0,1]$ and $f(a) = a$, monotonicity and (I.11.22) give $q_n \le a$ for all n, hence $q \le a$. Thus q is the least fixed point.

Write $g(z) = f(z) - z$. Since f is convex and $f'(1) = m$, if $m < 1$ then $g'(z) \le m - 1 < 0$ on $[0,1]$. Since $g(1) = 0$, one has $g(z) > 0$ for $z < 1$, and $q = 1$. If $m = 1$ and $\mathbb{P}(\xi = 1) < 1$, then $\mathbb{P}(\xi \ge 2) > 0$; otherwise the mean-one law would be concentrated at 1. Hence f is strictly convex on $(0,1)$, and its tangent at 1 gives

$$f(z) > f(1) + f'(1)(z-1) = z \quad (0 \le z < 1). \tag{I.11.23}$$

Again $q = 1$. If $m > 1$, differentiability at 1 gives $g(z) < 0$ for all $z < 1$ sufficiently close to 1, while $g(0) = \mathbb{P}(\xi = 0) \ge 0$. If $g(0) = 0$, then 0 is already a fixed point; otherwise the intermediate value theorem supplies one in $(0,1)$. In either case the least fixed point satisfies $q < 1$.

Conditionally on $\mathscr{F}_n$, Z_{n+1} is a sum of Z_n independent copies of ξ. Therefore

$$\mathbb{E}(Z_{n+1} \mid \mathscr{F}_n) = mZ_n, \qquad \mathrm{Var}(Z_{n+1} \mid \mathscr{F}_n) = \sigma^2 Z_n. \tag{I.11.24}$$

The first identity proves the martingale property of W_n, and induction gives $\mathbb{E}Z_n = m^n$. The conditional second-moment identity is

$$\mathbb{E}(Z_{n+1}^2 \mid \mathscr{F}_n) = m^2 Z_n^2 + \sigma^2 Z_n. \tag{I.11.25}$$

After division by m^{2n+2} and integration,

$$\mathbb{E}W_{n+1}^2 = \mathbb{E}W_n^2 + \frac{\sigma^2}{m^{n+2}}. \tag{I.11.26}$$

Summing the geometric series proves (I.11.20). Lemma 1.1 gives almost-sure and L^2 convergence to W, and L^2 convergence preserves the mean:

$$\mathbb{E}W = 1. \tag{I.11.27}$$

Let $W^{(1)}, W^{(2)}, \dots$ be the limits in the independent descendant trees of the root. The boundedness of ξ permits passage to the limit in the finite branching identity, giving

$$W = \frac{1}{m}\sum_{j=1}^{\xi} W^{(j)} \quad \text{almost surely.} \tag{I.11.28}$$

If $r = \mathbb{P}(W = 0)$, nonnegativity and independence in (I.11.28) imply

$$r = \mathbb{E}(r^{\xi}) = f(r). \tag{I.11.29}$$

Equation (I.11.27) gives $r < 1$. In the supercritical case convexity shows that $f(z) = z$ has exactly one solution below 1, namely q; hence $r = q$. Extinction plainly implies $W = 0$, and the two events have the same probability, so

$$\{W > 0\} = \mathrm{Surv} \quad \text{modulo a null set.} \tag{I.11.30}$$

On this event $Z_n = m^n W_n$ and $W_n \to W \in (0, \infty)$. Thus $\log W_n / n \to 0$, which proves the second assertion in (I.11.21). □

I.11.3 Weighted branching and its natural measure

Let $\mathcal{S} = (S_i)_{i \in \Lambda}$ be a finite similarity IFS on $\mathbb{R}^d$, with contraction ratios $0 < r_i < 1$. No separation hypothesis is needed in this section: only the ratios and the random tree enter its proof. SSC will be imposed in Section 4. To every $u \in \Lambda^*$ attach an i.i.d. random subset $A_u \subset \Lambda$. The active tree is

$$T = \{u_1 \cdots u_n \, : \, u_j \in A_{u|_{j-1}} \text{ for } 1 \le j \le n\}. \tag{I.11.31}$$

For $u = u_1 \cdots u_n$, put

$$S_u = S_{u_1} \circ \cdots \circ S_{u_n}, \qquad r_u = \prod_{j=1}^{n} r_{u_j}. \tag{I.11.32}$$

Writing $K_{\mathcal{S}}$ for the deterministic attractor, the random recursive limit is

$$K_T = \bigcap_{n \ge 1} \bigcup_{u \in T_n} S_u(K_{\mathcal{S}}). \tag{I.11.33}$$

Assume throughout this section that

$$\mathbb{E}\#A_\varnothing > 1. \tag{I.11.34}$$

Define the mean pressure

$$\Phi(t) = \mathbb{E} \sum_{i \in A_\varnothing} r_i^t \quad (t \geq 0). \tag{I.11.35}$$

It is continuous and strictly decreasing from $\Phi(0) = \mathbb{E}\#A_\varnothing > 1$ to 0. Thus there is a unique Malthusian exponent $s > 0$ such that

$$\Phi(s) = 1. \tag{I.11.36}$$

Since every $r_i^s \in (0, 1)$,

$$c := \Phi(2s) < \Phi(s) = 1. \tag{I.11.37}$$

For every word u, the full collection of descendant variables $(A_{uv})_{v \in \Lambda^*}$ defines a copy of the same random environment, whether or not u is active in the root tree.

Lemma — L-I.11.02 — Weighted branching martingale and natural measure [F]

Let

$$X_n = \sum_{u \in T_n} r_u^s. \tag{I.11.38}$$

Then (X_n) is L^2 bounded and converges almost surely and in L^2 to $X_\varnothing$, with

$$\mathbb{E}X_\varnothing = 1, \qquad \{X_\varnothing > 0\} = \mathrm{Surv} \quad \text{modulo a null set.} \tag{I.11.39}$$

For every u, let X_u be the corresponding limit in the descendant environment at u. On survival,

$$\nu_T([u]) = \frac{\mathbf{1}_{\{u \in T\}} r_u^s X_u}{X_\varnothing} \tag{I.11.40}$$

defines a probability measure on ∂T, and for ν_T-almost every ω,

$$\lim_{n\to\infty} \frac{\log \nu_T([\omega|_n])}{\log r_{\omega|_n}} = s. \tag{I.11.41}$$

Proof: martingale and second moment

Let $\mathscr{F}_n$ be generated by the variables A_u with $|u| < n$. For $u \in T_n$,

$$\mathbb{E}\left(\sum_{i \in A_u} r_{ui}^s \,\middle|\, \mathscr{F}_n \right) = r_u^s \Phi(s) = r_u^s. \tag{I.11.42}$$

Summing over T_n proves $\mathbb{E}(X_{n+1} \mid \mathscr{F}_n) = X_n$, and hence $\mathbb{E}X_n = 1$.
Set

$$a_i = \mathbf{1}_{\{i \in A_\varnothing\}} r_i^s, \qquad B = \mathbb{E} \sum_{i \neq j} a_i a_j. \tag{I.11.43}$$

If $X_n^{(i)}$ are independent depth-n weighted sums in the child environments, independent of (a_i), then

$$X_{n+1} = \sum_{i \in \Lambda} a_i X_n^{(i)}. \tag{I.11.44}$$

Since $\mathbb{E} X_n^{(i)} = 1$, squaring (I.11.44) and using independence gives

$$\mathbb{E} X_{n+1}^2 = c\, \mathbb{E} X_n^2 + B. \tag{I.11.45}$$

Both B and $X_0 = 1$ are finite. Iteration and (I.11.37) yield

$$\sup_n \mathbb{E} X_n^2 \leq 1 + \frac{B}{1-c} < \infty. \tag{I.11.46}$$

Lemma 1.1 gives the asserted almost-sure and L^2 limit, and L^2 convergence gives $\mathbb{E} X_\varnothing = 1$.

Proof: positivity is survival

Apply the preceding construction simultaneously to the countably many descendant environments. Passing to the limit in their finite recursions gives, on one probability-one event,

$$X_u = \sum_{i \in A_u} r_i^s X_{ui} \quad \text{for every } u \in \Lambda^*. \tag{I.11.47}$$

Let $h(z) = \mathbb{E} z^{\# A_\varnothing}$, and let $\rho = \mathbb{P}(X_\varnothing = 0)$. All summands in (I.11.47) are nonnegative, so independence gives

$$\rho = h(\rho). \tag{I.11.48}$$

Since $\mathbb{E} X_\varnothing = 1$, one has $\rho < 1$. The offspring mean is $h'(1) = \mathbb{E} \# A_\varnothing > 1$; the fixed-point argument in L-I.11.01 shows that the only fixed point of h below 1 is the extinction probability. Extinction implies $X_\varnothing = 0$, so equality of the two probabilities proves (I.11.39).

Proof: cylinder consistency

For $u \in T$, (I.11.47) implies

$$\sum_{i:\, ui \in T} r_{ui}^s X_{ui} = r_u^s X_u. \tag{I.11.49}$$

After division by $X_\varnothing$, this is exactly

$$\sum_{i:\, ui \in T} \nu_T([ui]) = \nu_T([u]), \qquad \nu_T([\varnothing]) = 1. \tag{I.11.50}$$

The consistent nonnegative masses on the finite cylinder algebras therefore define a unique Borel probability on the compact symbolic boundary. This is the cylinder-measure construction already proved in D/R-I.7; equation (I.11.50) verifies every hypothesis in the present random setting.

Proof: the annealed size-biased branch

On extinction choose a fixed probability on the full space $\Lambda^{\mathbb{N}}$, not on the empty boundary. On survival regard ν_T as a probability on that full space carried by ∂T. Its cylinder masses are measurable functions of the environment by (I.11.40); a monotone-class argument extends this to every Borel set. Consequently integration against this family defines a measure on the product sigma-algebra. On extinction its value is multiplied by $X_\varnothing = 0$. Define a probability $\mathbb{Q}_T$ on environments and branches by

$$\mathbb{Q}_T(B) = \mathbb{E}\left(X_\varnothing \int_{\Lambda^{\mathbb{N}}} \mathbf{1}_B(\omega_{\text{env}}, \omega)\, d\nu_T(\omega)\right). \tag{I.11.51}$$

This is an **annealed** law, meaning that it averages both the random environment and the sampled branch, with the displayed weight. Its total mass is $\mathbb{E}X_\varnothing = 1$. Put

$$\theta_i = \mathbb{P}(i \in A_\varnothing) r_i^s. \tag{I.11.52}$$

Equation (I.11.36) says $\sum_i \theta_i = 1$. If $u = u_1 \cdots u_n$, then X_u is independent of the ancestral activation events and has mean one. Using (I.11.40),

$$\begin{aligned}\mathbb{Q}_T([\omega|_n = u]) &= \mathbb{E}\big(\mathbf{1}_{\{u\in T\}} r_u^s X_u\big) \\ &= \prod_{j=1}^n \mathbb{P}(u_j \in A_\varnothing) r_{u_j}^s = \prod_{j=1}^n \theta_{u_j}.\end{aligned} \tag{I.11.53}$$

Thus the branch labels are i.i.d. with law (θ_i). The strong law gives

$$\frac{1}{n}\log r_{\omega|_n} = \frac{1}{n}\sum_{j=1}^n \log r_{\omega_j} \longrightarrow \lambda := \sum_{i\in\Lambda} \theta_i \log r_i < 0 \quad \mathbb{Q}_T\text{-a.s.} \tag{I.11.54}$$

The descendant limit at the selected vertex has the size-biased law of $X_\varnothing$. Indeed, for every Borel $C \subset [0,\infty)$,

$$\begin{aligned}\mathbb{Q}_T(X_{\omega|_n} \in C) &= \sum_{|u|=n} \mathbb{E}\big(\mathbf{1}_{\{u\in T\}} r_u^s X_u \mathbf{1}_{\{X_u\in C\}}\big) \\ &= \mathbb{E}\big(X_\varnothing \mathbf{1}_{\{X_\varnothing\in C\}}\big) \sum_{|u|=n}\prod_{j=1}^n \theta_{u_j} \\ &= \mathbb{E}\big(X_\varnothing \mathbf{1}_{\{X_\varnothing\in C\}}\big).\end{aligned} \tag{I.11.55}$$

Fix $\varepsilon > 0$. From (I.11.55),

$$\begin{aligned}\mathbb{Q}_T(X_{\omega|_n} < e^{-\varepsilon n}) &\le e^{-\varepsilon n}, \\ \mathbb{Q}_T(X_{\omega|_n} > e^{\varepsilon n}) &\le e^{-\varepsilon n}\mathbb{E}X_\varnothing^2.\end{aligned} \tag{I.11.56}$$

Both series are summable. Borel–Cantelli, first for rational $\varepsilon > 0$ and then by intersection, yields

$$\frac{1}{n} \log X_{\omega|_n} \longrightarrow 0 \quad \mathbb{Q}_T\text{-a.s.} \tag{I.11.57}$$

The selected limit is positive $\mathbb{Q}_T$-almost surely by its size-biased law, so the logarithm is finite. From (I.11.40),

$$\log \nu_T([\omega|_n]) = s \log r_{\omega|_n} + \log X_{\omega|_n} - \log X_{\varnothing}. \tag{I.11.58}$$

Equations (I.11.54), (I.11.57), and (I.11.58) prove (I.11.41) $\mathbb{Q}_T$-almost surely. If B is the exceptional branch set, then

$$0 = \mathbb{Q}_T(B) = \mathbb{E}\big(X_{\varnothing} \nu_T(B)\big). \tag{I.11.59}$$

On survival $X_{\varnothing} > 0$; hence $\nu_T(B) = 0$ for almost every surviving environment. This completes the proof. □

I.11.4 The finite separated random-recursive dimension theorem

For $\omega, \tau \in \Lambda^{\mathbb{N}}$, let $\omega \wedge \tau$ be their longest common prefix and define

$$\rho_{\mathcal{S}}(\omega, \tau) = \begin{cases} r_{\omega \wedge \tau}, & \omega \neq \tau, \\ 0, & \omega = \tau. \end{cases} \tag{I.11.60}$$

This is the weighted symbolic ultrametric of D/R-I.8. If $\pi_{\mathcal{S}}$ is the coding map and SSC holds, L-I.8.04 gives constants $0 < a \le A < \infty$ such that

$$a \rho_{\mathcal{S}}(\omega, \tau) \le |\pi_{\mathcal{S}}(\omega) - \pi_{\mathcal{S}}(\tau)| \le A \rho_{\mathcal{S}}(\omega, \tau). \tag{I.11.61}$$

Thus $\pi_{\mathcal{S}}$ is bi-Lipschitz on every subset of the symbolic space. The closed $\rho_{\mathcal{S}}$-ball of radius $r_{\omega|_n}$ centered at ω is exactly $[\omega|_n]$.

Theorem — T-I.11.03 — Finite separated random recursive dimension [F]

Under (I.11.31)–(I.11.37), assume additionally that the deterministic IFS satisfies SSC. Then almost surely conditioned on survival,

$$\dim_{\mathrm{H}} K_T = s. \tag{I.11.62}$$

Moreover,

$$\eta_T = (\pi_{\mathcal{S}})_* \nu_T \tag{I.11.63}$$

is exact dimensional of dimension s.

This is a fully proved finite-IFS SSC specialization of Mauldin–Williams, "Random Recursive Constructions: Asymptotic Geometric and Topological Properties," Theorems 1.1 and 1.3, local MW86-RR PDF pp. 4 and 7 / printed pp. 327 and 330; the random-measure architecture is compared with Theorem 3.6, local PDF p. 16 / printed p. 339. No theorem from that paper is used as a premise.

Proof: Hausdorff upper bound

Let $D = \operatorname{diam} K_{\mathcal{S}}$. For $t > s$, $\Phi(t) < 1$, and generation n gives the cover

$$K_T \subset \bigcup_{u\in T_n} S_u(K_{\mathcal{S}}), \qquad \operatorname{diam} S_u(K_{\mathcal{S}}) = Dr_u. \tag{I.11.64}$$

Independence along each word gives

$$\mathbb{E} \sum_{u\in T_n} r_u^t = \left(\mathbb{E} \sum_{i\in A_\varnothing} r_i^t\right)^n = \Phi(t)^n. \tag{I.11.65}$$

Therefore Tonelli implies

$$\mathbb{E} \sum_{n=1}^{\infty} \sum_{u\in T_n} (Dr_u)^t = D^t \sum_{n=1}^{\infty} \Phi(t)^n < \infty. \tag{I.11.66}$$

The double series is finite almost surely, so its generation-n summand tends to zero. If $r_{\max} = \max_i r_i < 1$, the mesh in (I.11.64) is at most $Dr_{\max}^n$. Hence

$$\mathscr{H}^t(K_T) = 0 \quad \text{almost surely.} \tag{I.11.67}$$

Intersecting the probability-one events over rational $t > s$ proves $\dim_{\mathrm{H}} K_T \le s$.

Proof: exact dimension of the natural measure

For a branch satisfying L-I.11.02,

$$\frac{\log \nu_T([\omega|_n])}{\log r_{\omega|_n}} \longrightarrow s. \tag{I.11.68}$$

Let $r_{\min} = \min_i r_i$. Consecutive logarithmic radii satisfy

$$0 < -\log r_{\min} \ge \log r_{\omega|_{n-1}} - \log r_{\omega|_n} \ge -\log r_{\max} > 0. \tag{I.11.69}$$

Also $-\log r_{\omega|_n} \ge n(-\log r_{\max}) \to \infty$. If $r_{\omega|_n} \le r < r_{\omega|_{n-1}}$, the ultrametric ball $\overline{B}_{\rho_{\mathcal{S}}}(\omega, r)$ is exactly $[\omega|_n]$: points outside that cylinder have distance at least $r_{\omega|_{n-1}} > r$, and points inside have distance at most $r_{\omega|_n} \le r$. Equations (I.11.68)–(I.11.69) therefore give

$$\lim_{r\downarrow 0} \frac{\log \nu_T(\overline{B}_{\rho_{\mathcal{S}}}(\omega, r))}{\log r} = s. \tag{I.11.70}$$

The bi-Lipschitz inequalities (I.11.61) sandwich Euclidean balls between symbolic balls whose radii differ by fixed factors. Since fixed multiplicative changes in r contribute $O(1)/\log r$, (I.11.70) implies

$$\lim_{r\downarrow 0} \frac{\log \eta_T(\overline{B}(\pi_{\mathcal{S}}\omega, r))}{\log r} = s \quad \text{for } \nu_T\text{-a.e. } \omega. \tag{I.11.71}$$

Thus η_T is exact dimensional of dimension s.

Finally, the coding identity

$$K_T = \pi_{\mathcal{S}}(\partial T) \tag{I.11.72}$$

follows by the nested compact construction in D/R-I.7: one inclusion follows from every active branch, and the other from the finite-branching argument (I.11.17) applied to the active prefixes whose cylinder sets contain the point. Hence $\eta_T(K_T) = 1$. The local-to-global theorem T-I.4.04 gives $\dim_{\mathrm{H}} K_T \geq \dim_{\mathrm{H}} \eta_T = s$. Together with (I.11.67), this proves (I.11.62). □

I.11.5 Homogeneous fractal percolation

Fix integers $M \geq 2$ and $d \geq 1$, put $b = M^d$, and choose $p \in [0, 1]$. Each retained M-adic cube keeps each of its b children independently with probability p. Let $\mathcal{Q}_n$ be the family of retained closed generation-n cubes and set

$$K_n = \bigcup_{Q\in\mathcal{Q}_n} Q, \qquad K = \bigcap_{n\geq 1} K_n. \tag{I.11.73}$$

The underlying retained-word tree has offspring law $\mathrm{Bin}(b, p)$ and mean

$$m = pb. \tag{I.11.74}$$

The use of closed cubes makes K compact but lets distinct retained cubes share faces. Every argument below records where this matters.

Corollary — C-I.11.04 — Fractal-percolation phase and dimensions [F]

One has

$$\mathbb{P}(K \neq \varnothing) > 0 \quad \Longleftrightarrow \quad pb > 1. \tag{I.11.75}$$

If $pb > 1$, then almost surely conditioned on $K \neq \varnothing$,

$$\dim_{\mathrm{H}} K = \underline{\dim}_{\mathrm{B}} K = \overline{\dim}_{\mathrm{B}} K = \frac{\log(pb)}{\log M} = d + \frac{\log p}{\log M}. \tag{I.11.76}$$

This is proved below. For provenance, the same phase and dimension formulas are recorded in RS13-RP, equations (2.2)–(2.3), local PDF p. 4, where the sources of the classical formula are also identified.

Proof: survival and the Hausdorff upper bound

If the retained tree becomes empty, then some K_n is empty and so is K. Conversely, if every generation is nonempty, finite branching gives an infinite retained branch by (I.11.17); the nested cubes on that branch have a common point. Thus

$$\{K \neq \varnothing\} = \mathsf{Surv}. \tag{I.11.77}$$

If $p = 0$, the first generation is empty almost surely. For $p > 0$, L-I.11.01 applied to $\mathrm{Bin}(b, p)$ proves (I.11.75). At the critical value $pb = 1$, the binomial offspring law is not concentrated at 1, since $b \geq 2$.

Assume $m > 1$ and set

$$s = \frac{\log m}{\log M}. \tag{I.11.78}$$

For $t > s$, the retained generation-n cubes cover K, have diameter $\sqrt{d}\, M^{-n}$, and have expected number m^n. Hence

$$\mathbb{E} \sum_{Q \in \mathcal{Q}_n} (\operatorname{diam} Q)^t = d^{t/2}(mM^{-t})^n. \tag{I.11.79}$$

The right side is summable in n. As in (I.11.66), Tonelli and the vanishing mesh give

$$\mathcal{H}^t(K) = 0 \quad \text{almost surely for every rational } t > s, \tag{I.11.80}$$

and therefore $\dim_{\mathrm{H}} K \leq s$.

Proof: the exponential count of cubes with infinite descendants

Let $Z_n = \#\mathcal{Q}_n$. On survival L-I.11.01 gives

$$\frac{Z_n}{m^n} \longrightarrow W \in (0, \infty), \qquad \frac{1}{n} \log Z_n \longrightarrow \log m. \tag{I.11.81}$$

Call a retained generation-n cube live if its descendant tree survives, and let V_n be the number of live cubes. If q is the extinction probability and $\vartheta = 1 - q > 0$, then, conditionally on $\mathcal{F}_n$,

$$V_n \sim \mathrm{Bin}(Z_n, \vartheta). \tag{I.11.82}$$

In particular, for $\varepsilon > 0$, conditional Chebyshev gives

$$\mathbb{P}\big(|V_n - \vartheta Z_n| > \varepsilon Z_n \bigm| \mathcal{F}_n\big) \leq \frac{\vartheta(1-\vartheta)}{\varepsilon^2 Z_n} \quad \text{on } \{Z_n > 0\}. \tag{I.11.83}$$

For $k \geq 1$, put $S_k = \{W \geq 1/k\}$. On S_k, (I.11.81) implies $Z_n \geq m^n/(2k)$ for all sufficiently large n. Moreover,

$$\sum_{n \geq 1} \mathbb{P}\Big(|V_n - \vartheta Z_n| > \varepsilon Z_n,\ Z_n \geq \frac{m^n}{2k}\Big) \leq \frac{2k\vartheta(1-\vartheta)}{\varepsilon^2} \sum_{n \geq 1} m^{-n} < \infty. \tag{I.11.84}$$

Borel–Cantelli, then a countable intersection over rational $\varepsilon > 0$, shows that

$$\frac{V_n}{Z_n} \longrightarrow \vartheta \quad \text{almost surely on } S_k. \tag{I.11.85}$$

Since $\mathrm{Surv} = \{W > 0\} = \bigcup_{k\geq 1} S_k$ modulo null sets,

$$\frac{1}{n}\log V_n \longrightarrow \log m \quad \text{almost surely on survival.} \tag{I.11.86}$$

Proof: lower and upper box dimensions, with boundary audit

The live generation-n cubes cover K, and each of them contains a point of K. Choose one such point from each live cube. A point of $\mathbb{R}^d$ lies in at most 2^d closed cubes of one fixed grid generation. Also a ball of radius M^{-n} meets at most a dimension-dependent number J_d of generation-n cubes. It follows that there are constants $0 < c_d \leq C_d < \infty$, independent of n and the realization, such that the covering number satisfies

$$c_d V_n \leq N(K, M^{-n}) \leq C_d V_n. \tag{I.11.87}$$

The left inequality follows by applying the two bounded-overlap observations to the chosen points. The right follows by covering each live cube by a bounded number of radius-M^{-n} balls. Thus shared faces can change the count only by a constant factor.

If $M^{-(n+1)} < r \leq M^{-n}$, monotonicity of covering numbers compares $N(K, r)$ with the two adjacent quantities in (I.11.87). Combining this with (I.11.86) proves

$$\underline{\dim}_{\mathrm{B}} K = \overline{\dim}_{\mathrm{B}} K = s \quad \text{almost surely on survival.} \tag{I.11.88}$$

Proof: a natural measure and the Hausdorff lower bound

In the equal-ratio random-recursive model, $r_i = M^{-1}$ and the Malthusian equation is $mM^{-s} = 1$. L-I.11.02 therefore gives a probability ν_T on infinite retained branches and descendant limits X_u. Define the unnormalized geometric measure

$$\lambda_T = X_\varnothing (\pi_M)_* \nu_T, \tag{I.11.89}$$

where π_M is the standard base-M cube coding. Its total mass is $X_\varnothing$, and, outside the countable union of grid boundaries,

$$\lambda_T(Q_u) = \mathbf{1}_{\{u\in T\}} m^{-n} X_u \quad (|u| = n). \tag{I.11.90}$$

We now justify both the boundary qualification and the Euclidean local dimension. For every symbolic generation-n cylinder,

$$\mathbb{E}\big(\mathbf{1}_{\{u\in T\}} m^{-n} X_u\big) = p^n m^{-n} = b^{-n}. \tag{I.11.91}$$

Apply (I.11.91) first on the symbolic space: the expected unnormalized branch measure agrees with the uniform b-ary probability on every cylinder, hence on all symbolic Borel sets by a monotone-class argument. That uniform probability pushes forward under π_M to d-dimensional Lebesgue measure on $[0,1]^d$. Taking preimages of Euclidean Borel sets therefore gives, without assuming anything yet about the random measure's boundary masses,

$$\mathbb{E}\lambda_T(A) = \mathscr{L}^d(A) \quad \text{for every Borel } A \subset [0,1]^d. \tag{I.11.92}$$

Every grid boundary has Lebesgue measure zero, so its nonnegative λ_T-mass has expectation zero. The countable union of all such boundaries therefore has λ_T-mass zero almost surely. This proves (I.11.90) for almost every point of λ_T.

Define the annealed probability

$$\mathbb{Q}_P(B) = \mathbb{E}\int \mathbf{1}_B(\omega, x)\, d\lambda_T(x). \tag{I.11.93}$$

Its x-marginal is Lebesgue measure by (I.11.92). Apply the size-biased calculation (I.11.55)–(I.11.57) to the generation-n cube $Q_n(x) = Q_{u_n}$ selected by x. It gives

$$\frac{1}{n}\log X_{u_n} \longrightarrow 0 \quad \mathbb{Q}_P\text{-almost surely.} \tag{I.11.94}$$

Fix $C \geq 1$. Let $\mathcal{N}_n(u)$ comprise the generation-n cubes meeting the CM^{-n}-neighborhood of Q_u. Its cardinality is bounded by $J = J(C, d)$. For $t > 0$, integrate over the selected cube and use a union bound:

$$\begin{aligned} &\mathbb{Q}_P\big(\max_{v\in\mathcal{N}_n(u_n)} X_v > t\big) \\ &\quad\leq \sum_{|u|=n}\sum_{v\in\mathcal{N}_n(u)} \mathbb{E}\Big(\mathbf{1}_{\{u\in T\}} m^{-n} X_u \mathbf{1}_{\{X_v>t\}}\Big). \end{aligned} \tag{I.11.95}$$

For $v = u$, the sum of the expectations is $\mathbb{E}(X_\varnothing \mathbf{1}_{\{X_\varnothing>t\}}) \leq t^{-1}\mathbb{E}X_\varnothing^2$. For $v \neq u$, the descendant variables X_u, X_v are independent of one another and of their ancestral activation events, so each summand is at most

$$\mathbb{E}\big(\mathbf{1}_{\{u\in T\}} m^{-n} X_u\big)\mathbb{P}(X_v > t) \leq b^{-n} t^{-1}. \tag{I.11.96}$$

There are at most Jb^n ordered neighboring pairs. Hence

$$\mathbb{Q}_P\big(\max_{v\in\mathcal{N}_n(u_n)} X_v > t\big) \leq \frac{\mathbb{E}X_\varnothing^2 + J}{t}. \tag{I.11.97}$$

Set $t = e^{\varepsilon n}$ and use Borel–Cantelli. For every $\varepsilon > 0$, eventually every neighboring descendant mass in (I.11.97) is at most $e^{\varepsilon n}$, for $\mathbb{Q}_P$-almost every sampled point. A ball of radius CM^{-n} meets at most J level-n cubes, so (I.11.90) gives

$$\lambda_T(B(x, CM^{-n})) \leq Jm^{-n}e^{\varepsilon n} \quad \text{eventually.} \tag{I.11.98}$$

On the other hand, $B(x, \sqrt{d}\, M^{-n})$ contains $Q_n(x)$ up to its zero-mass boundary, and (I.11.90), (I.11.94) give

$$\lambda_T(B(x, \sqrt{d}\, M^{-n})) \geq m^{-n}X_{u_n} = m^{-n}e^{o(n)}. \tag{I.11.99}$$

The two inequalities, followed by $\varepsilon \downarrow 0$, prove

$$\lim_{r\downarrow 0}\frac{\log \lambda_T(B(x,r))}{\log r} = \frac{\log m}{\log M} = s \quad \mathbb{Q}_P\text{-almost surely.} \tag{I.11.100}$$

The passage from the geometric sequence of radii to all $r \downarrow 0$ uses monotonicity between consecutive M-adic scales. If B_ω is the exceptional point set in environment ω, then

$$0 = \mathbb{Q}_P\{(\omega, x) : x \in B_\omega\} = \mathbb{E}\lambda_{T,\omega}(B_\omega). \tag{I.11.100a}$$

Thus $\lambda_{T,\omega}(B_\omega) = 0$ for almost every environment. Dividing λ_T by its positive total mass on survival does not change local dimensions. The normalized probability is supported on K, so T-I.4.04 gives $\dim_H K \geq s$. Together with (I.11.80) this proves the Hausdorff part of (I.11.76), and the proof is complete. □

I.11.6 Bounded scalar Mandelbrot cascades

Continue with $M, d, b = M^d$, but now use the full b-ary cube tree. The level cubes form the standard half-open M-adic partition, with only the exterior faces of $[0, 1]^d$ closed. Let $W \geq 0$ be bounded and satisfy

$$\mathbb{E}W = 1. \tag{I.11.101}$$

Attach independent copies W_u to all non-root words. On a level-n cube Q_u, let μ_n have constant density

$$D_n(x) = \prod_{j=1}^{n} W_{u|_j}, \qquad x \in Q_u. \tag{I.11.102}$$

Equivalently,

$$\mu_n(Q_u) = b^{-n} \prod_{j=1}^{n} W_{u|_j}. \tag{I.11.103}$$

Write $Y_n = \mu_n([0, 1]^d)$, and let $\mathscr{F}_n$ be generated by weights through generation n.

Lemma — L-I.11.05 — L^2 cascade convergence and non-degeneracy [F]

Assume

$$\mathbb{E}W^2 < b. \tag{I.11.104}$$

Then Y_n converges almost surely and in L^2 to Y, where

$$\mathbb{E}Y = 1, \qquad \mathbb{E}Y^2 = \frac{b-1}{b - \mathbb{E}W^2}. \tag{I.11.105}$$

On one probability-one event, μ_n converges weakly to a finite Borel measure μ and $\mu([0, 1]^d) = Y$. If $p_+ = \mathbb{P}(W > 0)$, then $bp_+ > 1$, and

$$\{Y > 0\} = \{\text{the positive-weight tree survives}\} \quad \text{modulo a null set}. \tag{I.11.106}$$

Thus the cascade is non-degenerate in the precise sense $\mathbb{P}(Y > 0) > 0$.
This is the bounded scalar L^2 specialization of Kahane–Peyrière, “Sur certaines martingales de Benoit Mandelbrot,” Theorem 1, local KP76-MC PDF p. 5 / source p. 4. The proof is complete below; the source theorem is not imported.

Proof: scalar martingales and the exact second moment

For continuous f, conditional expectation of the next-generation density gives

$$\mathbb{E}(D_{n+1}(x) \mid \mathscr{F}_n) = D_n(x) \quad \text{for Lebesgue-a.e. } x. \tag{I.11.107}$$

Conditional Fubini therefore yields

$$\mathbb{E}\Big(\int f\, d\mu_{n+1} \Big| \mathscr{F}_n\Big) = \int f\, d\mu_n. \tag{I.11.108}$$

For finite Borel measures on $[0,1]^d$, weak convergence $\sigma_n \Rightarrow \sigma$ means

$$\int f\, d\sigma_n \longrightarrow \int f\, d\sigma \quad \text{for every } f \in C([0,1]^d). \tag{I.11.108a}$$

In particular Y_n is a nonnegative martingale of mean one.

Let $Y_n^{(1)}, \dots, Y_n^{(b)}$ be the total masses of independent depth-n descendant cascades. The first-level decomposition is

$$Y_{n+1} = \frac{1}{b} \sum_{i=1}^{b} W_i Y_n^{(i)}. \tag{I.11.109}$$

All $W_i Y_n^{(i)}$ are independent and have mean one. Squaring gives

$$\mathbb{E}Y_{n+1}^2 = \frac{\mathbb{E}W^2}{b} \mathbb{E}Y_n^2 + \frac{b-1}{b}. \tag{I.11.110}$$

Put $a = \mathbb{E}W^2/b < 1$. Since $Y_0 = 1$,

$$\mathbb{E}Y_n^2 = a^n + \frac{b-1}{b} \sum_{j=0}^{n-1} a^j \longrightarrow \frac{b-1}{b - \mathbb{E}W^2}. \tag{I.11.111}$$

Lemma 1.1 proves almost-sure and L^2 convergence. The first formula in (I.11.105) follows from L^2 convergence, and the second from (I.11.111).

Proof: simultaneous construction of the weak limit

Choose a countable rational vector space $\mathscr{D} \subset C([0,1]^d)$ that contains 1 and is uniformly dense. For $f \in \mathscr{D}$, (I.11.108) and

$$\left| \int f\, d\mu_n \right| \le \|f\|_\infty Y_n \tag{I.11.112}$$

show that the martingale $\int f\, d\mu_n$ is L^2 bounded. Lemma 1.1 and countability give a common probability-one event on which all these integrals converge and $Y_n \to Y$. On this event $C_\omega = \sup_n Y_n < \infty$, and the rule

$$L_\omega(f) = \lim_{n\to\infty} \int f\, d\mu_n \quad (f \in \mathscr{D}) \tag{I.11.113}$$

satisfies $|L_\omega(f)| \le C_\omega \|f\|_\infty$. It extends uniquely by uniform density to a bounded real-linear functional on all of $C([0,1]^d)$.

If $g \ge 0$ is continuous and $\|f - g\|_\infty < \varepsilon$ with $f \in \mathscr{D}$, then

$$\int f \, d\mu_n \ge -\varepsilon Y_n. \tag{I.11.114}$$

First let $n \to \infty$, then $\varepsilon \downarrow 0$, to obtain $L_\omega(g) \ge 0$. Riesz representation, an allowed A-RA result, gives a unique finite Borel measure μ_ω with

$$L_\omega(g) = \int g \, d\mu_\omega \quad (g \in C([0,1]^d)). \tag{I.11.115}$$

Uniform approximation by $\mathscr{D}$, together with $\sup_n Y_n < \infty$, shows that (I.11.115) is the weak limit of the entire sequence μ_n, not merely a subsequence. Taking $g = 1$ gives $\mu_\omega([0,1]^d) = Y$.

Define $\mu_\omega = 0$ on the exceptional environment event. The maps $\omega \mapsto \int g \, d\mu_\omega$ are measurable for continuous g by their construction as limits. For a proper relatively open set U in the cube, the continuous functions $g_j(x) = \min\{1, j \operatorname{dist}(x, [0,1]^d \setminus U)\}$ increase to $\mathbf{1}_U$. Hence $\omega \mapsto \mu_\omega(U)$ is measurable; the whole cube is handled by Y. Complements and increasing disjoint unions extend measurability to all Borel sets by the monotone-class theorem. Thus the expected measure and the product-space integrals below are well-defined countably additive measures, by Tonelli.

For later use, this construction also determines the annealed mean measure. For every continuous g, Fubini and (I.11.101) give

$$\mathbb{E} \int g \, d\mu_n = \int g \, d\mathscr{L}^d. \tag{I.11.116}$$

For each fixed real continuous g, (I.11.108), (I.11.112) and Lemma 1.1 give convergence in L^2; its almost-sure limit is the integral against μ already constructed. Cauchy–Schwarz therefore permits passage of expectations to the limit and gives

$$\mathbb{E} \int g \, d\mu = \int g \, d\mathscr{L}^d. \tag{I.11.117}$$

Uniqueness in Riesz representation implies

$$\mathbb{E}\mu(A) = \mathscr{L}^d(A) \quad \text{for every Borel } A \subset [0,1]^d. \tag{I.11.118}$$

In particular, the countable union $\mathscr{B}_M$ of all M-adic cube boundaries satisfies

$$\mu(\mathscr{B}_M) = 0 \quad \text{almost surely.} \tag{I.11.119}$$

Construct descendant limits Y_u simultaneously for all words. For a level-n cube, weak convergence gives convergence of its masses because its boundary has zero limit mass. Here is the required implication without an extra probability theorem: continuous distance approximations from below to the indicator of a relatively open set give $\mu(U) \le \liminf_j \mu_j(U)$; taking complements, using convergence of total mass, gives $\limsup_j \mu_j(F) \le \mu(F)$ for relatively closed F. Sandwich Q_u between its relative interior and closure; their limit masses agree by (I.11.119). The finite descendant identity thus gives

$$\mu(Q_u) = b^{-n} \left(\prod_{j=1}^{n} W_{u|_j} \right) Y_u. \tag{I.11.120}$$

Proof: positivity and the positive-weight tree

Cauchy–Schwarz on $\{W > 0\}$ gives

$$1 = (\mathbb{E}W)^2 \le p_+ \mathbb{E}W^2, \qquad bp_+ \ge \frac{b}{\mathbb{E}W^2} > 1. \tag{I.11.121}$$

The positive-weight tree is therefore a supercritical Galton–Watson tree with $\mathrm{Bin}(b, p_+)$ offspring.

Passing to the limit in (I.11.109) gives the smoothing identity

$$Y = \frac{1}{b} \sum_{i=1}^{b} W_i Y_i, \tag{I.11.122}$$

where the pairs (W_i, Y_i) are independent and Y_i has the law of Y. If $r = \mathbb{P}(Y = 0)$, nonnegativity in (I.11.122) yields

$$r = (1 - p_+ + p_+ r)^b. \tag{I.11.123}$$

The right side is the offspring generating function of the positive-weight tree. Since $\mathbb{E}Y = 1$, $r < 1$. By the fixed-point proof in L-I.11.01, r equals the extinction probability of that tree. Extinction plainly forces $Y = 0$, so the two events coincide modulo a null set. This argument applies in every descendant environment; intersecting the resulting countably many probability-one events makes the equivalence simultaneous for all words u. This proves (I.11.106) and completes the proof. □

I.11.7 The Peyrière probability and spine laws

The Peyrière probability is the annealed measure on environments and sampled points defined by

$$\mathbb{Q}(B) = \mathbb{E} \int_{[0,1]^d} \mathbf{1}_B(\omega, x)\, d\mu_\omega(x). \tag{I.11.124}$$

It has total mass $\mathbb{E}Y = 1$, and its x-marginal is Lebesgue measure by (I.11.118). The half-open cube convention gives every x a unique sequence of cube labels $I_1(x), I_2(x), \ldots$. The size-biased law of W is

$$\mathbb{P}^*(W \in C) = \mathbb{E}\big(W \mathbf{1}_{\{W \in C\}}\big). \tag{I.11.125}$$

This is a probability because $\mathbb{E}W = 1$.

Lemma — L-I.11.06 — Peyrière spine law [F]

Under (I.11.101) and (I.11.104), the spine weights

$$W_{I_1},\ W_{I_1I_2}, \dots \tag{I.11.126}$$

are i.i.d. under $\mathbb{Q}$ with law $\mathbb{P}^*$. With $0 \log 0 = 0$,

$$\frac{1}{n}\sum_{j=1}^{n} \log W_{I_1\cdots I_j} \longrightarrow \mathbb{E}(W \log W) \quad \mathbb{Q}\text{-almost surely.} \tag{I.11.127}$$

The descendant total masses on the spine satisfy

$$\frac{1}{n}\log Y_{I_1\cdots I_n} \longrightarrow 0 \quad \mathbb{Q}\text{-almost surely.} \tag{I.11.128}$$

Proof: finite-dimensional spine distributions

Let $\varphi_1, \dots, \varphi_n$ be bounded Borel functions. Partition the integral in (I.11.124) over level-n cubes and use (I.11.120):

$$\begin{aligned} &\mathbb{E}\int \prod_{j=1}^{n} \varphi_j(W_{I_1\cdots I_j})\, d\mu \\ &\quad = \sum_{|u|=n} b^{-n}\mathbb{E}\Bigg[\Bigg(\prod_{j=1}^{n} W_{u|_j}\varphi_j(W_{u|_j})\Bigg) Y_u\Bigg]. \end{aligned} \tag{I.11.129}$$

The path weights are independent, Y_u is independent of them, and $\mathbb{E}Y_u = 1$. There are b^n words, so

$$\mathbb{E}\int \prod_{j=1}^{n} \varphi_j(W_{I_1\cdots I_j})\, d\mu = \prod_{j=1}^{n} \mathbb{E}(W\varphi_j(W)). \tag{I.11.130}$$

This proves the asserted i.i.d. size-biased law. It assigns no mass to $W = 0$. If $W \le C$, then

$$\mathbb{E}^*|\log W| = \mathbb{E}(W|\log W|) < \infty, \tag{I.11.131}$$

because $w|\log w|$ is bounded on $[0, C]$ after setting its value at zero to zero. The strong law applied under $\mathbb{P}^*$ now gives (I.11.127).

Proof: the terminal descendant mass

For a Borel set $C \subset [0, \infty)$, another use of (I.11.120) gives

$$\begin{aligned} \mathbb{Q}(Y_{I_1\cdots I_n} \in C) &= \sum_{|u|=n} b^{-n}\mathbb{E}\Bigg[\Bigg(\prod_{j=1}^{n} W_{u|_j}\Bigg) Y_u \mathbf{1}_{\{Y_u \in C\}}\Bigg] \\ &= \mathbb{E}\big(Y\mathbf{1}_{\{Y\in C\}}\big). \end{aligned} \tag{I.11.132}$$

Thus $Y_{I_1\cdots I_n}$ has the Y-size-biased law for every n. For $\varepsilon > 0$,

$$\begin{aligned} \mathbb{Q}(Y_{I_1\cdots I_n} < e^{-\varepsilon n}) &\le e^{-\varepsilon n}, \\ \mathbb{Q}(Y_{I_1\cdots I_n} > e^{\varepsilon n}) &\le e^{-\varepsilon n}\mathbb{E}Y^2. \end{aligned} \tag{I.11.133}$$

Both series are summable. Borel–Cantelli for rational $\varepsilon > 0$ gives (I.11.128). Equation (I.11.132) also shows that the selected $Y_{I_1\cdots I_n}$ is positive $\mathbb{Q}$-almost surely, so all logarithms in (I.11.128) are finite. □

I.11.8 Euclidean exact dimensionality of the cascade

Set

$$D = \frac{\log b - \mathbb{E}(W \log W)}{\log M} = d - \frac{\mathbb{E}(W \log W)}{\log M}, \tag{I.11.134}$$

with $0 \log 0 = 0$.

Theorem — T-I.11.07 — Exact dimension of a bounded scalar cascade [F]

Under (I.11.101) and (I.11.104), almost surely on $\{Y > 0\}$, the normalized probability

$$\overline{\mu} = Y^{-1}\mu \tag{I.11.135}$$

is exact dimensional and

$$\dim_{\mathrm{H}} \overline{\mu} = D. \tag{I.11.136}$$

Moreover $0 < D \le d$.
This is the bounded-L^2, d-dimensional cube analogue of Kahane–Peyrière, KP76-MC, Theorem 4, local PDF p. 5 / source p. 4; its proof architecture occupies local PDF pp. 14–18. The source statement uses one-dimensional interval exponents. The proof below supplies the d-dimensional construction and the extra passage to Euclidean balls; neither is inferred merely by citing that interval statement.

Proof: the M-adic cylinder exponent

For the selected generation-n cube $Q_n(x) = Q_{I_1\cdots I_n}$, (I.11.120) gives

$$\log \mu(Q_n(x)) = -n \log b + \sum_{j=1}^{n} \log W_{I_1\cdots I_j} + \log Y_{I_1\cdots I_n}. \tag{I.11.137}$$

The selected weights and terminal mass are positive $\mathbb{Q}$-almost surely. By L-I.11.06,

$$\frac{\log \mu(Q_n(x))}{-n \log M} \longrightarrow D \quad \mathbb{Q}\text{-almost surely.} \tag{I.11.138}$$

The zero-boundary statement (I.11.119) ensures that the half-open and closed versions of $Q_n(x)$ have equal μ-mass at every level, outside one common null set.

Proof: sublinear digit carries

We need a uniform comparison with all neighboring cubes; fixed-grid cylinders alone do not imply Euclidean local dimensions. Index a generation-n cube by d integers in $\{0, \dots, M^n - 1\}$. Fix $C \geq 1$. Let $R_n(x)$ be the least nonnegative integer with the following property: every level-n cube meeting $B(x, CM^{-n})$ has a symbolic common prefix with $Q_n(x)$ of length at least $n - R_n(x) - 1$.

There is a constant $A = A(C, d, M)$ such that, for every $1 \leq \ell \leq n$,

$$\mathscr{L}^d\{x : R_n(x) \geq \ell\} \leq AM^{-\ell}. \tag{I.11.139}$$

Here is a geometric count with explicit constants. Ignore the grid boundaries themselves, a Lebesgue-null set. If $R_n(x) \geq \ell$, a cube meeting $B(x, CM^{-n})$ lies in a different generation-$(n-\ell)$ ancestor from $Q_n(x)$. A segment from x to a point of that cube in the ball crosses a generation-$(n-\ell)$ grid hyperplane. Thus, for some coordinate j and integer h,

$$|x_j - hM^{-(n-\ell)}| \leq CM^{-n}. \tag{I.11.139a}$$

There are at most $M^{n-\ell} + 1$ such hyperplanes in each coordinate. Each strip inside the unit cube has volume at most $2CM^{-n}$. Subadditivity therefore bounds the exceptional volume by $2dC(M^{n-\ell}+1)M^{-n} \leq 4dCM^{-\ell}$. This proves (I.11.139), for example with $A = 4dC$. When $\ell = n$, the exceptional event is empty; the same bound remains valid.

The x-marginal of $\mathbb{Q}$ is Lebesgue measure. Apply (I.11.139) with $\ell = \lceil \varepsilon n \rceil$, use Borel–Cantelli, and then intersect over rational $\varepsilon > 0$. We obtain

$$\frac{R_n(x)}{n} \longrightarrow 0 \quad \mathbb{Q}\text{-almost surely.} \tag{I.11.140}$$

Thus every level-n cube Q_v meeting $B(x, CM^{-n})$ shares with $Q_n(x)$ a prefix w of length

$$k = k(n, v) \geq n - o(n). \tag{I.11.141}$$

Proof: neighboring terminal masses

Let $\mathcal{N}_n(u)$ be all level-n cubes meeting the CM^{-n}-neighborhood of Q_u. Its size is at most a constant $J = J(C, d)$. For $t > 0$, partition the Peyrière integral over the selected cube:

$$\begin{aligned} &\mathbb{Q}\Big(\max_{v \in \mathcal{N}_n(I|_n)} Y_v > t\Big) \\ &\quad \leq \sum_{|u|=n} \sum_{v \in \mathcal{N}_n(u)} \mathbb{E}\big(\mu(Q_u) \mathbf{1}_{\{Y_v > t\}}\big). \end{aligned} \tag{I.11.142}$$

For $v = u$, summing (I.11.120) over u, using independence and $\mathbb{E}W = 1$, gives

$$\sum_{|u|=n} \mathbb{E}\big(\mu(Q_u) \mathbf{1}_{\{Y_u > t\}}\big) = \mathbb{E}\big(Y \mathbf{1}_{\{Y > t\}}\big) \leq t^{-1} \mathbb{E}Y^2. \tag{I.11.143}$$

For $v \neq u$, the descendant variable Y_v is independent of $\mu(Q_u)$; therefore

$$\mathbb{E}\big(\mu(Q_u) \mathbf{1}_{\{Y_v > t\}}\big) = b^{-n} \mathbb{P}(Y > t) \leq b^{-n} t^{-1}. \tag{I.11.144}$$

There are at most Jb^n ordered neighboring pairs. Consequently,

$$\mathbb{Q}\big(\max_{v\in\mathcal{N}_n(I|_n)} Y_v > t\big) \le \frac{\mathbb{E}Y^2 + J}{t}. \tag{I.11.145}$$

Taking $t = e^{\varepsilon n}$, applying Borel–Cantelli, and intersecting over rational $\varepsilon > 0$ yields

$$\max_{v\in\mathcal{N}_n(I|_n)} Y_v \le e^{o(n)} \quad \mathbb{Q}\text{-almost surely.} \tag{I.11.146}$$

Proof: upper mass bound for Euclidean balls

Let $C_W = \max\{1, \|W\|_\infty/b\}$. If Q_v is a neighboring level-n cube and w is its common prefix with $Q_n(x)$, then (I.11.120) and boundedness of W give

$$\begin{aligned}\mu(Q_v) &= \left(b^{-k}\prod_{j=1}^{k} W_{I_1\cdots I_j}\right)\left(b^{-(n-k)}\prod_{j=k+1}^{n} W_{v|_j}\right)Y_v\\ &\le \left(b^{-k}\prod_{j=1}^{k} W_{I_1\cdots I_j}\right)C_W^{n-k}Y_v.\end{aligned} \tag{I.11.147}$$

The strong law in L-I.11.06 gives, along every subsequence $k\to\infty$,

$$\frac{1}{k}\log\left(b^{-k}\prod_{j=1}^{k} W_{I_1\cdots I_j}\right) \longrightarrow -\log b + \mathbb{E}(W\log W) = -D\log M. \tag{I.11.148}$$

Equations (I.11.141), (I.11.146)–(I.11.148) show, uniformly over the bounded family of neighbors, that

$$\mu(Q_v) \le \exp\big(-Dn\log M + o(n)\big). \tag{I.11.149}$$

A ball $B(x, CM^{-n})$ meets at most J such cubes. Hence

$$\mu(B(x, CM^{-n})) \le \exp\big(-Dn\log M + o(n)\big). \tag{I.11.150}$$

Proof: lower mass bound and disintegration

The ball $B(x, \sqrt{d}\,M^{-n})$ contains $Q_n(x)$, modulo its zero-mass boundary. Equation (I.11.138) therefore gives

$$\mu(B(x, \sqrt{d}\,M^{-n})) \ge \exp\big(-Dn\log M + o(n)\big). \tag{I.11.151}$$

Use (I.11.150) with any fixed C large enough for one scale and (I.11.151), then interpolate arbitrary r between consecutive M-adic scales by monotonicity. The negative sign of $\log r$ reverses the inequality coming from (I.11.150), and the two bounds yield

$$\lim_{r\downarrow 0}\frac{\log\mu(B(x,r))}{\log r} = D \quad \mathbb{Q}\text{-almost surely.} \tag{I.11.152}$$

If B_ω is the exceptional set of points in environment ω, then

$$0 = \mathbb{Q}\{(\omega, x) : x \in B_\omega\} = \mathbb{E}\mu_\omega(B_\omega). \tag{I.11.153}$$

Thus $\mu_\omega(B_\omega) = 0$ for almost every environment. On $\{Y > 0\}$, normalization by Y changes $\log \mu(B(x, r))$ by the constant $-\log Y$, so $\overline{\mu}$ is exact dimensional with dimension D.

Finally, convexity of $w \mapsto w \log w$ and $\mathbb{E}W = 1$ gives $\mathbb{E}(W \log W) \geq 0$. Under the size-biased law, Jensen gives

$$\mathbb{E}(W \log W) = \mathbb{E}^*(\log W) \leq \log \mathbb{E}^* W = \log \mathbb{E}W^2 < \log b. \tag{I.11.154}$$

Therefore $0 < D \leq d$, completing the proof. □

I.11.9 The deletion cascade and natural percolation measure

Let $b^{-1} < p \leq 1$ and take

$$W = \begin{cases} p^{-1}, & \text{with probability } p, \\ 0, & \text{with probability } 1 - p. \end{cases} \tag{I.11.155}$$

Then

$$\mathbb{E}W = 1, \qquad \mathbb{E}W^2 = p^{-1} < b. \tag{I.11.156}$$

Corollary — C-I.11.08 — Deletion cascade is the natural percolation measure [F]

The positive-weight tree for (I.11.155) is homogeneous M-adic fractal percolation with parameter p. Almost surely on survival,

$$\operatorname{supp} \overline{\mu} = K, \qquad \overline{\mu} = (\pi_M)_* \nu_T, \tag{I.11.157}$$

where ν_T is the natural weighted branching probability of L-I.11.02. Moreover,

$$\dim_{\mathrm{H}} \overline{\mu} = \dim_{\mathrm{H}} K = d + \frac{\log p}{\log M}. \tag{I.11.158}$$

Proof

A child has positive weight exactly when it is retained, independently with probability p. Thus the positive tree and the percolation retained tree coincide. If Z_n is their generation size and $m = pb$, then the cascade mass and the weighted branching sum are the same random variable:

$$Y_n = b^{-n} p^{-n} Z_n = m^{-n} Z_n = X_n. \tag{I.11.159}$$

The same identity holds in every descendant tree, so $Y_u = X_u$ simultaneously. For an active word u of length n, (I.11.120) gives

$$\mu(Q_u) = m^{-n} Y_u = m^{-n} X_u. \tag{I.11.160}$$

After division by $Y = X_\varnothing$, this is precisely the cylinder formula (I.11.40). The grid-boundary mass is zero by (I.11.119), proving the measure identity in (I.11.157).

For every n, all later approximating measures vanish outside the closed set K_n. If g is continuous with compact support in K_n^c, then $\int g\, d\mu_j = 0$ for every $j \geq n$, so weak convergence gives $\int g\, d\mu = 0$. Inner regularity and continuous approximation of compact subsets of the open set K_n^c, both A-RA facts, yield $\mu(K_n^c) = 0$. Hence $\operatorname{supp}\mu \subset K = \bigcap_n K_n$. Conversely, take $x \in K$. Finite branching supplies an infinite retained branch whose nested closed cubes contain x. On the common probability-one event from L-I.11.05, the descendant cascade mass is positive at every vertex whose positive tree survives. Every prefix u of this branch therefore has

$$\mu(Q_u) = m^{-|u|} Y_u > 0. \tag{I.11.161}$$

The diameters of these cubes tend to zero. Every neighborhood of x contains one of them and consequently has positive μ-mass. Thus $x \in \operatorname{supp}\mu$, proving the support identity.

Finally,

$$\mathbb{E}(W \log W) = \log(p^{-1}), \tag{I.11.162}$$

so T-I.11.07 gives

$$\dim_{\mathrm{H}} \overline{\mu} = d - \frac{\log(p^{-1})}{\log M} = d + \frac{\log p}{\log M}. \tag{I.11.163}$$

C-I.11.04 gives the same dimension for K, completing the proof. □

I.11.10 Precisely sourced terminal previews

The next two results are not proved in this chapter. They are terminal: nothing later in the F subgraph may cite them as a premise.

Guided result — G-I.11.09 — Rams–Simon all-direction projection theorem [G]

Currency: Published. **Source:** Rams and Simon [RS14, Theorem 2 and Corollary 3, p. 5]. **Omitted proof blocks:** the uniform line-slice counting estimate, its large-deviation argument, the non-coordinate projection deduction, and the coordinate-direction results combined in Corollary 3.

Let $K \subset [0,1]^2$ be planar homogeneous M-adic fractal percolation with $p > M^{-2}$. Almost surely conditioned on $K \neq \varnothing$, one common event works for every line $\ell \subset \mathbb{R}^2$, and on that event

$$\dim_{\mathrm{H}} P_\ell K = \min\{1, \dim_{\mathrm{H}} K\}. \tag{I.11.164}$$

Here P_ℓ is orthogonal projection onto ℓ. The order of quantifiers is

$$\mathbb{P}(\forall \ell,\ \dim_{\mathrm{H}} P_\ell K = \min\{1, \dim_{\mathrm{H}} K\} \,|\, K \neq \varnothing) = 1. \tag{I.11.165}$$

This node has no outgoing premise edge.

Guided proof architecture, not a proof

For non-coordinate directions, the source bounds uniformly the number of retained level squares met by each line, using a slicing estimate and a large-deviation inequality. A cover of the projected set with anomalously small dimension would contradict those uniform counts together with the known growth of retained squares. The source then combines the non-coordinate theorem with separately established coordinate-direction results to obtain Corollary 3. Since those quantitative slice and coordinate blocks are omitted here, this paragraph is explanatory only and does not change the G label.

Guided result — G-I.11.10 — Rams–Simon interval theorem for independent sums [G]

Currency: Published. **Source:** Rams and Simon [RS14, Theorem 17, p. 15]. **Omitted proof blocks:** the probability-thinning reduction, the hyperplane/cube intersection estimate, the bounded-dependence coloring argument, and the multiscale interval-production step in Section 4 of the source.
Fix $M \geq 2$ and $k \geq 2$. Let $E^{(1)}, \ldots, E^{(k)} \subset [0,1]$ be independent one-dimensional M-adic fractal percolations with parameters

$$p_j > M^{-1} \quad (1 \leq j \leq k), \qquad \prod_{j=1}^{k} p_j > M^{-k+1}. \tag{I.11.166}$$

For each fixed coefficient vector $a = (a_1, \ldots, a_k) \in (\mathbb{R} \setminus \{0\})^k$,

$$a_1 E^{(1)} + \cdots + a_k E^{(k)} = \left\{ \sum_{j=1}^{k} a_j x_j \ :\ x_j \in E^{(j)} \text{ for every } j \right\}. \tag{I.11.166a}$$

For this set,

$$\mathbb{P}\left(a_1 E^{(1)} + \cdots + a_k E^{(k)} \text{ contains a non-degenerate interval} \,\middle|\, \bigcap_{j=1}^{k} \{E^{(j)} \neq \varnothing\} \right) = 1. \tag{I.11.167}$$

The source does not assert in this theorem that one probability-one event works simultaneously for every uncountable choice of a, and neither does the chapter. This node has no outgoing premise edge.

Guided proof architecture, not a proof

The product $E^{(1)} \times \cdots \times E^{(k)}$ is a dependent selection of M-adic cubes in $[0,1]^k$, and the algebraic sum is its projection in the direction a. The product condition in (I.11.166) places the expected dimension above one. Rams–Simon first thin the probabilities to a convenient strict regime, bound how many dependent product cubes a hyperplane can meet, color the dependency graph into controlled independent classes, and iterate a projection-survival estimate to create an interval. Those four blocks are not reproduced here, so this remains a G result.

I.11.11 Terminal boundary of the chapter

Research boundary — Q-I.11.11 — Deferred random-geometry regimes [Q]

The following questions are outside the proved kernel. We specify the terms so this boundary is not a list of unexplained names.

(1) For planar percolation let $\mathscr{C}$ be the event that K contains a connected compact subset meeting both vertical sides of the unit square. The crossing **connectivity threshold** is $p_c = \inf\{p \in [0,1] : \mathbb{P}_p(\mathscr{C}) > 0\}$, where $\mathbb{P}_p$ denotes the percolation law with parameter p. No value of p_c, or behavior at that value, is asserted here.

(2) A projection has a **Hölder density** if $(P_\ell)_*\mu = f_\ell \mathscr{L}^1|_\ell$ with $|f_\ell(x) - f_\ell(y)| \leq C|x-y|^\beta$ for some $C < \infty$ and $0 < \beta \leq 1$. No claim of such densities, much less one event valid for all directions, follows from G-I.11.09.

(3) Given nonnegative random densities f_n and a prescribed family of finite Borel measures $(\eta_t)_{t\in\Gamma}$, intersection averages mean $Y_n(t) = \int f_n \, d\eta_t$. No uniform-in-$t$ convergence or regularity theorem for these averages is proved here.

(4) For a compactly supported probability μ on $\mathbb{R}^d$, write $\hat{\mu}(\xi) = \int e^{-2\pi i\langle \xi, x\rangle} \, d\mu(x)$. In this discussion, **Salem measure** means that for every $0 < \sigma < \dim_{\mathrm{H}} \mu$ there is $C_\sigma < \infty$ such that

$$|\hat{\mu}(\xi)| \leq C_\sigma (1+|\xi|)^{-\sigma/2} \quad (\xi \in \mathbb{R}^d).$$

No almost-sure assertion of this property is made for our random measures.

(5) A **spatially independent martingale**, in the convention of Shmerkin and Suomala [SS18, Definition 3.1, p. 10], is a sequence of nonnegative random Borel densities f_n with an increasing filtration $(\mathscr{F}_n)$, satisfying: f_0 is deterministic, bounded and compactly supported; f_n is adapted and $\mathbb{E}(f_{n+1}(x) \mid \mathscr{F}_n) = f_n(x)$; and there is a deterministic $C < \infty$ such that

$$f_{n+1}(x) \leq C f_n(x) \quad (n \geq 0, \ x \in \mathbb{R}^d).$$

In addition, for any finite family of generation-$(n+1)$ dyadic cubes with pairwise set distances at least $C2^{-n}$, their next-density restrictions are conditionally independent given $\mathscr{F}_n$. Precisely, if U_j is a bounded random variable measurable with respect to the restriction on the j-th cube, then

$$\mathbb{E}\left(\prod_{j=1}^{k} U_j \,\middle|\, \mathscr{F}_n\right) = \prod_{j=1}^{k} \mathbb{E}(U_j \mid \mathscr{F}_n).$$

This definition imports none of that framework's convergence or intersection theorems.

(6) A **random affine code tree** here means a random rooted tree carrying, at each vertex u, a finite list of contracting affine maps $S_{u,i}(x) = A_{u,i}x + b_{u,i}$. Along a branch one composes the maps in ancestral order. When contraction bounds are uniform and a common compact set is mapped into itself, nested images define the associated limit set. No dimension or pressure formula for this broader model is asserted.
For the later intersection and Fourier owner, the frozen boundary source is Shmerkin and Suomala [SS18, Theorem 4.1, pp. 18–23, Theorem 7.1, pp. 39–40, and Theorem 10.1 with Corollary 10.2, pp. 58–60]. None is asserted or used here. This Q node has no outgoing edge.

I.11.12 How the proofs fit together

The proved result sequence is

$$\begin{gathered}\text{L-I.11.01} \longrightarrow \text{L-I.11.02} \longrightarrow \text{T-I.11.03},\\ \text{L-I.11.01}, \text{L-I.11.02} \longrightarrow \text{C-I.11.04},\\ \text{L-I.11.05} \longrightarrow \text{L-I.11.06} \longrightarrow \text{T-I.11.07} \longrightarrow \text{C-I.11.08}.\end{gathered} \tag{I.11.168}$$

C-I.11.08 also uses the already proved C-I.11.04 only to identify the set dimension. Every other premise is one of D/R-I.1–I.4, D/R-I.7–I.8, the three exact allowed-book probability facts in Section 1, or a locally proved lemma. The two Rams–Simon previews and the final research roadmap are not used in any of these proofs.

I.11.13 Intuition and strategy

The following discussion explains the geometric ideas behind the proofs.

I.11.13.1 One random tree, several geometric models

A Galton–Watson process records how many pieces remain after each generation. Fractal percolation adds a rigid geometric address to every child. A random recursive IFS replaces the common grid ratio by symbol-dependent similarity ratios. A multiplicative cascade keeps the entire address tree but assigns random mass multipliers to its edges.

The common mechanism is a normalized generation sum. Its expectation stays constant because the expected contribution of all children of one vertex is one. The second-moment hypothesis says that two descendants chosen from the same parent do not correlate strongly enough to destroy this normalization.

I.11.13.2 Why positivity is a fixed-point question

A limiting mass is zero precisely when every child contribution is zero. Independence turns that sentence into the same generating-function equation that describes extinction of the positive-child tree. Mean preservation rules out the fixed point one. The smaller fixed point is therefore both the probability of extinction and the probability that the limiting mass vanishes.

This is the elementary core of several much more general non-degeneracy theorems. The chapter keeps the bounded second-moment regime so that the entire argument can be seen without a Kesten–Stigum or general smoothing-transform theorem.

I.11.13.3 The natural random-recursive measure

The Malthusian exponent makes the expected sum of the s-powers of the child ratios equal one. The limiting mass below a cylinder measures how much viable geometry remains in that cylinder. Multiplying its deterministic scale weight by that remaining random mass produces consistent cylinder weights.

A typical branch for this measure is not sampled with the original child law. Children are tilted by their expected s-mass. Under this tilted law the logarithms of contraction ratios are i.i.d., while the terminal subtree mass contributes only a sublinear logarithmic error. That is why the local dimension is exactly s.

I.11.13.4 Why separation makes the first dimension theorem clean

SSC makes the coding map bi-Lipschitz when the symbolic space is equipped with the contraction-weighted ultrametric. A symbolic cylinder is then a ball at the correct geometric scale. The upper Hausdorff bound comes from expected covers; the lower bound comes from the exact-dimensional natural measure.

The grid cubes in fractal percolation touch, so this shortcut is deliberately not used there. The proof instead checks that shared faces cost only a bounded multiplicity and that the natural measure assigns zero mass to all grid boundaries.

I.11.13.5 Retained cubes versus live cubes

Generation n may contain many cubes whose descendants later die. Box dimension is controlled by live cubes: cubes whose descendant trees survive forever. Given the generation-n tree, liveness is an independent Bernoulli thinning with a fixed positive parameter. The thinning changes the count by a constant proportion, not its exponential rate.

This distinction is easy to overlook. Counting all retained cubes proves a covering upper bound, but the lower box bound needs cubes that actually meet the limiting set.

I.11.13.6 Peyrière sampling

The Peyrière law first samples the random measure and then samples a point from it, without normalizing each realization. A path edge is therefore seen in proportion to the mass multiplier carried by that edge. This is ordinary size biasing:

$$\text{original law of } W \quad \longmapsto \quad \text{law with density } W.$$

Under the new law, average logarithmic weight becomes $\mathbb{E}(W \log W)$. Subtracting it from the deterministic volume loss $\log b$ gives the cascade dimension.

I.11.13.7 Why cylinder dimension is not automatically Euclidean dimension

An M-adic neighbor of the cube containing x can differ far back in its address if adding one to a cube index creates a long digit carry. For a Lebesgue-typical point, such carries have sublinear length. The annealed point distribution under the Peyrière law is Lebesgue, so the same conclusion holds for the sampled cascade point.

After the common prefix, only sublinearly many bounded weights remain. Neighboring terminal subtree masses are also subexponential by a size-biased/ordinary tail calculation. Thus a Euclidean ball, although it may meet several M-adic cubes, has the same exponential mass as the selected cylinder.

I.11.13.8 Deletion cascades

For the deletion weight, every retained edge contributes p^{-1} and every deleted edge contributes zero. The cascade mass of a retained generation-n cube is exactly the normalized Galton–Watson descendant mass. Consequently the cascade probability and the natural fractal-percolation probability are the same object, not merely measures with the same dimension.

I.11.13.9 What the two previews add

The proved kernel explains dimensions of the random sets and their natural measures. Rams–Simon goes further in two directions: one realization has the correct projection dimension in every planar direction, and each fixed nontrivial linear combination of sufficiently large independent one-dimensional percolations contains an interval.

These conclusions require uniform slice counts and multiscale dependency control that are not part of the foundational kernel. They remain terminal G results. Shmerkin–Suomala's spatially independent martingales provide a far broader later framework for uniform intersections, densities, and random Fourier decay; that theory belongs to Volume VII.

I.11.14 Exercises

Every exercise uses only the accepted incoming nodes and earlier results in I.11. The G and Q nodes are never used as mathematical premises.

I.11.14.1 Exercise I.11.1 — Fixed points of an offspring generating function

Let ξ be a bounded nonnegative integer-valued random variable and $f(z) = \mathbb{E}z^{\xi}$.

1. Prove that if $\mathbb{E}\xi > 1$, then $f(z) = z$ has exactly two solutions in $[0, 1]$, except that the smaller one may be 0.
2. Prove directly that the smaller solution is $\lim_{n\to\infty} f^n(0)$.

Dependencies: Section 2 and elementary convexity.

I.11.14.2 Exercise I.11.2 — Orthogonality of martingale differences

Let $(M_n, \mathscr{F}_n)$ be an L^2 martingale and $\Delta_n = M_n - M_{n-1}$.

1. Prove that $\Delta_j \perp \Delta_k$ in L^2 for $j \neq k$.
2. Deduce
$$\mathbb{E}|M_n - M_r|^2 = \sum_{k=r+1}^{n} \mathbb{E}|\Delta_k|^2.$$
3. Explain exactly why L^2-boundedness makes (M_n) L^2-Cauchy.

Dependencies: conditional expectation from ARA-KB1.

I.11.14.3 Exercise I.11.3 — A binary Galton–Watson example

Suppose

$$\mathbb{P}(\xi = 0) = 1 - p, \qquad \mathbb{P}(\xi = 2) = p, \qquad p > \tfrac{1}{2}.$$

Compute the extinction probability, the mean and variance, and $\lim_n \mathbb{E}W_n^2$. Verify every answer from L-I.11.01.

Dependencies: L-I.11.01.

I.11.14.4 Exercise I.11.4 — Exact weighted-branching second moment

In Section 3, let

$$c = \Phi(2s), \qquad B = \mathbb{E} \sum_{i \neq j} \mathbf{1}_{\{i,j \in A_\varnothing\}} r_i^s r_j^s.$$

Solve recurrence (I.11.45) exactly and compute $\mathbb{E}X_\varnothing^2$.

Dependencies: L-I.11.02.

I.11.14.5 Exercise I.11.5 — The annealed branch law

Prove from the cylinder formula, without quoting (I.11.53), that under $\mathbb{Q}_T$:

1. the first label has law $\theta_i = \mathbb{P}(i \in A_\varnothing) r_i^s$;
2. conditional on the first n labels, the next label again has law (θ_i); and
3. the descendant limit at the selected generation-n vertex has the $X_\varnothing$-size-biased law.

Dependencies: L-I.11.02.

I.11.14.6 Exercise I.11.6 — Random-recursive upper covers

Let $t > s$. Prove that

$$\sum_{n \geq 1} \sum_{u \in T_n} (\operatorname{diam} S_u(K_{\mathcal{S}}))^t < \infty \quad \text{almost surely.}$$

Deduce $\mathscr{H}^t(K_T) = 0$, explicitly checking that the mesh tends to zero.

Dependencies: T-I.11.03 and Tonelli.

I.11.14.7 Exercise I.11.7 — Live-cube thinning

For supercritical M-adic fractal percolation, let Z_n be the number of retained generation-n cubes and V_n the number with infinite descendants.

1. Prove the conditional law (I.11.82).
2. Starting from conditional Chebyshev, prove $V_n/Z_n \to 1-q$ on survival.
3. Explain why using Z_n alone would not prove the lower box bound.

Dependencies: L-I.11.01 and C-I.11.04.

I.11.14.8 Exercise I.11.8 — Annealed Lebesgue measure

For the unnormalized natural percolation measure λ_T, prove

$$\mathbb{E}\lambda_T = \mathscr{L}^d|_{[0,1]^d}$$

first for the unnormalized symbolic measure on word cylinders, then by pushing forward to all Euclidean Borel sets. Deduce that every M-adic boundary has zero λ_T-mass almost surely.

Dependencies: L-I.11.02 and the monotone-class/Riesz uniqueness facts in A-RA.

I.11.14.9 Exercise I.11.9 — Cascade second moments

For the scalar cascade, derive (I.11.110) from the first-level decomposition and solve it. Show that

$$\sup_n \mathbb{E}Y_n^2 < \infty \quad \Longleftrightarrow \quad \mathbb{E}W^2 < b$$

under the standing hypothesis $b \geq 2$.

Dependencies: L-I.11.05.

I.11.14.10 Exercise I.11.10 — Two size-biased laws

Under the Peyrière law, prove the formulas

$$\mathbb{Q}(W_{I_1} \in C) = \mathbb{E}(W\mathbf{1}_{\{W\in C\}}),$$

and

$$\mathbb{Q}(Y_{I|_n} \in C) = \mathbb{E}(Y\mathbf{1}_{\{Y\in C\}}).$$

Use them to reproduce both Borel–Cantelli bounds in (I.11.133).

Dependencies: L-I.11.05–L-I.11.06.

I.11.14.11 Exercise I.11.11 — A digit-carry estimate

In one dimension, fix an integer $A \geq 1$ and $1 \leq \ell \leq n$. For a uniformly selected $j \in \{0,\dots,M^n-1\}$, prove that the probability that adding some integer a with $|a| \leq A$ changes a digit earlier than the final ℓ base-M digits is at most $C_{A,M}M^{-\ell}$, after excluding indices for which $j + a$ leaves the range. Extend the estimate to d coordinates and deduce (I.11.140).

Dependencies: elementary counting and Borel–Cantelli.

I.11.14.12 Exercise I.11.12 — Neighboring cascade masses

Fill in the independence step in (I.11.144). In particular, prove that for distinct words u, v of the same length, the terminal variable Y_v is independent of $\mu(Q_u)$, even when u and v share a long prefix. Then derive (I.11.145).

Dependencies: the tree-indexed independence in D-I.11.07 and L-I.11.05.

I.11.14.13 Exercise I.11.13 — Deletion cascade identity

For the deletion law (I.11.155), prove at every finite level that the cascade total mass, the normalized Galton–Watson population, and the weighted branching sum are equal. Pass to descendant limits and prove the measure identity in (I.11.157).

Dependencies: L-I.11.02, L-I.11.05, and C-I.11.08.

I.11.14.14 Exercise I.11.14 — Quantifier audit for the two G results

Write the logical forms of (I.11.165) and (I.11.167). Explain why

$$\forall a\ \mathbb{P}(E_a) = 1$$

does not in general imply

$$\mathbb{P}(\forall a\ E_a) = 1$$

when the parameter set is uncountable. Give a concrete probability-space example.

Dependencies: only D-I.11.01 and the exact G statements; no G conclusion is used as a premise.

Chapter I.12

Sceneries and the Unity of the Subject

The last chapter returns to the operation of magnification. Tangent measures record limits of rescaled measures, while scenery distributions record the statistics of many successive views. Symbolic components provide a discrete counterpart. The aim is to make these objects and their normalization precise and to relate the dimension certificates developed earlier in the volume. The final outlook identifies further questions without using later-volume results as proof premises.

The definitions below fix this chapter's conventions. Geometric intuition and proof strategy are discussed separately after the main development; exercises and their solutions provide a second route through the material.

Definitions and notation

Previously accepted dimension, measure, symbolic, IFS, entropy, and dynamical-system terminology retains its meaning. Every term particular to sceneries or CP magnification is defined below.

Definition — I.12.01: Measure topologies and push-forwards

Let $\mathscr{M}(\mathbb{R}^d)$ be the set of nonnegative Radon measures on $\mathbb{R}^d$. A sequence μ_n **converges vaguely** to μ, written $\mu_n \to \mu$ vaguely, if

$$\int f\, d\mu_n \longrightarrow \int f\, d\mu \qquad \text{for every } f \in C_c(\mathbb{R}^d).$$

If K is compact metric, $\mathscr{P}(K)$ is the set of Borel probability measures on K. A sequence ν_n **converges weakly** to ν in $\mathscr{P}(K)$ if

$$\int_K f\, d\nu_n \longrightarrow \int_K f\, d\nu \qquad \text{for every } f \in C(K).$$

For a Borel map $F : X \to Y$ and a measure μ on X, the **push-forward** is $F_*\mu(A) = \mu(F^{-1}A)$. The Dirac probability at x is δ_x.

Definition — I.12.02: Blow-ups and tangent measures

For $x \in \mathbb{R}^d$ and $r > 0$, define

$$T_{x,r}(y) = \frac{y-x}{r}.$$

Let $\mu \in \mathscr{M}(\mathbb{R}^d)$ and $x \in \operatorname{supp}(\mu)$. A nonzero measure $\nu \in \mathscr{M}(\mathbb{R}^d)$ is a **tangent measure of** μ **at** x if there are $r_n \downarrow 0$ and constants $c_n > 0$ such that

$$c_n (T_{x,r_n})_*\mu \longrightarrow \nu \quad \text{vaguely.}$$

The collection of all such measures is the **tangent cone** $\operatorname{Tan}(\mu, x)$. The normalizing constants are part of the data; tangent measures are not required to be probabilities.

Definition — I.12.03: Centered sceneries and tangent distributions

Fix the observation window $B = [-1,1]^d$ and put $S_t(y) = e^t y$. For $\mu \in \mathscr{M}(\mathbb{R}^d)$, $x \in \operatorname{supp}(\mu)$, and $t \geq 0$, the **normalized centered scenery** is

$$\mu^{\square}_{x,t} = \frac{\big((S_t)_*(T_x)_*\mu\big)|_B}{\mu(x + e^{-t}B)} \in \mathscr{P}(B), \qquad T_x(y) = y - x.$$

Its **scenery distribution up to time** $T > 0$ is

$$\langle \mu \rangle_{x,T} = \frac{1}{T}\int_0^T \delta_{\mu^{\square}_{x,t}}\, dt \in \mathscr{P}(\mathscr{P}(B)).$$

A weak accumulation point of these distributions as $T \to \infty$ is a **tangent distribution of** μ **at** x. If the distributions converge to P, then μ **generates** P **at** x. The measure μ is a **scaling measure** if it generates some P_x at μ-almost every x; it is a **uniformly scaling measure** if the same distribution P is generated at μ-almost every x.
These terms do not assert that a tangent distribution is a fractal distribution; that is part of Q-I.12.09 and its later proof owner.

Definition — I.12.04: Empirical distributions in compact dynamics

Let X be compact metric. For a continuous map $F : X \to X$, the **length-N empirical distribution of** z is

$$A_N(z) = \frac{1}{N}\sum_{n=0}^{N-1} \delta_{F^n z}.$$

For a continuous flow $(F_t)_{t\in R}$ on X, its **time-T empirical distribution** is

$$A_T(z) = \frac{1}{T}\int_0^T \delta_{F_t z}\,dt.$$

A probability P is **F-invariant** if $F_*P = P$; it is **flow-invariant** if $(F_t)_*P = P$ for every $t \in R$.

Definition — I.12.05: Symbolic components and marked measures

Let Λ be a nonempty finite alphabet, $\Sigma = \Lambda^{\mathbb{N}}$, and $\sigma(\omega_1\omega_2\ldots) = \omega_2\omega_3\ldots$. For $i \in \Lambda$, let $[i] = \{\omega : \omega_1 = i\}$. If $\nu \in \mathscr{P}(\Sigma)$ and $\nu([i]) > 0$, its **first-level conditional component** is

$$\nu_i(A) = \frac{\nu(iA)}{\nu([i])}, \qquad iA = \{i\eta : \eta \in A\}.$$

For a word $u = u_1\ldots u_n$ with $\nu([u]) > 0$, define the **component at u** by

$$\nu_u(A) = \frac{\nu(uA)}{\nu([u])}.$$

For the empty word, set $\nu_\varnothing = \nu$.
The compact **marked-measure ambient space** is

$$Y_\Sigma = \mathscr{P}(\Sigma) \times \Sigma.$$

A pair (ν, ω) is **admissible** if $\omega \in \operatorname{supp}(\nu)$. The admissible set need not be closed in Y_Σ: an atom marking ω can lose all its mass in a weak limit. It is therefore not called a compact state space.

Definition — I.12.06: Symbolic magnification and adaptedness

Fix once and for all a reference probability $\vartheta \in \mathscr{P}(\Sigma)$. The **symbolic magnification map** is the Borel map

$$M : Y_\Sigma \longrightarrow Y_\Sigma, \qquad M(\nu,\omega) = \begin{cases} (\nu_{\omega_1}, \sigma\omega), & \nu([\omega_1]) > 0, \\ (\vartheta, \sigma\omega), & \nu([\omega_1]) = 0. \end{cases}$$

The second line only makes M total. Let $Q \in \mathscr{P}(Y_\Sigma)$ and let $\bar{Q}$ be its first marginal. The distribution Q is **adapted** if, for every $F \in C(\mathscr{P}(\Sigma)x\Sigma)$,

$$\int F(\nu,\omega)\,dQ(\nu,\omega) = \int\left(\int F(\nu,\eta)\,d\nu(\eta)\right)d\bar{Q}(\nu).$$

The conditional-probability interpretation is that, given the first coordinate ν, the second coordinate has law ν; the displayed identity is the formal definition.

An adapted, M-invariant probability on Y_Σ is called an **invariant adapted symbolic magnification distribution**. This is the symbolic model used in I.12, not a definition of every geometric CP-distribution in the literature. Adaptedness implies that admissible pairs have full Q-measure. For $\nu \in \mathscr{P}(\Sigma)$, its **canonical adapted seed** is

$$Q_\nu = \int_\Sigma \delta_{(\nu,\omega)}\, d\nu(\omega).$$

Definition — I.12.07: Boundary-null normalized restriction

For $\lambda \in \mathscr{M}(\mathbb{R}^d)$ satisfying $\lambda(B) > 0$, define

$$\mathscr{N}_B(\lambda) = \frac{\lambda|_B}{\lambda(B)} \in \mathscr{P}(B).$$

The measure λ is a **continuity point for normalization on** B if

$$\lambda(B) > 0 \quad \text{and} \quad \lambda(\partial B) = 0.$$

The boundary-null clause is essential: restriction to a closed window is not continuous at a measure charging its boundary.

Definition — I.12.08: Three dimension certificates

Let $(S_i)_{i\in\Lambda}$ be a finite similarity IFS with ratios r_i and attractor K, and let s solve $\sum_i r_i^s = 1$. Put $p_i = r_i^s$ and let μ be the corresponding self-similar probability.
The **covering certificate** is the level-n identity

$$\sum_{|u|=n} \operatorname{diam}(S_u K)^t = \operatorname{diam}(K)^t \left(\sum_i r_i^t\right)^n.$$

The **energy certificate below** s is finiteness of the Riesz energy $I_t(\mu)$ for every $0 < t < \min\{s, d\}$, obtained from an s-Frostman bound. The capstone's SSC hypotheses imply $s \le d$, so its range is exactly $0 < t < s$.
The **entropy certificate** is the quotient

$$\frac{h(p)}{\chi(p)}, \qquad h(p) = -\sum_i p_i \log p_i, \qquad \chi(p) = -\sum_i p_i \log r_i.$$

These are called certificates only after their hypotheses and conclusions are proved; they are not interchangeable slogans.

Terminal roadmap terminology

Q-I.12.09 prints the full definitions of normalized full-space scaling, quasi-Palm probabilities, fractal distributions and their restricted versions, geometric base-b CP-distributions in a disjoint unit-cube convention, and discrete/continuous centering. Those definitions precede the roadmap's use of the terms. None is an incoming premise for an I.12 F proof, and no later structure theorem is smuggled into a definition.

Every statement labelled F is proved below. Q-I.12.09 is a later-owner boundary, not a theorem assertion. Intuition and historical commentary follow the main development.

The only allowed-book facts used without local proof are the following:

1. separability of $C(K)$ for compact metric K: ARA-KB1, Corollary 2.59, printed p. 127 / local PDF p. 149;
2. Urysohn's lemma: ARA-KB1, Theorem 10.42, printed p. 474 / local PDF p. 495;
3. the Tietze extension theorem: ARA-KB1, Problems 18–20, printed p. 483 / local PDF p. 504;
4. Riesz representation and regularity of the representing Borel measure: ARA-KB1, Theorem 11.1 and the preceding definition, printed p. 490 / local PDF p. 511; and
5. the strong law for i.i.d. integrable random variables: ARA-KA1, Theorem 9.8, printed pp. 394–398 / local PDF pp. 417–421.

All other non-incoming assertions used by an I.12 F result are proved below.

I.12.1 Measure topologies and compact empirical dynamics

Continuous test functions in this chapter are real-valued. Testing their real and imaginary parts gives the equivalent convention with complex-valued functions.

Let $\mathscr{M}(\mathbb{R}^d)$ denote the nonnegative Radon measures on $\mathbb{R}^d$. A sequence (λ_n) **converges vaguely** to λ when

$$\int f\,d\lambda_n \longrightarrow \int f\,d\lambda \qquad (f \in C_c(\mathbb{R}^d)). \tag{I.12.1}$$

For a compact metric space K, let $\mathscr{P}(K)$ be its Borel probabilities. Weak convergence in $\mathscr{P}(K)$ means

$$\nu_n \longrightarrow \nu \quad \Longleftrightarrow \quad \int_K f\,d\nu_n \longrightarrow \int_K f\,d\nu \quad (f \in C(K)). \tag{I.12.2}$$

If $F : X \to Y$ is Borel and μ is a measure on X, its push-forward is

$$F_*\mu(A) = \mu(F^{-1}A). \tag{I.12.3}$$

The probability concentrated at x is δ_x.

Let $T : K \to K$ be continuous. The length-N empirical distribution of $z \in K$ is

$$A_N^T(z) = \frac{1}{N}\sum_{n=0}^{N-1} \delta_{T^n z}. \tag{I.12.4}$$

If $(T_t)_{t\in\mathbb{R}}$ is a continuous flow, meaning that $(t,z) \mapsto T_t z$ is jointly continuous and

$$T_0 = \mathrm{id}_K, \qquad T_{s+t} = T_s \circ T_t, \tag{I.12.5}$$

its time-R empirical distribution is

$$A_R^{\mathrm{T}}(z) = \frac{1}{R}\int_0^R \delta_{T_t z}\, dt. \tag{I.12.6}$$

A probability P is T-invariant when $T_* P = P$, and is invariant under the flow when $(T_t)_* P = P$ for every $t \in \mathbb{R}$.

Lemma — L-I.12.01 — Compactness and empirical invariance [F]

If K is compact metric, then $\mathscr{P}(K)$ is compact and metrizable in the weak topology. Every weak accumulation point of (I.12.4) is T-invariant. Every weak accumulation point of (I.12.6), as $R \to \infty$, is invariant under the flow.

Proof: weak compactness and metrizability

By the first allowed-book item above, $C(K)$, with the uniform norm, is separable. Choose a sequence $(f_j)_{j\geq 1}$ dense in its closed unit ball. Given $(\nu_n) \subset \mathscr{P}(K)$, each scalar sequence

$$\left(\int f_j \, d\nu_n\right)_{n\geq 1} \subset [-1, 1] \tag{I.12.7}$$

has a convergent subsequence. Successive extraction and diagonalization produce (ν_{n_k}) for which (I.12.7) converges for every j.

For $f \in C(K)$, put $a = \max\{1, \|f\|_\infty\}$. Given $\varepsilon > 0$, choose one j with $\|f - af_j\|_\infty < \varepsilon$. Since every ν_{n_k} has mass one,

$$\left|\int (f - g)\, d\nu_{n_k}\right| \leq \|f - g\|_\infty. \tag{I.12.8}$$

The scalar sequence $\int f_j \, d\nu_{n_k}$ converges, so (I.12.8), first with k, ℓ large and then with $\varepsilon \downarrow 0$, shows that $\int f \, d\nu_{n_k}$ is Cauchy for every $f \in C(K)$. Define

$$L(f) = \lim_{k\to\infty} \int f \, d\nu_{n_k}. \tag{I.12.9}$$

Then L is linear, $L(f) \geq 0$ for $f \geq 0$, and $L(1) = 1$. By the fourth allowed-book item above, there is a unique $\nu \in \mathscr{P}(K)$ satisfying

$$L(f) = \int f \, d\nu \qquad (f \in C(K)). \tag{I.12.10}$$

Thus every sequence in $\mathscr{P}(K)$ has a weakly convergent subsequence.

After replacing (f_j) by a dense sequence in the unit ball that contains the constant function, define

$$d_{\mathrm{w}}(\nu, \eta) = \sum_{j=1}^{\infty} 2^{-j-1} \left|\int f_j \, d\nu - \int f_j \, d\eta\right|. \tag{I.12.11}$$

The series converges because every summand before its coefficient is at most 2. Symmetry and the triangle inequality hold term by term. If $d_{\mathrm{w}}(\nu, \eta) = 0$, the two measures integrate every f_j equally; density and (I.12.8) extend the equality to every $f \in C(K)$, and

Riesz uniqueness gives $\nu = \eta$. Thus (I.12.11) is a metric. If $d_{\mathrm{w}}(\nu_n, \nu) \to 0$, density and (I.12.8) imply weak convergence. The converse follows by splitting the series into a finite head and a uniformly small tail. Hence (I.12.11) metrizes the weak topology. Sequential compactness in this metric is compactness.

Proof: discrete empirical invariance

For $f \in C(K)$,

$$\begin{aligned} \int f\, d(T_* A_N^T(z)) - \int f\, dA_N^T(z) &= \frac{1}{N} \sum_{n=0}^{N-1} \big(f(T^{n+1}z) - f(T^n z)\big) \\ &= \frac{f(T^N z) - f(z)}{N}. \end{aligned} \tag{I.12.12}$$

The absolute value is at most $2\|f\|_\infty / N$. If $A_{N_k}^T(z) \to P$ weakly, continuity of $f \circ T$ and (I.12.12) give

$$\int f\, d(T_* P) = \int f\, dP \qquad (f \in C(K)). \tag{I.12.13}$$

Therefore $T_* P = P$.

Proof: continuous-flow empirical invariance

Fix $s \geq 0$. The flow identity and a change of variable give

$$\begin{aligned} &\int f\, d((T_s)_* A_R^{\mathrm{T}}(z)) - \int f\, dA_R^{\mathrm{T}}(z) \\ &\quad = \frac{1}{R} \left(\int_R^{R+s} f(T_u z)\, du - \int_0^s f(T_u z)\, du \right). \end{aligned} \tag{I.12.14}$$

Its absolute value is at most $2s\|f\|_\infty / R$. Passing along any weakly convergent subsequence yields $(T_s)_* P = P$. Since this holds for every $s \geq 0$, and T_{-s} is the inverse of T_s, it holds for every real s. This proves the lemma. □

I.12.2 Tangent measures and the tangent cone

For $x \in \mathbb{R}^d$ and $r > 0$, put

$$T_{x,r}(y) = \frac{y - x}{r}. \tag{I.12.15}$$

Let $\mu \in \mathscr{M}(\mathbb{R}^d)$ and $x \in \operatorname{supp} \mu$. A nonzero $\nu \in \mathscr{M}(\mathbb{R}^d)$ is a **tangent measure of μ at x** if

$$c_n (T_{x,r_n})_* \mu \longrightarrow \nu \quad \text{vaguely} \tag{I.12.16}$$

for some $r_n \downarrow 0$ and $c_n > 0$. The set of tangent measures is $\operatorname{Tan}(\mu, x)$. The constants c_n are not prescribed; in particular, tangent measures need not be probabilities.

For $a > 0$, let

$$D_a(y) = a^{-1}y. \tag{I.12.17}$$

Lemma — L-I.12.02 — Algebra of tangent measures [F]

For every $c, a > 0$,

$$\nu \in \mathrm{Tan}(\mu, x) \quad \Longrightarrow \quad c(D_a)_*\nu \in \mathrm{Tan}(\mu, x). \tag{I.12.18}$$

If $\nu_j \in \mathrm{Tan}(\mu, x)$, $\nu_j \to \nu$ vaguely, and $\nu \neq 0$, then $\nu \in \mathrm{Tan}(\mu, x)$.

Proof: the cone and dilation identities

Multiplying every normalizing constant in (I.12.16) by c proves closure under positive scalar multiplication. Moreover,

$$T_{x,ar} = D_a \circ T_{x,r}. \tag{I.12.19}$$

Push-forward commutes with vague limits under the homeomorphism D_a. Thus (I.12.16) and (I.12.19), with scales $ar_n \downarrow 0$, prove dilation closure.

Proof: the diagonal closed-limit assertion

Choose a countable family $(\varphi_\ell)_{\ell \geq 1} \subset C_c(\mathbb{R}^d)$ with the following properties:

1. for every m, it contains a uniformly dense subset of the continuous functions supported in $B(0, m + 1)$; and
2. for every integer $m \geq 1$, the family contains a function $\psi_m \geq 0$ satisfying

$$\psi_m = 1 \text{ on } \overline{B}(0, m), \qquad \operatorname{supp} \psi_m \subset B(0, m + 1). \tag{I.12.20}$$

Such a family is obtained from separability of the continuous functions on the compact balls and a countable collection of cutoffs.

For each j, tangent-measure convergence supplies a blow-up

$$\lambda_j = c_j(T_{x,r_j})_*\mu, \qquad 0 < r_j < \min\{j^{-1}, r_{j-1}/2\}, \quad r_0 = 1, \tag{I.12.21}$$

so far enough along its defining sequence that

$$\left| \int \varphi_\ell \, d\lambda_j - \int \varphi_\ell \, d\nu_j \right| < j^{-1} \qquad (1 \leq \ell \leq j). \tag{I.12.22}$$

Since $\nu_j \to \nu$ vaguely, (I.12.22) implies

$$\int \varphi_\ell \, d\lambda_j \longrightarrow \int \varphi_\ell \, d\nu \qquad (\ell \geq 1). \tag{I.12.23}$$

In particular, (I.12.20) and (I.12.23) give, for every fixed m,

$$\sup_{j \geq j(m)} \lambda_j(\overline{B}(0,m)) \leq \sup_{j \geq j(m)} \int \psi_m \, d\lambda_j < \infty. \tag{I.12.24}$$

Let $f \in C_c(\mathbb{R}^d)$, supported in $B(0, m+1)$. Approximate f uniformly on $\mathbb{R}^d$ by members of the declared family supported in that same ball. The local mass bound (I.12.24) controls the approximation error uniformly in large j. Together with (I.12.23), this proves

$$\int f \, d\lambda_j \longrightarrow \int f \, d\nu. \tag{I.12.25}$$

Thus $\lambda_j \to \nu$ vaguely. The scales in (I.12.21) tend to zero, and $\nu \neq 0$, so (I.12.21) is a tangent sequence for ν. □

I.12.3 Normalized sceneries and their accumulation distributions

Fix

$$B = [-1,1]^d, \qquad S_t(y) = e^t y, \qquad T_x(y) = y - x. \tag{I.12.26}$$

For $\mu \in \mathscr{M}(\mathbb{R}^d)$, $x \in \operatorname{supp} \mu$, and $t \geq 0$, define the **normalized centered scenery**

$$\mu^{\square}_{x,t} = \frac{\big((S_t)_*(T_x)_*\mu\big)|_B}{\mu(x + e^{-t}B)} \in \mathscr{P}(B). \tag{I.12.27}$$

The denominator is finite because $x + e^{-t}B$ is compact. It is positive because this cube is a neighborhood of the support point x.

For $R > 0$, the **scenery distribution up to time** R is

$$\langle \mu \rangle_{x,R} = \frac{1}{R} \int_0^R \delta_{\mu^{\square}_{x,t}} \, dt \in \mathscr{P}(\mathscr{P}(B)). \tag{I.12.28}$$

A weak accumulation point of (I.12.28), as $R \to \infty$, is a **tangent distribution of** μ **at** x. If $\langle \mu \rangle_{x,R} \to P$, then μ **generates** P **at** x. It is a **scaling measure** if it generates some P_x at μ-almost every x, and a **uniformly scaling measure** if one common P is generated at μ-almost every x.

These definitions make no invariance or spatial resampling assertion; the precise later terminology is defined in the terminal discussion.

Lemma — L-I.12.03 — Existence of scenery accumulation distributions [F]

The map

$$[0,\infty) \longrightarrow \mathscr{P}(B), \qquad t \longmapsto \mu^{\square}_{x,t}, \tag{I.12.29}$$

is Borel. The family $\{\langle \mu \rangle_{x,R} : R \geq 1\}$ has nonempty compact closure in $\mathscr{P}(\mathscr{P}(B))$. Hence every support point has at least one tangent distribution.

Proof

For $g \in C(B)$, let $\hat{g} : \mathbb{R}^d \to \mathbb{R}$ equal g on B and zero off B. This extension is bounded and Borel, though it need not be continuous on ∂B. Equation (I.12.27) gives

$$\int_B g \, d\mu_{x,t}^{\square} = \frac{\int_{\mathbb{R}^d} \hat{g}(e^t(y-x)) \, d\mu(y)}{\int_{\mathbb{R}^d} \mathbf{1}_B(e^t(y-x)) \, d\mu(y)}. \tag{I.12.30}$$

The two integrands are Borel functions of (t, y). On every bounded t-interval they vanish outside a fixed compact set in the y variable and are bounded. Approximation by nonnegative simple functions, followed by monotone convergence, proves measurability for the denominator and separately for the positive and negative parts of the numerator. Thus both integrals in (I.12.30) are Borel functions of t. The denominator is positive, so the quotient is Borel.

Choose a countable dense family in $C(B)$. The weak Borel structure on $\mathscr{P}(B)$ is generated by the corresponding integration maps, by the metric (I.12.11). Equation (I.12.30) therefore proves (I.12.29) is Borel, so (I.12.28) is well defined.

The set B is compact. Lemma L-I.12.01 applied first to B and then to the compact metric space $\mathscr{P}(B)$ gives

$$\mathscr{P}(B) \text{ compact metric}, \qquad \mathscr{P}(\mathscr{P}(B)) \text{ compact metric}. \tag{I.12.31}$$

Thus every sequence $R_j \to \infty$ has a subsequence on which (I.12.28) converges. □

I.12.3.1 Why compactness alone does not prove scenery invariance

For $\lambda \in \mathscr{M}(\mathbb{R}^d)$ with $\lambda(B) > 0$, put

$$\mathcal{N}_B(\lambda) = \frac{\lambda|_B}{\lambda(B)}. \tag{I.12.32}$$

Although unnormalized dilation is continuous in the vague topology, $\lambda \mapsto \mathcal{N}_B(\lambda)$ can be discontinuous when $\lambda(\partial B) > 0$. Therefore the endpoint identity for empirical measures cannot be passed to arbitrary scenery limits without additional work. This obstruction is recorded in HOC13-FD, local PDF p. 3 and the discussion on local PDF p. 6. No result from that source is used here.

Lemma — L-I.12.04 — Boundary-null tangent/scenery bridge [F]

Suppose $r_n \downarrow 0$, $c_n > 0$, and

$$\lambda_n = c_n (T_{x,r_n})_* \mu \longrightarrow \nu \quad \text{vaguely}, \tag{I.12.33}$$

where

$$\nu(B) > 0, \qquad \nu(\partial B) = 0. \tag{I.12.34}$$

Discard the finitely many scales with $r_n > 1$, so $t_n = -\log r_n \geq 0$. Then

$$\mu^{\square}_{x,t_n} \longrightarrow \mathcal{N}_B(\nu) \quad \text{weakly in } \mathscr{P}(B). \tag{I.12.35}$$

Proof

For $\delta > 0$, let

$$C_\delta = \{y : \operatorname{dist}(y, \partial B) \le \delta\}. \tag{I.12.36a}$$

For $0 < \delta \le 1$, the compact sets C_δ decrease to ∂B as $\delta \downarrow 0$; $\nu(C_1) < \infty$ because ν is Radon. By (I.12.34) and continuity from above,

$$\nu(C_\delta) \longrightarrow 0. \tag{I.12.36b}$$

We use the following direct compact-set consequence of vague convergence. If C is compact, outer regularity and Urysohn's lemma give, for every $\varepsilon > 0$, a function $0 \le \phi \in C_c(\mathbb{R}^d)$ with $\phi = 1$ on C and $\int \phi\, d\nu \le \nu(C) + \varepsilon$. Therefore

$$\limsup_{n\to\infty} \lambda_n(C) \le \lim_{n\to\infty} \int \phi\, d\lambda_n = \int \phi\, d\nu \le \nu(C) + \varepsilon. \tag{I.12.36c}$$

Letting $\varepsilon \downarrow 0$ proves the compact-set upper bound.

Choose $\chi_\delta \in C_c(\mathbb{R}^d)$ such that

$$0 \le \chi_\delta \le \mathbf{1}_B, \qquad \chi_\delta = 1 \text{ on } B \setminus C_\delta. \tag{I.12.36d}$$

Vague convergence applies to χ_δ, while (I.12.36c) controls the error inside C_δ. Letting first $n \to \infty$ and then $\delta \downarrow 0$ using (I.12.36b) gives

$$\lambda_n(B) \longrightarrow \nu(B). \tag{I.12.36}$$

Let $g \in C(B)$. By the Tietze extension theorem and multiplication by a compactly supported cutoff equal to one near B, choose $\tilde{g} \in C_c(\mathbb{R}^d)$ with $\tilde{g}|_B = g$. The same cutoff calculation gives

$$\left| \int_B \tilde{g}\, d\lambda_n - \int \tilde{g}\chi_\delta\, d\lambda_n \right| \le \|\tilde{g}\|_\infty \lambda_n(C_\delta), \tag{I.12.37a}$$

The analogous bound holds for ν. Vague convergence applies to $\tilde{g}\chi_\delta$. Therefore (I.12.36b)–(I.12.36c), followed by $\delta \downarrow 0$, yield

$$\int_B g\, d\lambda_n = \int \tilde{g}\mathbf{1}_B\, d\lambda_n \longrightarrow \int \tilde{g}\mathbf{1}_B\, d\nu = \int_B g\, d\nu. \tag{I.12.37}$$

For all n, support of x gives

$$\lambda_n(B) = c_n \mu(x + r_n B) > 0. \tag{I.12.38}$$

Divide (I.12.37) by (I.12.36). Since scalar normalization cancels,

$$\mathcal{N}_B(\lambda_n) = \frac{((T_{x,r_n})_*\mu)|_B}{\mu(x + r_n B)} = \mu^{\square}_{x,t_n}. \tag{I.12.39}$$

Equations (I.12.36)–(I.12.39) prove (I.12.35). □

I.12.4 Symbolic components and controlled magnification

Let Λ be a nonempty finite alphabet and put

$$\Sigma = \Lambda^{\mathbb{N}}, \qquad \sigma(\omega_1\omega_2\omega_3 \dots) = \omega_2\omega_3 \dots . \tag{I.12.40}$$

For a finite word $u = u_1 \cdots u_n$, its cylinder is

$$[u] = \{\omega \in \Sigma : \omega_1 = u_1, \dots, \omega_n = u_n\}. \tag{I.12.41}$$

Endow Σ with its product topology. Every cylinder is both open and closed. If $\nu \in \mathcal{P}(\Sigma)$ and $\nu([u]) > 0$, the **conditional component at** u is

$$\nu_u(A) = \frac{\nu(uA)}{\nu([u])}, \qquad uA = \{u\eta : \eta \in A\}. \tag{I.12.42}$$

For the empty word, set $\nu_\varnothing = \nu$; this fixes the generation-zero convention in the component averages below.

Equivalently,

$$\nu_u = (\sigma^n)_* \left(\frac{\nu|_{[u]}}{\nu([u])} \right). \tag{I.12.43}$$

The compact **marked-measure ambient space** is

$$Y_\Sigma = \mathcal{P}(\Sigma) \times \Sigma. \tag{I.12.44}$$

A pair (ν, ω) is **admissible** when $\omega \in \operatorname{supp} \nu$. If $\#\Lambda \geq 2$, admissible pairs do not form a closed set: choose $\omega \neq \eta$; then

$$\nu_n = (1 - n^{-1})\delta_\eta + n^{-1}\delta_\omega, \qquad (\nu_n, \omega) \longrightarrow (\delta_\eta, \omega), \tag{I.12.45}$$

and the limiting pair is not admissible. For a one-symbol alphabet the whole marked space is a single admissible point. This is why compactness is asserted for (I.12.44), not for the support-pair subset.

Fix a reference $\vartheta \in \mathcal{P}(\Sigma)$. Define the **symbolic magnification map** $M : Y_\Sigma \to Y_\Sigma$ by

$$M(\nu, \omega) = \begin{cases} (\nu_{\omega_1}, \sigma\omega), & \nu([\omega_1]) > 0, \\ (\vartheta, \sigma\omega), & \nu([\omega_1]) = 0. \end{cases} \tag{I.12.46}$$

The second line only makes M total. Its probabilistic irrelevance is proved after the required class of marked laws is defined in Lemma 5.1.

Lemma — L-I.12.05 — Controlled symbolic magnification [F]

The space Y_Σ is compact metric. The map M is Borel and is continuous at every (ν, ω) for which $\nu([\omega_1]) > 0$. Hence

$$\mathrm{Disc}(M) \subset Z := \{(\nu, \omega) : \nu([\omega_1]) = 0\}. \tag{I.12.47}$$

For every admissible pair and $n \geq 1$,

$$M^n(\nu, \omega) = \left(\nu_{\omega|_n}, \sigma^n \omega\right). \tag{I.12.48}$$

Proof

The countable product of the finite discrete alphabet is metrized by

$$d_{1/2}(\omega, \eta) = \begin{cases} 0, & \omega = \eta, \\ 2^{-|\omega \wedge \eta|}, & \omega \neq \eta, \end{cases} \tag{I.12.44a}$$

where $|\omega \wedge \eta|$ is the length of the longest common prefix. Given any sequence in Σ, successively choose subsequences on which the first, second, and later coordinates are constant; the diagonal subsequence converges in (I.12.44a). Thus Σ is sequentially compact and therefore compact metric. Lemma L-I.12.01 makes $\mathscr{P}(\Sigma)$ compact metric, and the product of two compact metric spaces makes (I.12.44) compact metric.

For $i \in \Lambda$, the map

$$\nu \longmapsto \nu([i]) \tag{I.12.49}$$

is continuous because $[i]$ is clopen. If $f \in C(\Sigma)$ and $\nu([i]) > 0$, then

$$\int f \, d\nu_i = \frac{1}{\nu([i])} \int_{[i]} f(\sigma\eta) \, d\nu(\eta). \tag{I.12.50}$$

The numerator is continuous in ν, because $\mathbf{1}_{[i]}(f \circ \sigma) \in C(\Sigma)$; the denominator is positive and continuous. Thus $\nu \mapsto \nu_i$ is weakly continuous on $\{\nu : \nu([i]) > 0\}$.

Suppose $(\nu_k, \omega^{(k)}) \to (\nu, \omega)$. Then $\omega_1^{(k)} = \omega_1$ for all large k. Equations (I.12.49)–(I.12.50) prove continuity of M at every positive-cell pair. They also make each branch of (I.12.46) Borel, proving Borel measurability and (I.12.47).

If (ν, ω) is admissible, every cylinder $[\omega|_n]$ has positive ν-mass. The component identity

$$(\nu_u)_v = \nu_{uv} \quad \text{whenever } \nu([uv]) > 0 \tag{I.12.51}$$

follows immediately from (I.12.42). Induction using (I.12.51) proves (I.12.48). □

I.12.5 Adapted distributions and invariant symbolic limits

Let $Q \in \mathscr{P}(Y_\Sigma)$, and let $\overline{Q}$ be its first marginal. We define Q to be **adapted** when, for every $F \in C(Y_\Sigma)$,

$$\int F(\nu,\omega)\,dQ(\nu,\omega) = \int_{\mathscr{P}(\Sigma)} \left(\int_{\Sigma} F(\nu,\eta)\,d\nu(\eta) \right) d\overline{Q}(\nu). \tag{I.12.52}$$

Thus, conditioned on the measure coordinate ν, the marked point has law ν. Formula (I.12.52), rather than this verbal interpretation, is the formal definition.

For $\nu \in \mathscr{P}(\Sigma)$, its **canonical adapted seed** is

$$Q_\nu = \int_{\Sigma} \delta_{(\nu,\omega)}\,d\nu(\omega). \tag{I.12.53}$$

An adapted probability Q satisfying $M_* Q = Q$ is an **invariant adapted symbolic magnification distribution**.

Lemma — 5.1 — Adapted laws avoid small and zero cells

If Q is adapted and $m = \#\Lambda$, then, for every $a > 0$,

$$Q\{(\nu,\omega) : \nu([\omega_1]) < a\} \le ma. \tag{I.12.54}$$

In particular,

$$Q(Z) = 0, \qquad Q\{(\nu,\omega) : \omega \in \operatorname{supp}\nu\} = 1. \tag{I.12.55}$$

Proof

For fixed ν,

$$\nu\{\omega : \nu([\omega_1]) < a\} = \sum_{\substack{i\in\Lambda \\ \nu([i])<a}} \nu([i]) \le ma. \tag{I.12.56}$$

The identity (I.12.52) says that Q equals the probability obtained by first sampling ν with law $\overline{Q}$ and then sampling ω with law ν. For cylinders, $\nu \mapsto \nu([u])$ is continuous. The class of sets whose evaluation map is Borel contains the cylinder algebra and is closed under increasing unions and decreasing intersections; the monotone-class theorem therefore gives every Borel set of Σ. Formally, measurability of the resulting kernel integral holds first for Borel rectangles and then for all Borel subsets of Y_Σ by a monotone-class argument. It therefore defines a Borel probability on Y_Σ. Equality on all continuous functions in (I.12.52) and Riesz uniqueness make it equal to Q. Therefore (I.12.56) integrates to (I.12.54). Letting $a \downarrow 0$ gives $Q(Z) = 0$.

To see the support assertion without another import, fix a countable base (U_j) of Σ. The complement of $\operatorname{supp}\nu$ is the union of those U_j for which $\nu(U_j) = 0$, and hence has ν-mass zero. Sampling ω from ν and integrating this assertion over $\overline{Q}$ proves the second part of (I.12.55). □

Lemma — 5.2 — Components preserve adaptedness

For $n \geq 0$,

$$M_*^n Q_\nu = \sum_{\substack{u \in \Lambda^n \\ \nu([u])>0}} \nu([u]) Q_{\nu_u}. \tag{I.12.57}$$

Consequently, $M_*^n Q_\nu$ is adapted.

Proof

Under Q_ν, the point ω has law ν. On the event $\omega|_n = u$, which has probability $\nu([u])$, the tail $\sigma^n\omega$ has law ν_u, and (I.12.48) makes the measure coordinate equal to ν_u. This proves (I.12.57). Each Q_{ν_u} satisfies (I.12.52) directly, and convex combinations preserve that identity. □

Lemma — 5.3 — Adaptedness is weakly closed

If $Q_k \to Q$ weakly in $\mathscr{P}(Y_\Sigma)$ and every Q_k is adapted, then Q is adapted.

Proof

For $F \in C(Y_\Sigma)$, define

$$J_F(\nu) = \int_\Sigma F(\nu, \eta)\, d\nu(\eta). \tag{I.12.58}$$

If $\nu_k \to \nu$, uniform continuity of F on the compact product gives

$$\sup_{\eta \in \Sigma} |F(\nu_k, \eta) - F(\nu, \eta)| \longrightarrow 0. \tag{I.12.59}$$

Hence

$$\begin{aligned} |J_F(\nu_k) - J_F(\nu)| &\leq \sup_\eta |F(\nu_k, \eta) - F(\nu, \eta)| \\ &\quad + \left| \int F(\nu, \eta)\, d\nu_k(\eta) - \int F(\nu, \eta)\, d\nu(\eta) \right| \longrightarrow 0. \end{aligned} \tag{I.12.60}$$

Thus J_F is continuous. The first marginals satisfy $\overline{Q}_k \to \overline{Q}$, because coordinate projection is continuous. Passing to the limit in (I.12.52), using (I.12.58)–(I.12.60), proves (I.12.52) for Q. □

Theorem — T-I.12.06 — Invariant adapted symbolic limits [F]

For $\nu \in \mathscr{P}(\Sigma)$, put

$$\mathscr{A}_N(\nu) = \frac{1}{N} \sum_{n=0}^{N-1} M_*^n Q_\nu. \tag{I.12.61}$$

The family $(\mathscr{A}_N(\nu))$ has weak accumulation points, and every one is an invariant adapted symbolic magnification distribution.

Proof: compactness and adaptedness

Every term of (I.12.61) is adapted by Lemma 5.2, hence so is every Cesaro average. Since Y_Σ is compact metric, L-I.12.01 gives compactness of $\mathscr{P}(Y_\Sigma)$. Choose

$$\mathscr{A}_{N_k}(\nu) \longrightarrow Q. \tag{I.12.62}$$

Lemma 5.3 makes Q adapted.

Proof: passing the discontinuous magnification to the limit

Let

$$q(\theta, \omega) = \theta([\omega_1]). \tag{I.12.63}$$

This function is continuous: on each clopen set $\mathscr{P}(\Sigma) \times [i]$ it equals the continuous function $\theta \mapsto \theta([i])$.

Fix $f \in C(Y_\Sigma)$ and $0 < a < 1/2$. Choose continuous $\chi_a : [0, 1] \to [0, 1]$ satisfying

$$\chi_a(s) = 0 \ (0 \le s \le a), \qquad \chi_a(s) = 1 \ (2a \le s \le 1). \tag{I.12.64}$$

Define

$$h_a(\theta, \omega) = \begin{cases} \chi_a(q(\theta, \omega)) f(M(\theta, \omega)), & q(\theta, \omega) > 0, \\ 0, & q(\theta, \omega) = 0. \end{cases} \tag{I.12.65}$$

On $\{q > 0\}$, L-I.12.05 makes M continuous. On $\{q \le a\}$, (I.12.65) is identically zero. Hence $h_a \in C(Y_\Sigma)$. For every adapted probability R, Lemma 5.1 gives

$$\left| \int f \circ M \, dR - \int h_a \, dR \right| \le \|f\|_\infty R\{q < 2a\} \le 2ma\|f\|_\infty. \tag{I.12.66}$$

Apply (I.12.66) to $R = \mathscr{A}_{N_k}(\nu)$ and to the adapted limit $R = Q$. Weak convergence and continuity of h_a give

$$\limsup_{k\to\infty} \left| \int f \circ M \, d\mathscr{A}_{N_k}(\nu) - \int f \circ M \, dQ \right| \le 4ma\|f\|_\infty. \tag{I.12.67}$$

Let $a \downarrow 0$. Thus

$$M_* \mathscr{A}_{N_k}(\nu) \longrightarrow M_* Q \quad \text{weakly.} \tag{I.12.68}$$

Finally, endpoint cancellation holds as an identity of signed measures:

$$M_* \mathscr{A}_N(\nu) - \mathscr{A}_N(\nu) = \frac{M_*^N Q_\nu - Q_\nu}{N}. \tag{I.12.69}$$

Testing (I.12.69) against f bounds its right side by $2\|f\|_\infty / N$. Pass along (I.12.62) using (I.12.68). Then

$$\int f\,d(M_*Q) = \int f\,dQ \qquad (f \in C(Y_\Sigma)), \tag{I.12.70}$$

so $M_*Q = Q$. □

I.12.6 The Bernoulli fixed point

Let $p = (p_i)_{i\in\Lambda}$ satisfy

$$p_i > 0, \qquad \sum_{i\in\Lambda} p_i = 1. \tag{I.12.71}$$

The **Bernoulli probability** β_p is determined by

$$\beta_p([u_1 \cdots u_n]) = \prod_{j=1}^{n} p_{u_j}. \tag{I.12.72}$$

For $0 < \rho < 1$, let

$$d_\rho(\omega,\eta) = \begin{cases} 0, & \omega = \eta, \\ \rho^{|\omega\wedge\eta|}, & \omega \neq \eta, \end{cases} \tag{I.12.73}$$

where $|\omega \wedge \eta|$ is the length of their longest common prefix. This is the **symbolic ultrametric with scale** ρ.

Corollary — C-I.12.07 — Bernoulli fixed point and dimension [F]

The probability

$$Q_{\beta_p} = \delta_{\beta_p} \otimes \beta_p \tag{I.12.74}$$

is adapted and M-invariant. In fact,

$$\mathscr{A}_N(\beta_p) = Q_{\beta_p} \qquad (N \geq 1). \tag{I.12.75}$$

The measure β_p is exact dimensional in (Σ, d_ρ), with

$$\dim_{\mathrm{H}} \beta_p = \frac{h(p)}{|\log \rho|}, \qquad h(p) = -\sum_{i\in\Lambda} p_i \log p_i. \tag{I.12.76}$$

If Σ codes a common-ratio ρ similarity IFS satisfying SSC, then its geometric projection $\pi_*\beta_p$ is exact dimensional with the same dimension.

Proof: stationarity

For $i \in \Lambda$ and every word u,

$$(\beta_p)_i([u]) = \frac{\beta_p([iu])}{\beta_p([i])} = \beta_p([u]). \tag{I.12.77}$$

Cylinders determine Borel probabilities, so $(\beta_p)_i = \beta_p$. Under Q_{β_p}, the point has law β_p, its tail also has law β_p, and the measure coordinate stays equal to β_p. Hence (I.12.74) is adapted and

$$M_* Q_{\beta_p} = Q_{\beta_p}. \tag{I.12.78}$$

Equation (I.12.75) follows.

Proof: symbolic exact dimension

For β_p-almost every ω, the fifth allowed-book item applied to the i.i.d. variables $\log p_{\omega_j}$ gives

$$\frac{1}{n} \log \beta_p([\omega|_n]) = \frac{1}{n} \sum_{j=1}^{n} \log p_{\omega_j} \longrightarrow \sum_i p_i \log p_i = -h(p). \tag{I.12.79}$$

Ultrametric balls at radii between ρ^{n+1} and ρ^n are one of the two adjacent prefix cylinders; the one-index ambiguity disappears after division by n. Therefore

$$\lim_{r \downarrow 0} \frac{\log \beta_p(B_{d_\rho}(\omega, r))}{\log r} = \frac{-h(p)}{\log \rho} = \frac{h(p)}{|\log \rho|} \tag{I.12.80}$$

at β_p-almost every ω. This is exact dimensionality.

Under SSC and common ratio ρ, the coding map is bi-Lipschitz between (Σ, d_ρ) and the attractor: the upper estimate follows by diameter scaling, and the lower estimate follows by scaling the positive first-level separation at the longest common prefix. Bi-Lipschitz maps preserve the two local-dimension limits, proving the geometric assertion. □

I.12.7 Capstone: three dimension certificates

For a one-map IFS the attractor is its fixed point, $s = 0$, and the canonical probability is a point mass. Its set and measure dimensions are zero, the entropy is zero, and there is no energy parameter $0 < t < s$. The covering and entropy conclusions are immediate in that case. We now give the nontrivial proof for at least two maps.

Let $(S_i)_{i \in \Lambda}$ be a finite similarity IFS on $\mathbb{R}^d$ with $\#\Lambda \geq 2$, ratios $0 < r_i < 1$, attractor K, and SSC:

$$S_i(K) \cap S_j(K) = \varnothing \qquad (i \neq j). \tag{I.12.81}$$

There is a unique $s > 0$ such that

$$\sum_{i \in \Lambda} r_i^s = 1. \tag{I.12.82}$$

Put $p_i = r_i^s$, and let μ be the self-similar probability

$$\mu = \sum_{i \in \Lambda} p_i (S_i)_* \mu. \tag{I.12.83}$$

For $u = u_1 \cdots u_n$, write

$$r_u = \prod_{j=1}^{n} r_{u_j}, \qquad K_u = S_u(K), \qquad \mu(K_u) = p_u = r_u^s. \tag{I.12.84}$$

The last equality uses SSC, which makes the coding injective.

Theorem — T-I.12.08 — Covering, energy, and entropy routes agree [F]

For the system above, $0 < s \le d$ and

$$\begin{aligned} \dim_{\mathrm{H}} K = s, \qquad &\mu \text{ is exact dimensional with dimension } s, \\ &I_t(\mu) < \infty \qquad (0 < t < s). \end{aligned} \tag{I.12.85}$$

Here

$$I_t(\mu) = \iint |x-y|^{-t}\, d\mu(x) d\mu(y) \tag{I.12.86}$$

is the t-Riesz energy defined in I.3.

Proof: the covering certificate

The level-n cylinder sets cover K, and

$$\sum_{|u|=n} \operatorname{diam}(K_u)^t = \operatorname{diam}(K)^t \left(\sum_i r_i^t \right)^n . \tag{I.12.87}$$

If $t > s$, strict monotonicity of $t \mapsto \sum_i r_i^t$ and (I.12.82) give $\sum_i r_i^t < 1$. The right side of (I.12.87) tends to zero, while the largest cylinder diameter tends to zero. Hence

$$\mathscr{H}^t(K) = 0 \quad (t > s), \qquad \dim_{\mathrm{H}} K \le s. \tag{I.12.88}$$

Proof: the Frostman and energy certificate

Let $r_{\min} = \min_i r_i$. For $0 < r < \min\{1, \operatorname{diam} K\}$, define the stopping antichain

$$\mathscr{W}(r) = \{u : r_u \le r < r_{u^-}\}, \tag{I.12.89}$$

where only nonempty words are used, u^- is obtained by deleting the final symbol, and $r_\varnothing = 1$. Then

$$r_{\min} r < r_u \le r \qquad (u \in \mathscr{W}(r)). \tag{I.12.90}$$

Let

$$\Delta = \min_{i \ne j} \operatorname{dist}(S_i K, S_j K) > 0. \tag{I.12.91}$$

If $u, v \in \mathscr{W}(r)$ are distinct, let w be their longest common prefix. Scaling (I.12.91) inside S_w gives

$$\operatorname{dist}(K_u, K_v) \geq r_w\Delta \geq r_{\min} r\Delta. \tag{I.12.92}$$

Choose one point from each stopping cylinder that meets a fixed Euclidean ball $B(x,r)$. These points lie in $B(x, r + \operatorname{diam} K\, r)$ and are separated by $r_{\min}\Delta r$. Comparing the disjoint balls of radius $r_{\min}\Delta r/3$ around them with one containing Euclidean ball shows that their number is at most a constant

$$N_0 = N_0(d, K, r_{\min}, \Delta) < \infty, \tag{I.12.93}$$

independent of x and r. From (I.12.84) and (I.12.90),

$$\mu(B(x,r)) \leq \sum_{\substack{u\in\mathscr{W}(r)\\ K_u\cap B(x,r)\neq\emptyset}} \mu(K_u) \leq N_0 r^s. \tag{I.12.94}$$

After increasing the constant, (I.12.94) holds for every $r > 0$. Thus μ is an s-Frostman probability. The mass-distribution theorem proved in I.3 gives

$$\dim_{\mathrm{H}} K \geq s. \tag{I.12.95}$$

Together with (I.12.88), this proves $\dim_{\mathrm{H}} K = s$. Since $K \subset \mathbb{R}^d$, its Hausdorff dimension is at most d, so $s \leq d$. In particular every $0 < t < s$ lies in the defining range $0 < t < d$ of the Riesz energies in I.3.

For $0 < t < s$, the layer-cake identity gives, for every x,

$$\begin{aligned}\int |x-y|^{-t}\, d\mu(y) &= t\int_0^\infty r^{-t-1}\mu(B(x,r))\, dr \\ &\leq tN_0\int_0^1 r^{s-t-1}\, dr + t\int_1^\infty r^{-t-1}\, dr \\ &= \frac{tN_0}{s-t} + 1 < \infty,\end{aligned} \tag{I.12.96}$$

Here the Frostman constant has already been enlarged to apply at every radius, and the bound for $r \geq 1$ uses only $\mu(\mathbb{R}^d) = 1$; no rescaling is needed. Integrating in x proves $I_t(\mu) < \infty$.

Proof: the entropy certificate

The Shannon entropy and Lyapunov exponent are

$$h(p) = -\sum_i p_i \log p_i, \qquad \chi(p) = -\sum_i p_i \log r_i. \tag{I.12.97}$$

Since $p_i = r_i^s$,

$$h(p) = -\sum_i r_i^s \log(r_i^s) = -s\sum_i r_i^s \log r_i = s\chi(p). \tag{I.12.98}$$

Thus

$$\frac{h(p)}{\chi(p)} = s. \tag{I.12.99}$$

For completeness, the local quotient is already exact on every cylinder:

$$\frac{\log \mu(K_{\omega|_n})}{\log r_{\omega|_n}} = \frac{\log r^s_{\omega|_n}}{\log r_{\omega|_n}} = s. \tag{I.12.100}$$

For the reverse ball estimate, fix $x = \pi(\omega)$, take $0 < r < \operatorname{diam} K$, and choose the first n for which

$$\operatorname{diam}(K) r_{\omega|_n} < r. \tag{I.12.100a}$$

The strict inequality ensures $K_{\omega|_n} \subset B(x,r)$ even for the open-ball convention. Minimality and the final-symbol ratio give

$$r_{\omega|_n} \geq \frac{r_{\min} r}{\operatorname{diam}(K)}. \tag{I.12.100b}$$

Consequently,

$$\mu(B(x,r)) \geq \mu(K_{\omega|_n}) = r^s_{\omega|_n} \geq \left(\frac{r_{\min}}{\operatorname{diam}(K)}\right)^s r^s. \tag{I.12.100c}$$

Together, (I.12.94) and (I.12.100c) convert (I.12.100) into

$$\lim_{r \downarrow 0} \frac{\log \mu(B(x,r))}{\log r} = s \quad \text{for } \mu\text{-almost every } x. \tag{I.12.101}$$

This is exact dimensionality, and (I.12.98)–(I.12.101) identify its value with the entropy/Lyapunov quotient. All three certificates therefore return the same exponent s. □

I.12.8 Terminal boundary: what compactness does not supply

Research boundary — Q-I.12.09 — Full fractal-distribution and CP-centering theory [Q]

Here are the definitions needed to read the roadmap; none supplies the deeper theorem. Retain $B = [-1,1]^d$, and set

$$\mathscr{M}^* = \{\lambda \in \mathscr{M}(\mathbb{R}^d) \,:\, \lambda(B) = 1,\ 0 \in \operatorname{supp}\lambda\}, \qquad \lambda^* = \lambda/\lambda(B) \quad (0 < \lambda(B) < \infty).$$

For $\lambda \in \mathscr{M}^*$ define $S_t^*\lambda = ((S_t)_*\lambda)^*$, $t \in \mathbb{R}$. For $x \in \operatorname{supp}\lambda$, define $T_x^*\lambda = ((T_x)_*\lambda)^*$. A probability P on $\mathscr{M}^*$, with the Borel structure inherited from vague convergence, is **scale-invariant** if $(S_t^*)_*P = P$ for every $t \in \mathbb{R}$. It is **quasi-Palm** if, for every bounded open neighborhood U of zero, the measure

$$\mathscr{D}_U(P)(A) = \int\int_U \mathbf{1}_A(T_x^*\lambda)\, d\lambda(x)\, dP(\lambda)$$

has exactly the same null Borel sets as P: $\mathscr{D}_U(P)(A) = 0$ if and only if $P(A) = 0$. The inner integration can be restricted to support points; its complement has zero λ-mass. No normalization of $\lambda|_U$ is inserted in this definition, and the averaged measure can have infinite total mass. A **fractal distribution** is a scale-invariant quasi-Palm probability. Its **restricted version** is $(\mathscr{N}_B)_*P$ on $\mathscr{P}(B)$. These are the normalized full-space and restricted conventions of Hochman, Definitions 1.1–1.2.

To define the geometric counterpart of symbolic magnification, use the unit cube $D = [0,1]^d$ and an integer $b \geq 2$. Partition D into the b^d cells $C_i = \prod_{j=1}^d [i_j/b, (i_j+1)/b)$, $i_j \in \{0, \dots, b-1\}$, closing only faces on the outer boundary of D. Let $C(x)$ be the unique cell containing x, and put $L_{C_i}(y) = by - i$. For $\nu(C(x)) > 0$, define

$$M_b^D(\nu, x) = \left((L_{C(x)})_* \frac{\nu|_{C(x)}}{\nu(C(x))}, L_{C(x)}x \right).$$

For a zero-mass selected cell use an arbitrary fixed probability as the first coordinate; this makes the map total. A probability Q on $\mathscr{P}(D) \times D$ is **adapted** when

$$\int F(\nu, x)\, dQ = \int\int F(\nu, y)\, d\nu(y)\, d\overline{Q}(\nu) \quad (F \in C(\mathscr{P}(D) \times D)).$$

It is a **restricted geometric base-b CP-distribution** if it is adapted and invariant under M_b^D. The initials CP mean conditional probability. Its associated stationary magnification process is obtained by sampling (ν, x) with law Q and iterating M_b^D; this is the CP-chain viewpoint. The unit-cube convention is affinely related to the centered-window convention in Hochman's Definitions 1.10–1.12. The later proof must check the boundary conventions when translating between them; the elementary symbolic theorem does not perform that translation.

Finally, for a law $\widetilde{Q}$ on full-space measure/point pairs (λ, x) with $x \in \operatorname{supp}\lambda$, its **discrete centering** and its **continuous centering at base b** are

$$\operatorname{cent}_0 \widetilde{Q} = ((\lambda, x) \mapsto T_x^*\lambda)_*\widetilde{Q}, \qquad \operatorname{cent}_b \widetilde{Q} = \frac{1}{\log b} \int_0^{\log b} (S_t^*)_* \operatorname{cent}_0 \widetilde{Q}\, dt.$$

These formulas define operations, not their conclusions. Constructing a compatible full-space extension of a geometric CP-distribution and proving that its continuous centering is a fractal distribution are later tasks.
I.12 makes no theorem claim that an arbitrary tangent distribution is scale-invariant or quasi-Palm. It also makes no theorem claim that every geometric CP-distribution centers to a fractal distribution, or conversely. The later full-proof owner IV.2 must establish all of the following from its own full proofs:
(1) the quasi-Palm property and its compatibility with scale invariance;
(2) the theorem that, at μ-almost every base point, every scenery accumulation distribution is a restricted fractal distribution;
(3) geometric base-b CP-distributions with their cell-boundary discontinuities; and
(4) continuous and discrete centering, generation, and ergodic decomposition.
The scope locators are Hochman [Hoc13, Definitions 1.1–1.5, pp. 4–5, Definition 1.10 and Theorem 1.7, p. 6, Definitions 1.11–1.13 and Theorems 1.14–1.15, p. 7, and Theorems 5.1–5.2, p. 33]. Furstenberg's definitions supply historical CP-chain concordance [Fur08, Definitions 1.2–1.5 and 5.1–5.2, pp. 407 and 419–420]. None of these source results is asserted or used in an I.12 F proof.

I.12.9 How the proofs fit together

The proved dependency order is

$$\begin{gathered} \text{L-I.12.01} \longrightarrow \text{L-I.12.03}, \\ \text{L-I.12.02} \longrightarrow \text{L-I.12.04}, \\ \text{L-I.12.01} \longrightarrow \text{L-I.12.05} \longrightarrow \text{T-I.12.06} \longrightarrow \text{C-I.12.07}, \\ \text{D/R-I.1–I.4}, \text{D/R-I.7–I.8} \longrightarrow \text{T-I.12.08}. \end{gathered} \tag{I.12.102}$$

L-I.12.03 uses L-I.12.01 only for two nested compact-probability spaces. L-I.12.04 uses the tangent definition and a locally displayed continuity-set argument. T-I.12.06 does not assume global continuity of M; equations (I.12.54) and (I.12.63)–(I.12.68) discharge the exact discontinuity. The full fractal-distribution roadmap is not used in any of the chapter's proofs.

I.12.10 Intuition and strategy

The following discussion explains the geometric ideas behind the proofs.

I.12.10.1 A tangent measure and a tangent distribution answer different questions

A tangent measure freezes one selected sequence of spatial scales. One chooses radii, renormalizes mass, and asks for a single limiting measure on all of Euclidean space. The freedom in the mass constants is why tangent measures form a cone.

A tangent distribution instead averages the entire zooming orbit over a long time interval and then takes a limit in a space whose points are themselves probability measures. It records how often different local pictures occur. A measure can therefore have many tangent measures while generating one tangent distribution, or can fail to generate a unique distribution at all.

I.12.10.2 Why the observation window causes trouble

To keep every scenery in one compact state space, the view is cut down to a fixed cube and normalized to mass one. A small movement of mass across the boundary of that cube can then produce a jump in the normalized view. The problem is not dilation itself; dilation is continuous before restriction. The discontinuity is created by the hard window.

This is why the proofs do not turn "Cesaro averages are nearly invariant" directly into "every scenery limit is invariant." That tempting argument is valid only when the dynamics is continuous at almost every state seen by the limiting law. The boundary-null bridge isolates one elementary case in which normalized restriction is continuous.

I.12.10.3 The marked point is the conditional-probability coordinate

In symbolic magnification, a state contains a measure and a point sampled from that measure. The first digit of the point selects a cylinder; the measure is conditioned on that cylinder and the selected digit is deleted. The word "adapted" means that the point remains distributed according to the measure displayed next to it.

This coupling is stronger than keeping an arbitrary measure and arbitrary point in the same pair. It supplies the estimate that makes the elementary theory work: the probability of selecting a cell whose mass is below a is at most the number of first-level cells times a.

I.12.10.4 Why support pairs were not declared compact

It is easy to mark a point with an atom of mass $1/n$. The point belongs to the support at every finite stage, but the atom disappears in the limit. A compactness argument that treats support pairs as a closed state space is therefore incorrect.

The repaired construction keeps the full product of measures and points as the compact ambient space. Magnification is discontinuous on zero-mass selected cells, but adapted laws never choose such a cell. Better still, they choose very small cells with uniformly small probability. That quantitative fact is what permits the endpoint-cancellation argument to survive weak limits.

I.12.10.5 Bernoulli measures are stationary under symbolic zooming

For a Bernoulli measure, conditioning on the first digit and deleting it returns the original measure. The scenery of the measure coordinate is therefore completely stationary. The point coordinate still moves by the shift, but its law is shift invariant.

The same product formula that gives stationarity also gives dimension. A typical length-n cylinder has logarithmic mass approximately $-nh(p)$, while its logarithmic diameter is $n \log \rho$. Their quotient is the entropy divided by the geometric contraction rate.

I.12.10.6 Three certificates, one exponent

The capstone deliberately separates three kinds of evidence.

- Coverings see how the total t-cost of all level cylinders changes with depth. This is naturally an upper-bound mechanism.
- A Frostman estimate controls how much mass can enter a small ball. It gives a lower dimension bound and, after integrating over scales, finite Riesz energy below the critical exponent.
- Entropy records the typical exponential loss of mass, while the Lyapunov exponent records the typical exponential loss of diameter. Their quotient is a local dimension.

For the canonical separated self-similar measure, the choice $p_i = r_i^s$ makes the entropy identity exact rather than asymptotic. The three methods agree because they are measuring the same multiplicative tree from three different directions: all cylinders, mass in balls, and a typical branch.

I.12.10.7 What the later CP/fractal-distribution theory adds

The elementary chapter produces compactness and adapted invariant limits in a symbolic model. The later theory has to prove substantially more. It must control geometric cell boundaries, convert discrete magnification to continuous centering, and impose a spatial Palm-type condition that says the view remains statistically consistent after moving to a typical nearby point.

The closing formal section defines quasi-Palm distributions, geometric CP-distributions and centering. The assertions just described are structural theorems, not consequences of those definitions. They are postponed to IV.2 because using them here would hide the hardest part of the scenery theory inside the word "invariant."

I.12.10.8 The role of I.12 in the series

Earlier chapters introduced dimension through covers, measures, energies, symbolic coding, and pressure. I.12 supplies the dynamical act of zooming. Later volumes can then ask whether entropy grows under convolution, whether projections preserve dimension, whether nonlinear maps change a scenery distribution, and whether Fourier decay can be read from multiscale structure.

The chapter is therefore a junction, not a survey endpoint. Its proved exports are intentionally elementary enough to be reliable premises; the deep bridge theorems remain attached to their later full-proof owners.

I.12.11 Exercises

Every exercise uses only accepted incoming nodes or definitions/results proved in the I.12 formal notes. No exercise is a premise for a theorem.

I.12.11.1 E-I.12.01 — A concrete weak metric

Let K be compact metric and let (f_j) be dense in the closed unit ball of $C(K)$. Prove directly that

$$d(\mu,\nu) = \sum_{j\geq 1} 2^{-j-1} \left| \int f_j \, d\mu - \int f_j \, d\nu \right|$$

is a metric on $\mathscr{P}(K)$ and induces weak convergence.

I.12.11.2 E-I.12.02 — A periodic empirical distribution

Let $T : K \to K$ be continuous and suppose $T^q z = z$. Compute the weak limit of $A_N^T(z)$ and verify its invariance without invoking L-I.12.01.

I.12.11.3 E-I.12.03 — Flow averages over shifted intervals

For a continuous flow (T_t) on compact K, define

$$A_{R,a}(z) = \frac{1}{R} \int_a^{a+R} \delta_{T_t z} \, dt.$$

Show that if $a_R/R \to 0$, then $A_{R,a_R}(z) - A_R^{\mathsf{T}}(z)$ tends weakly to zero as $R \to \infty$.

I.12.11.4 E-I.12.04 — Tangents of Lebesgue measure

Let $\mathscr{L}^d$ be Lebesgue measure on $\mathbb{R}^d$. Determine $\mathrm{Tan}(\mathscr{L}^d, x)$ for every x.

I.12.11.5 E-I.12.05 — Existence under Ahlfors regularity

Suppose μ is s-Ahlfors regular on its support, $s > 0$, and $x \in \mathrm{supp}(\mu)$. Show directly that $\mathrm{Tan}(\mu, x)$ is nonempty. Use the normalization

$$\lambda_r = \frac{(T_{x,r})_* \mu}{\mu(B(x,r))}.$$

I.12.11.6 E-I.12.06 — Why the boundary-null clause is necessary

In R, let $B = [-1, 1]$ and

$$\lambda_n = \delta_0 + \delta_{1+1/n}.$$

Find the vague limit of λ_n, compute every normalized restriction $N_B(\lambda_n)$, and compute the normalized restriction of the limit.

I.12.11.7 E-I.12.07 — The scenery of Lebesgue measure

For $\mu = \mathscr{L}^d$ on $\mathbb{R}^d$, compute $\mu_{x,t}^{\square}$ and $< \mu >_{x,R}$ for every x, t, R. Identify the unique tangent distribution.

I.12.11.8 E-I.12.08 — Support pairs are not closed

Let $\omega \neq \eta$ in Σ and

$$\nu_n = (1 - n^{-1})\delta_\eta + n^{-1}\delta_\omega.$$

Verify that every (ν_n, ω) is admissible but its limit is not. Explain exactly which attempted compactness proof this invalidates.

I.12.11.9 E-I.12.09 — Composition of components

If u, v are finite words and $\nu([uv]) > 0$, prove

$$(\nu_u)_v = \nu_{uv}.$$

Deduce the iterate formula (I.12.48).

I.12.11.10 E-I.12.10 — The exact small-cell identity

Let Q be adapted and let $q(\nu, \omega) = \nu([\omega_1])$. Prove

$$Q\{q < a\} = \int_{\mathscr{P}(\Sigma)} \sum_{\substack{i \in \Lambda \\ \nu([i]) < a}} \nu([i])\, d\bar{Q}(\nu),$$

and recover the bound $Q\{q < a\} \leq \#\Lambda a$.

I.12.11.11 E-I.12.11 — Closedness of adaptedness

Give a self-contained proof that

$$J_F(\nu) = \int F(\nu, \eta)\, d\nu(\eta)$$

is continuous for $F \in C(\mathscr{P}(\Sigma) x \Sigma)$. Use it to prove that the set of adapted probabilities is weakly closed and hence compact.

I.12.11.12 E-I.12.12 — Bernoulli scenery and entropy

For a Bernoulli probability β_p, prove directly that every component $(\beta_p)_u$ equals β_p. Then prove its symbolic exact-dimensionality formula in the metric d_ρ.

I.12.11.13 E-I.12.13 — A boundary-null tangent computation

Let μ be a nonnegative Radon measure, $x \in \operatorname{supp}(\mu)$, $c_n > 0$, and $r_n \downarrow 0$. Assume

$$c_n (T_{x, r_n})_* \mu \longrightarrow c\mathscr{L}^d$$

vaguely for some $c > 0$. Compute the corresponding limit of the normalized sceneries at $t_n = -\log r_n$.

I.12.11.14 E-I.12.14 — The canonical Moran energy threshold

In the SSC capstone, prove from the Frostman bound that $I_t(\mu) < \infty$ for $0 < t < s$. Then use $p_i = r_i^s$ to verify $h(p) = s\chi(p)$ without appealing to any dimension theorem.

Solutions to Exercises

Solutions to Chapter I.1

References to numbered results below refer to Chapter I.1.

Solution 1 — Similarities

The assumed equality makes f λ-Lipschitz. Theorem 4.1 gives

$$\mathscr{H}^s(f(E)) \leq \lambda^s \mathscr{H}^s(E). \tag{S1.1}$$

The equality also makes f injective and makes $f^{-1} : f(E) \to E$ λ^{-1}-Lipschitz. Applying Theorem 4.1 to the inverse gives

$$\mathscr{H}^s(E) \leq \lambda^{-s} \mathscr{H}^s(f(E)). \tag{S1.2}$$

Multiplying (S1.2) by λ^s and combining with (S1.1) yields

$$\mathscr{H}^s(f(E)) = \lambda^s \mathscr{H}^s(E).$$

The dimension equality follows either from this identity for every $s > 0$ or from Corollary 4.2. □

Solution 2 — Hölder maps

Let $(U_i)_i$ be a δ-cover of E and set $V_i = f(E \cap U_i)$. Then $(V_i)_i$ covers $f(E)$ and

$$\operatorname{diam} V_i \leq L(\operatorname{diam} U_i)^\theta \leq L\delta^\theta.$$

For every $t > 0$,

$$\begin{aligned} \mathscr{H}^t_{L\delta^\theta}(f(E)) &\leq \sum_i (\operatorname{diam} V_i)^t \\ &\leq L^t \sum_i (\operatorname{diam} U_i)^{\theta t}. \end{aligned}$$

Taking the infimum over δ-covers and then letting $\delta \downarrow 0$ gives

$$\mathscr{H}^t(f(E)) \leq L^t \mathscr{H}^{\theta t}(E). \tag{S2.1}$$

If $t > \theta^{-1} \dim_{\mathrm{H}}(E)$, then $\theta t > \dim_{\mathrm{H}}(E)$, so Theorem 3.3 gives $\mathscr{H}^{\theta t}(E) = 0$. Equation (S2.1) gives $\mathscr{H}^t(f(E)) = 0$. Taking the infimum over such t proves

$$\dim_{\mathrm{H}} f(E) \le \theta^{-1} \dim_{\mathrm{H}} E.$$

The statement is also valid when the right side is infinite. □

Solution 3 — Snowflaked metrics

For $a, b \ge 0$ and $0 < \theta \le 1$,

$$(a+b)^\theta \le a^\theta + b^\theta. \tag{S3.1}$$

If $a+b > 0$, divide by $(a+b)^\theta$ and put $u = a/(a+b)$, $v = b/(a+b)$. Since $u, v \in [0,1]$, $u^\theta \ge u$ and $v^\theta \ge v$, so $u^\theta + v^\theta \ge u + v = 1$; this proves (S3.1). The case $a = b = 0$ is direct. Applying (S3.1) to the triangle inequality for d proves

$$d_\theta(x,z) \le d_\theta(x,y) + d_\theta(y,z).$$

The other metric axioms follow directly, so d_θ is a metric.

For every $U \subset X$, monotonicity and continuity of $r \mapsto r^\theta$ give

$$\operatorname{diam}_{d_\theta} U = (\operatorname{diam}_d U)^\theta. \tag{S3.2}$$

Thus a family is a δ-cover in d_θ exactly when it is a $\delta^{1/\theta}$-cover in d, and its s-cost in d_θ equals its θs-cost in d. Hence

$$\mathscr{H}^s_{d_\theta,\delta}(E) = \mathscr{H}^{\theta s}_{d,\delta^{1/\theta}}(E).$$

Letting $\delta \downarrow 0$ yields

$$\mathscr{H}^s_{d_\theta}(E) = \mathscr{H}^{\theta s}_{d}(E).$$

Taking the infimum of the zero-measure exponents on both sides proves

$$\dim_{\mathrm{H}}^{d_\theta} E = \theta^{-1} \dim_{\mathrm{H}}^{d} E,$$

with the same conclusion for dimensions zero or infinity. □

Solution 4 — Open subsets of the line

Every nonempty open $U \subset R$ contains a nondegenerate closed interval $[a,b]$. Proposition 5.1 and monotonicity give

$$\dim_{\mathrm{H}} U \ge \dim_{\mathrm{H}}[a,b] = 1. \tag{S4.1}$$

On the other hand,

$$\mathbb{R} = \bigcup_{n=1}^{\infty} [-n,n].$$

Proposition 5.1 and countable stability give $\dim_{\mathrm{H}}(R) = 1$. Since $U \subset R$, monotonicity gives $\dim_{\mathrm{H}}(U) \le 1$, which together with (S4.1) proves equality. □

Solution 5 — Restricting the measure to the set

Define a finite Borel measure ν by

$$\nu(A) = \mu(A \cap E)$$

for every Borel $A \subset X$. Then ν is carried by E, $\nu(E) = \mu(E)$, and, for $x \in E$ and $0 < r \leq r_0$,

$$\nu(\overline{B}(x,r)) = \mu(E \cap \overline{B}(x,r)) \leq Cg(r).$$

Theorem 4.3 applied to ν gives

$$\mathscr{H}^g(E) \geq \frac{\nu(E)}{C} = \frac{\mu(E)}{C}.$$

□

Solution 6 — Positive subsets

Define $\nu(B) = \mu(B \cap A)$ for Borel $B \subset X$. The measure ν is carried by A, has $\nu(A) = \mu(A) > 0$, and, for $x \in A \subset E$,

$$\nu(\overline{B}(x,r)) \leq \mu(\overline{B}(x,r)) \leq Cg(r) \qquad (0 < r \leq r_0).$$

Theorem 4.3 gives $\mathscr{H}^g(A) \geq \mu(A)/C$. □

Solution 7 — Countable modifications

Proposition 5.2 gives $\dim_{\mathrm{H}}(N) = 0$. Theorem 3.4 therefore gives

$$\dim_{\mathrm{H}}(E \cup N) = \max\{\dim_{\mathrm{H}} E, \dim_{\mathrm{H}} N\} = \dim_{\mathrm{H}} E.$$

For the failure with closure, take $E = \emptyset$ and $N = Q \cap [0,1]$. Then $\dim_{\mathrm{H}}(E \cup N) = 0$, while

$$\dim_{\mathrm{H}}(E \cup \overline{N}) = \dim_{\mathrm{H}}[0,1] = 1.$$

□

Solution 8 — Products with countable sets

Write $F = \{y_1, y_2, \ldots\}$, allowing a finite enumeration if F is finite. By Proposition 5.4,

$$\dim_{\mathrm{H}}(E \times \{y_j\}) = \dim_{\mathrm{H}} E \qquad (j \geq 1).$$

Since

$$E \times F = \bigcup_j E \times \{y_j\},$$

countable stability gives

$$\dim_H(E \times F) = \sup_j \dim_H(E \times \{y_j\}) = \dim_H E.$$

□

Solution 9 — Pinned distance sets

The reverse triangle inequality gives

$$|d(x_0, x) - d(x_0, y)| \le d(x, y) \qquad (x, y \in X).$$

Thus $x \mapsto d(x_0, x)$ is 1-Lipschitz. Its image of E is precisely $\Delta_{x_0}(E)$, so Corollary 4.2 gives

$$\dim_H \Delta_{x_0}(E) \le \dim_H E.$$

□

Solution 10 — The middle-thirds Cantor set

At level n, the set C lies in 2^n closed basic intervals, each of length 3^{-n}. Since $3^\alpha = 2$, their total α-cost is

$$2^n(3^{-n})^\alpha = 1. \tag{S10.1}$$

For every $\delta > 0$, choose n with $3^{-n} \le \delta$. Equation (S10.1) gives $\mathscr{H}^\alpha_\delta(C) \le 1$; hence

$$\mathscr{H}^\alpha(C) \le 1. \tag{S10.2}$$

The Bernoulli product measure assigns mass 2^{-n} to each level-n cylinder, and its push-forward μ is a Borel probability measure carried by C. Let $0 < r \le 1/2$, and choose $n \ge 0$ so that

$$3^{-(n+1)} < 2r \le 3^{-n}. \tag{S10.3}$$

Distinct level-n basic intervals are separated by gaps of length at least 3^{-n}. An interval of length $2r \le 3^{-n}$ therefore meets at most two of them. Consequently

$$\mu(\overline{B}(x, r)) \le 2 \cdot 2^{-n} = 2(3^{-n})^\alpha. \tag{S10.4}$$

The left inequality in (S10.3) gives $3^{-n} < 6r$; substituting this into (S10.4) yields

$$\mu(\overline{B}(x, r)) \le 2 \cdot 6^\alpha r^\alpha. \tag{S10.5}$$

For $1/2 < r \le 1$, the same inequality follows from

$$\mu(\overline{B}(x, r)) \le 1 \le 2 \cdot 6^\alpha r^\alpha.$$

Theorem 4.3 applied with $g(r) = r^\alpha$ and $C_0 = 2 * 6^\alpha$ gives

$$\mathscr{H}^{\alpha}(C) \geq (2 \cdot 6^{\alpha})^{-1} > 0. \tag{S10.6}$$

Equations (S10.2) and (S10.6) show that the critical measure is positive and finite. Theorem 3.3 then gives

$$\dim_{\mathrm{H}} C = \alpha = \frac{\log 2}{\log 3}.$$

The existence and cylinder normalisation of the fair Bernoulli product measure are available in A-RA; exact source routes are ARA-KB1, Chapter V §7 (product measure), and ARA-KA1, Chapter IX §1 (measure-theoretic probability). □

Solutions to Chapter I.2

Solution 1

Every cover counted by $N_r(E)$ is allowed for $N'_r(E)$, so

$$N'_r(E) \le N_r(E).$$

Take an ambient-centred cover of E by $N'_{r/2}(E)$ balls of radius $r/2$. Discard balls missing E. From each remaining ball choose a point of E. Every point of E in the old ball is within distance at most r of the new centre, whence

$$N_r(E) \le N'_{r/2}(E).$$

The logarithmic denominator changes by only $\log 2$ when the scale is halved. The squeeze therefore gives the same liminf and limsup.

Solution 2

For $r_k = c^k$,

$$\frac{r_{k+1}}{r_k} = c > 0.$$

If $r_{k+1} < r \le r_k$, then

$$N_{r_k}(E) \le N_r(E) \le N_{r_{k+1}}(E),$$

while

$$(-\log r_{k+1}) - (-\log r_k) = -\log c.$$

Dividing and squeezing proves the claim exactly as in L-I.2.02. For $r_k = 2^{-k^2}$,

$$\frac{r_{k+1}}{r_k} = 2^{-(2k+1)} \longrightarrow 0,$$

so the bounded-ratio hypothesis of Lemma 1.2 fails. Nevertheless, the gaps are small relative to the logarithmic scale itself:

$$\frac{-\log r_{k+1}}{-\log r_k} = \frac{(k+1)^2}{k^2} \longrightarrow 1.$$

Put $L_k = -\log r_k$ and $a_k = \log N_{r_k}(E)/L_k$. For $r_{k+1} < r \le r_k$, monotonicity gives

$$a_k \frac{L_k}{L_{k+1}} \le \frac{\log N_r(E)}{-\log r} \le a_{k+1} \frac{L_{k+1}}{L_k}.$$

All terms are nonnegative, and the multiplying ratios tend to one. Consequently the full-scale lower limit is at least $\liminf_k a_k$, and the full-scale upper limit is at most $\limsup_k a_k$, including infinite limits. The reverse inequalities follow by restricting to the sequence r_k. Thus both box dimensions are unchanged by this sampling.

Solution 3

Write $x_n = n^{-\alpha}$. The mean value theorem gives constants $c_\alpha, C_\alpha > 0$ such that

$$c_\alpha n^{-\alpha-1} \le x_n - x_{n+1} \le C_\alpha n^{-\alpha-1} \qquad (n \ge 1). \tag{S3.1}$$

Choose a sufficiently small constant $c > 0$ and put

$$m = \left\lfloor c r^{-1/(\alpha+1)} \right\rfloor.$$

For small r, (S3.1) and the choice of c ensure $x_n - x_{n+1} > 2r$ for every $n \le m$. Hence $\{x_1, \ldots, x_m\}$ is $2r$-separated, and L-I.2.01 gives

$$N_r(E_\alpha) \ge m \ge c' r^{-1/(\alpha+1)}. \tag{S3.2}$$

For the upper bound, choose a sufficiently large constant C and put

$$M = \left\lceil C r^{-1/(\alpha+1)} \right\rceil.$$

Cover the first M points individually. Partition $[0, x_M]$ into intervals of length at most r; in each interval meeting E_α, choose a point of E_α as centre. Radius-r balls about these centres cover the tail and their number is at most $x_M/r + 2$. Consequently

$$\begin{aligned} N_r(E_\alpha) &\le M + \frac{x_M}{r} + 2 \\ &\le C' r^{-1/(\alpha+1)}, \end{aligned} \tag{S3.3}$$

because $x_M = M^{-\alpha}$. Taking logarithms in (S3.2)–(S3.3) proves

$$\dim_{\mathrm{B}} E_\alpha = \frac{1}{1+\alpha}.$$

Solution 4

Every centred δ-packing of D consists of disjoint closed intervals contained in $[-\delta, 1+\delta]$. Their total diameter is at most the length of that interval, so

$$\mathscr{P}^1_\delta(D) \le 1 + 2\delta. \tag{S4.1}$$

Conversely, fix $\varepsilon \in (0,1)$. For a sufficiently large integer N,

$$r = \frac{1-\epsilon}{2N} \leq \delta.$$

The rational centres

$$q_i = \frac{2i-1}{2N}, \qquad 1 \leq i \leq N,$$

have mutual spacing $1/N > 2r$. Their closed radius-r intervals are disjoint and have total diameter

$$N(2r) = 1 - \epsilon.$$

Thus $P^1_\delta(D) \geq 1 - \varepsilon$. Letting $\varepsilon \downarrow 0$, then $\delta \downarrow 0$ in (S4.1), gives

$$\mathscr{P}^1_0(D) = 1.$$

A centred packing of $\{x\}$ contains at most one ball, so

$$\mathscr{P}^1_\delta(\{x\}) \leq 2\delta \quad \Longrightarrow \quad \mathscr{P}^1_0(\{x\}) = 0.$$

Since D is a countable union of its singletons, P^1_0 is not countably subadditive. The singleton cover in the definition of packing outer measure has total premeasure zero, so $P^1(D) = 0$.

Solution 5

Let $s_n = \#(S \cap \{1, ..., n\})$. There are 2^{s_n} permitted level-n ternary cylinders, each of diameter at most 3^{-n}, hence

$$N_{3^{-n}}(K_S) \leq 2^{s_n}. \tag{S5.1}$$

The points obtained from the permitted length-n words by setting the tail equal to zero are 3^{-n}-separated, as proved in Proposition 3.5. Therefore

$$2^{s_n} \leq N_{3^{-n}/2}(K_S). \tag{S5.2}$$

The bounded-ratio scale lemma applied to (S5.1)–(S5.2) gives

$$\underline{\dim}_{\mathrm{B}} K_S = \frac{\log 2}{\log 3} \liminf_n \frac{s_n}{n}, \qquad \overline{\dim}_{\mathrm{B}} K_S = \frac{\log 2}{\log 3} \limsup_n \frac{s_n}{n}.$$

Removing a finite prefix and shifting indices does not alter lower or upper asymptotic density. Every nonempty relatively open subset of K_S contains a permitted cylinder, which is a similarity image of such a tail set. It therefore has the same upper box dimension as K_S.

For an arbitrary countable bounded cover of K_S, close the intersections of the covering sets with K_S in K_S. Baire category gives one member with nonempty relative interior. Monotonicity and closure invariance force that member to have upper box dimension at least that of K_S. Hence

$$\overline{\dim}_{\mathrm{MB}} K_S = \overline{\dim}_{\mathrm{B}} K_S.$$

T-I.2.09 converts this equality to the claimed packing-dimension formula.

Solution 6

Every singleton has packing dimension zero, so countable stability gives $\dim_P C = 0$. Applying countable stability again,

$$\dim_P(E \cup C) = \max\{\dim_P E, \dim_P C\} = \dim_P E.$$

Take $E = \emptyset$ and $C = Q \cap [0, 1]$. The set C has both box dimensions one, whereas the empty set has both values zero. Thus countable modification can change either box dimension.

Solution 7

If radius-r balls centred on E cover E, the Hölder condition shows that radius-Lr^θ balls centred on $f(E)$ cover the image. Thus, for $L > 0$,

$$N_{Lr^\theta}(f(E)) \le N_r(E). \tag{S7.1}$$

Set $\rho = Lr^\theta$. Then

$$-\log \rho = \theta(-\log r) - \log L,$$

and division in (S7.1) followed by lower and upper limits gives the two factors θ^{-1}. If $L = 0$, the image is a singleton and the conclusions are immediate.

Solution 8

The interval has box dimension one by Theorem 1.3. Repeated application of T-I.2.13 gives

$$\dim_B [0, 1]^k = k.$$

Applying the same product theorem once more to E and the cube yields

$$\dim_B(E \times [0, 1]^k) = \dim_B E + \dim_B [0, 1]^k = s + k.$$

Solution 9

T-I.2.13 gives

$$\underline{\dim}_B(E \times F) \in [a + b, \min\{a + B, A + b\}],$$

and

$$\overline{\dim}_B(E \times F) \in [\max\{A + b, a + B\}, A + B].$$

The first interval collapses exactly when

$$\min\{B - b, A - a\} = 0,$$

and the second collapses under the same condition. Thus each interval is a singleton precisely when at least one factor has a box dimension. If exactly one factor has a box dimension and the other oscillates, the two singleton values for lower and upper product dimensions are different. The product is forced to have a box dimension when both factors do.

Solution 10

Choose $a > \dim_{\mathrm{H}} E$ and any $b > 0$. Since $\overline{\dim}_{\mathrm{B}} F = 0 < b$, there is $r_0 > 0$ with $N_r(F) \le r^{-b}$ for $r \le r_0$. Construct a positive-radius ball cover $(B(x_i, r_i))$ of E with $\sum_i r_i^a < \varepsilon$, as in Lemma 4.2. Cover F by $N_{r_i}(F)$ balls at each corresponding scale. The product cover has total $(a+b)$-cost at most

$$2^{a+b} \sum_i r_i^a < 2^{a+b} \epsilon.$$

Thus $\mathscr{H}^{a+b}(E \times F) = 0$ and $\dim_{\mathrm{H}}(E \times F) \le a + b$. First let $a \downarrow \dim_{\mathrm{H}} E$, then let $b \downarrow 0$.

Solution 11

For nonnegative u, v,

$$\max\{u, v\} \le (u^2 + v^2)^{1/2} \le \sqrt{2} \max\{u, v\}.$$

Substitute $u = d_X(x, x')$ and $v = d_Y(y, y')$. The identity map from the max-metric product to the d_2 product and its inverse are Lipschitz. Hausdorff dimension (I.1), both box dimensions (T-I.2.03), and packing dimension for Euclidean subsets (C-I.2.10) are therefore identical for the two metrics.

Solution 12

The critical-jump proof compares different exponents:

$$\mathscr{P}_\delta^t(E) \le (2\delta)^{t-s} \mathscr{P}_\delta^s(E), \qquad t > s.$$

The vanishing factor $(2\delta)^{t-s}$ is what forces the jump. At the endpoint $t = s = p$, that factor is one and supplies no information. The definition of p only locates the boundary of the zero set of exponents; it does not say whether that boundary belongs to the zero set, whether the measure there is finite, or whether it is infinite. Thus all three displayed endpoint values are consistent with the definition and with every implication proved in C-I.2.06 and T-I.2.07. An attempted universal endpoint proof fails exactly when it sets $t = s$ in the strict-exponent comparison.

Solutions to Chapter I.3

Solution 1

The product measure gives the singleton (x_0, x_0) mass $\mu(\{x_0\})^2 > 0$. Since the Riesz kernel equals $+\infty$ on the diagonal,

$$I_s(\mu) \geq \int_{\{(x_0,x_0)\}} |x-y|^{-s}\, d(\mu \times \mu)(x,y) = +\infty.$$

Thus finite energy rules out every positive-mass singleton, which is exactly nonatomicity.

Solution 2

Orthogonality gives

$$|T(x) - T(y)| = |a|\,|O(x-y)| = |a|\,|x-y|.$$

The push-forward integration formula and Tonelli therefore yield

$$\begin{aligned} I_s(T_\# \mu) &= \iint |T(x) - T(y)|^{-s}\, d\mu(x) d\mu(y) \\ &= |a|^{-s} I_s(\mu). \end{aligned}$$

Push-forward by T is a bijection from $\mathrm{Prob}(E)$ onto $\mathrm{Prob}(T(E))$. Taking the infimum of the displayed identity over $\mathrm{Prob}(E)$ and then taking reciprocals proves

$$\mathrm{Cap}_s(T(E)) = |a|^s\, \mathrm{Cap}_s(E),$$

including the extended endpoint conventions.

Solution 3

For $0 < r \leq D$,

$$r^{-s} = r^{t-s} r^{-t} \leq D^{t-s} r^{-t}.$$

All pairs in the support of $\mu \times \mu$ have distance at most D. Integrating the preceding inequality, with the kernel values on the diagonal understood in the extended sense, gives

$$I_s(\mu) \leq D^{t-s} I_t(\mu).$$

Hence finite t-energy implies finite s-energy for every smaller positive exponent. The finite-energy exponents are therefore downward closed, hence an interval with a possibly included or excluded upper endpoint.

Solution 4

If $D > 0$, then $|x - y| \leq D$ for $x, y \in E$, so every $\nu \in \mathrm{Prob}(E)$ satisfies

$$I_s(\nu) \geq D^{-s}\nu(E)^2 = D^{-s}.$$

Taking an infimum and a reciprocal gives $\mathrm{Cap}_s(E) \leq D^s$. If $D = 0$, then E is a singleton. Every probability on it has an atom, so Solution 1 gives infinite energy and zero capacity, as required.

Now suppose E is countable. If $\nu \in \mathrm{Prob}(E)$, countable additivity gives

$$1 = \nu(E) = \sum_{x \in E} \nu(\{x\}).$$

At least one summand is positive. Solution 1 gives $I_s(\nu) = +\infty$. Thus the infimum over $\mathrm{Prob}(E)$ is infinite and $\mathrm{Cap}_s(E) = 0$.

Solution 5

The integrand is nonnegative, so Tonelli and symmetry about the diagonal give

$$\begin{aligned} I_s(\lambda) &= 2\int_0^1 \int_0^x (x-y)^{-s}\,dy\,dx \\ &= \frac{2}{1-s}\int_0^1 x^{1-s}\,dx \\ &= \frac{2}{(1-s)(2-s)}. \end{aligned}$$

This is finite for every $0 < s < 1$. The energy criterion gives $\dim_{\mathrm{H}}[0,1] \geq s$ for all such s, hence $\dim_{\mathrm{H}}[0,1] \geq 1$. The reverse inequality follows because $[0,1] \subset R$, so equality holds.

Solution 6

Each digit function b_k is Borel measurable and takes values in $\{0,1\}$. The series defining T converges uniformly, so T is Borel measurable and takes values in C. Thus its push-forward is a probability carried by C. For any prescribed first n binary digits, their preimage in $[0,1)$ is a half-open interval of length 2^{-n}: the left endpoint is $\sum_{k=1}^{n} a_k 2^{-k}$. Distinct level-$n$ ternary basic intervals are disjoint, and $T(u)$ lies in the one prescribed by the first n digits of u. Consequently every such basic interval has mass 2^{-n}.

For $0 < r < 1$, choose $n \geq 0$ so that

$$3^{-(n+1)} < r \leq 3^{-n}. \tag{S6.1}$$

The level-n construction consists of 2^n closed intervals of length 3^{-n}, each of mass 2^{-n}, and distinct such intervals are separated apart from their ordering by gaps of length at least 3^{-n}. An interval of length $2r \le 2 \cdot 3^{-n}$ meets at most four level-n basic intervals. Consequently

$$\mu(\overline{B}(x,r)) \le 4 \cdot 2^{-n}.$$

Since $q = \log 2 / \log 3$, equation (S6.1) gives

$$r^q > 3^{-(n+1)q} = 2^{-(n+1)},$$

and therefore

$$\mu(\overline{B}(x,r)) \le 8r^q.$$

For $r \ge 1$, the same inequality follows from $\mu(B) \le 1 \le r^q$. Thus μ is a q-Frostman probability and the mass-distribution principle gives $\dim_{\mathrm{H}} C \ge q$.

For the reverse inequality, the 2^n level-n intervals cover C and have diameter 3^{-n}. Their total q-cost is

$$2^n (3^{-n})^q = 1.$$

For every $u > q$, the same covers have total u-cost $2^n 3^{-\nu} = (2 \cdot 3^{-u})^n$, which tends to zero. Hence $\mathscr{H}^u(C) = 0$, so $\dim_{\mathrm{H}} C \le q$. Equality follows.

Solution 7

For $\xi \ne 0$, direct integration gives

$$\begin{aligned} \hat{\lambda}(\xi) &= \int_0^1 e^{-2\pi i x \xi}\, dx \\ &= \frac{1 - e^{-2\pi i \xi}}{2\pi i \xi} = e^{-\pi i \xi} \frac{\sin(\pi \xi)}{\pi \xi}. \end{aligned}$$

The continuous value at zero is one. On $|\xi| \le 1$, the transform is bounded by one and $|\xi|^{s-1}$ is integrable because $s > 0$. On $|\xi| > 1$,

$$|\hat{\lambda}(\xi)| \le (\pi |\xi|)^{-1},$$

so the integrand is bounded by $\pi^{-2} |\xi|^{s-3}$, which is integrable because $s < 1$. Theorem 4.1 and Solution 5 now give the exact identity

$$\int_{\mathbb{R}} |\xi|^{s-1} |\hat{\lambda}(\xi)|^2 \, d\xi = \frac{2}{c_{1,s}(1-s)(2-s)}.$$

Solution 8

Since $\mu(\mathbb{R}^d) = 1$,

$$|\hat{\mu}(\xi)| \le 1.$$

On the unit ball, therefore,

$$\int_{|\xi|\le 1} |\xi|^{s-d} |\hat{\mu}(\xi)|^2 \, d\xi \le \int_{|\xi|\le 1} |\xi|^{s-d} \, d\xi < +\infty,$$

because polar integration reduces the last integral to a constant times $\int_0^1 r^{s-1} dr$. On $|\xi| > 1$, the exponent $s - d$ is negative, so

$$\int_{|\xi|>1} |\xi|^{s-d} |\hat{\mu}(\xi)|^2 \, d\xi \le \int_{\mathbb{R}^d} |\widehat{f}(\xi)|^2 \, d\xi = \int_{\mathbb{R}^d} |f(x)|^2 \, dx < +\infty$$

by Plancherel. Theorem 4.1 gives $I_s(\mu) < +\infty$.

Solution 9

Write $I = I_s(\mu)$. The potential is nonnegative and $\int U_s^\mu d\mu = I$. Markov's inequality gives

$$\mu\{x \,:\, U_s^\mu(x) > 2I\} \le \frac{1}{2}.$$

Thus

$$F = E \cap \{x \,:\, U_s^\mu(x) \le 2I\}$$

has $\mu(F) \ge 1/2$. Put $\nu = \mu|F$. For $x \in F$ and $r > 0$,

$$2I \ge U_s^\mu(x) \ge r^{-s} \mu(\overline{B}(x,r)) \ge r^{-s} \nu(\overline{B}(x,r)).$$

The mass-distribution principle, applied to ν, yields

$$\mathscr{H}^s(E) \ge \mathscr{H}^s(F) \ge \frac{\nu(F)}{2I} \ge \frac{1}{4I}.$$

Solution 10

Markov's inequality on the probability space (S^1, σ) gives

$$\begin{aligned}\sigma\{e \,:\, I_t(\mu_e) > L\} &\le L^{-1} \int_{S^1} I_t(\mu_e) \, d\sigma(e) \\ &= \frac{A_t I_t(\mu)}{L},\end{aligned}$$

where the last equality is Lemma 5.1. If $I_t(\mu)$ is finite, the upper bound tends to zero with L^{-1}. In particular, the union over integers L of the sets where $I_t(\mu_e) \le L$ has full measure, equivalently the projected energy is finite almost everywhere.

Solution 11

Since $\pi_{-e}(x) = -\pi_e(x)$, the measure μ_{-e} is the push-forward of μ_e by the isometry $u \mapsto -u$. Distances are unchanged, so

$$I_t(\mu_{-e}) = I_t(\mu_e).$$

The same isometry maps $\pi_e(E)$ onto $\pi_{-e}(E)$. Bi-Lipschitz invariance therefore gives equality of their Hausdorff dimensions.

The quotient map q in the exercise satisfies, for every Borel subset $A \subset \mathscr{G}$,

$$(q_\#\sigma)(A) = \sigma(q^{-1}(A)).$$

In particular, A is null exactly when its inverse image is null. To apply this without a measurability assumption on an arbitrary image, choose a Borel null set $N \subset S^1$ containing the exceptional directions from Theorem 5.2 and replace it by $N \cup (-N)$. This is still Borel and null. Its image under q is Borel: on either open semicircle the quotient has a continuous inverse, and finitely many such semicircles cover the circle. Moreover,

$$q^{-1}(q(N \cup (-N))) = N \cup (-N).$$

Thus the exceptional lines lie in a Borel $q_\#\sigma$-null set. Conversely, the inverse image of a null set of lines is a null set of directions. The equality of the two dimension statements follows from the isometry already proved.

Solution 12

With $e = (\cos\theta, \sin\theta)$, normalized arclength gives

$$A_1 = \frac{1}{2\pi}\int_0^{2\pi} |\cos\theta|^{-1}\,d\theta.$$

Near $\theta = \pi/2$, the mean value theorem shows that $|\cos\theta|$ is comparable to $|\theta - \pi/2|$. The integral therefore dominates a positive constant times

$$\int_0^{\varepsilon} \frac{du}{u} = +\infty.$$

Thus equation (5.4) has an infinite directional constant at $t = 1$. Even for a nonzero measure of finite planar 1-energy, the average of the projected kernel integrals

$$\iint |u - v|^{-1}\,d\mu_e(u)d\mu_e(v)$$

is infinite by Tonelli and the same directional calculation. These are written as kernel integrals here because the chapter defines Riesz energy in ambient dimension one only for exponents strictly below one. The proof of Theorem 5.2 avoids this endpoint by using a sequence $t_j < 1$.

Finally, Hausdorff dimension one is only a critical-exponent statement. It does not specify whether the one-dimensional Hausdorff measure at the critical exponent is zero, finite positive, or infinite. Therefore the dimension conclusion alone requires a separate measure argument. Here is an explicit example showing that such an implication is false. Let

$$F = \left\{ \sum_{k=1}^{\infty} a_k 2^{-k} \;:\; a_k \in \{0,1\}, \quad a_{j^2} = 0 \; (j \ge 1) \right\}.$$

This set is compact by diagonal selection of its digits and the uniform tail bound 2^{-n}. There are $2^{n-\lfloor\sqrt{n}\rfloor}$ possible prefixes of length n. The associated closed dyadic intervals of length 2^{-n} cover F, so

$$\mathscr{L}^1(F) \le 2^{-\lfloor\sqrt{n}\rfloor} \longrightarrow 0.$$

For the dimension lower bound, list the nonsquare positive integers in increasing order as $k_1, k_2, \ldots$. Map $u \in [0,1)$ to $\sum_{j\ge1} b_j(u)2^{-k_j}$, using the binary digit functions from Exercise 6, and call the resulting probability measure ν. Each admissible prefix of length n has probability $2^{-n+\lfloor\sqrt{n}\rfloor}$. Boundary points cause no extra mass: a point has at most two binary expansions, and the probability of either fixed infinite digit sequence is at most $2^{-n+\lfloor\sqrt{n}\rfloor} \to 0$.

If $2^{-n-1} < r \le 2^{-n}$, a closed interval of length $2r$ meets at most four dyadic intervals of level n. Hence

$$\nu(\overline{B}(x,r)) \le 4\,2^{-n+\lfloor\sqrt{n}\rfloor}.$$

Fix $0 < s < 1$. Since $\lfloor\sqrt{n}\rfloor - (1-s)n$ is bounded above, there is $C_s < \infty$ such that this is at most $C_s r^s$ for all $0 < r \le 1$; enlarge the constant for $r > 1$. The mass-distribution principle gives $\dim_{\mathrm{H}} F \ge s$. Letting s increase to one and using $F \subset R$ gives $\dim_{\mathrm{H}} F = 1$, although F has Lebesgue measure zero.

Solutions to Chapter I.4

Solution 1

If $a = \mu(\{x\}) > 0$, then for every $r > 0$,

$$a \le \mu(\overline{B}(x,r)) \le 1.$$

Thus

$$0 \le \frac{\log \mu(\overline{B}(x,r))}{\log r} \le \frac{\log a}{\log r}$$

for $0 < r < 1$; both extreme quantities tend to zero. The two local dimensions are zero.

If x is outside the closed support, some ball $B(x, r_0)$ has zero measure. Every smaller ball also has zero measure, so the quotient in D-I.4.01 is $+\infty$ for all sufficiently small radii. Both limits are infinite.

Solution 2

The metric inequalities give the ball inclusions

$$\overline{B}_{d_1}(x, r/C) \subset \overline{B}_{d_2}(x,r) \subset \overline{B}_{d_1}(x, r/c).$$

Take masses, negative logarithms, and divide by $-\log r$. Replacing r by r/C or r/c changes the denominator by an additive constant. The squeeze argument from Lemma 1.1 therefore gives equality of both liminf and limsup.

Solution 3

Write the countable carrier as $E = \{x_1, x_2, \ldots\}$ after removing points of zero mass. Countable additivity gives

$$1 = \sum_i \mu(\{x_i\}),$$

so μ-almost every point is an atom. Solution 1 gives both local dimensions zero almost everywhere. Theorems 1.4 and 2.3 make all four set-carried dimensions zero. If the support is compact, Corollary 4.3 applies to this exact-dimensional measure and gives entropy dimension zero.

Solution 4

For $x \in \operatorname{spt} \mu$ and small r, the assumed inequalities imply

$$\alpha + \frac{\log C}{\log r} \le \frac{\log \mu(\overline{B}(x,r))}{\log r} \le \alpha + \frac{\log c}{\log r},$$

where the direction reverses because $\log r < 0$. Both error terms tend to zero, so the local dimension is α at every support point. Thus μ is exact dimensional. Corollaries 3.1 and 4.3 show that its lower and upper Hausdorff, packing, and entropy dimensions all equal α.

Solution 5

For every $z \in [0,1]^2$ and all sufficiently small r, the area of $\overline{B}(z,r) \cap [0,1]^2$ lies between one quarter of πr^2 and πr^2. Hence

$$\frac{\log \mu(\overline{B}(z,r))}{\log r} \longrightarrow 2.$$

Thus μ is exact dimensional of dimension two. Its first-coordinate push-forward is normalized Lebesgue measure on $[0,1]$. The length of a small ball intersected with the interval lies between r and $2r$, so its local dimension is one at every point of the interval. This proves the claim and gives a strict drop from two to one.

Solution 6

Weighted Jensen's inequality for the concave logarithm gives

$$\sum_{i\,:\,p_i>0} p_i \log(1/p_i) \le \log\left(\sum_{i\,:\,p_i>0} p_i(1/p_i)\right) \le \log N.$$

This is the desired bound. Equality in both Jensen and the final cardinality bound holds exactly when all N probabilities are positive and equal to $1/N$.

In the proof of (4.4), each conditional probability vector has at most M nonzero entries, so its entropy is at most $\log M$. Averaging these bounds over atoms of P gives

$$H(\mu, \mathscr{P} \vee \mathscr{Q}) \le H(\mu, \mathscr{P}) + \log M.$$

Since Q is coarser than the common refinement, refinement monotonicity recovers (4.1).

Solution 7

On the support of μ_i, sufficiently small balls see only the ith component. Multiplication of ball mass by p_i does not change a local dimension. Hence the mixture has local dimension α_i on a set of mass p_i. The dictionaries give

$$\underline{\dim}_{\mathrm{H}} \mu = \underline{\dim}_{\mathrm{P}} \mu = \min_i \alpha_i,$$

and

$$\overline{\dim}_{\mathrm{H}}\mu = \overline{\dim}_{\mathrm{P}}\mu = \max_i \alpha_i.$$

Iterating the finite-mixture entropy inequality changes the scale entropy by at most the fixed label entropy $-\sum_i p_i \log p_i$. More precisely, it compares the mixture entropy with $\sum_i p_i H(\mu_i, \mathscr{D}_n)$. Corollary 4.3 and division by the scale denominator give

$$\dim_{\mathrm{e}} \mu = \sum_i p_i \alpha_i.$$

Solution 8

The point mass is exact dimensional of dimension zero, while normalized Lebesgue measure on $[1, 2]$ is exact dimensional of dimension one. Their supports are positively separated, so Theorem 5.1 gives

$$\underline{\dim}_{\mathrm{H}}\mu = \underline{\dim}_{\mathrm{P}}\mu = 0,$$

$$\overline{\dim}_{\mathrm{H}}\mu = \overline{\dim}_{\mathrm{P}}\mu = 1,$$

and

$$\dim_{\mathrm{e}} \mu = 1 - p.$$

Thus positive-mass carriers select the minimum component dimension, full-mass carriers select the maximum, and entropy takes the weighted average.

Solution 9

At a fixed level, a cube of the original grid meets at most a dimension- and base-dependent constant number of cubes of the translated grid, and the same holds with the roles reversed. The bounded-intersection inequality (4.1) therefore gives

$$|H(\mu, \mathscr{D}_n^{(b)}) - H(\mu, v + \mathscr{D}_n^{(b)})| \leq C_{b,d}$$

for every n. After division by $n \log b$, the error tends to zero. Both the lower and upper entropy dimensions agree.

Solution 10

Periodicity gives the ordinary Cesaro limit

$$\frac{1}{n}\sum_{k=1}^{n} h(p_k) \longrightarrow \frac{1}{m}\sum_{k=1}^{m} h(p_k).$$

Thus the two quantities a, b in Theorem 5.3 coincide with

$$\alpha = \frac{h(p_1) + \cdots + h(p_m)}{m \log 3}.$$

The same theorem gives lower and upper local dimension α almost everywhere, so the measure is exact dimensional, and it gives entropy dimension α as well.

Solution 11

Set $p = 1/4$ and $q = 1/2$ in Corollary 5.4. Since $h(1/2) = \log 2$, the two almost-sure local dimensions are

$$a = \frac{-\frac{1}{4}\log\frac{1}{4} - \frac{3}{4}\log\frac{3}{4}}{\log 3}, \qquad b = \frac{\log 2}{\log 3}.$$

The lower and upper local dimensions equal a and b almost everywhere. Both Hausdorff measure dimensions equal a, both packing measure dimensions equal b, and the lower and upper entropy dimensions equal a and b. Because h is strictly increasing on $(0, 1/2)$, $a < b$; the local limit does not exist almost everywhere, so the measure is not exact dimensional.

Solution 12

For the first direction, mix two positively separated exact-dimensional measures of dimensions $\alpha < \beta$. Theorem 5.1 gives

$$\overline{\dim}_{\mathrm{H}}\mu = \beta > \alpha = \underline{\dim}_{\mathrm{P}}\mu.$$

For the reverse direction, take the oscillating inhomogeneous Bernoulli measure from Solution 11. Equation (5.18) gives

$$\overline{\dim}_{\mathrm{H}}\mu = a < b = \underline{\dim}_{\mathrm{P}}\mu.$$

Thus neither middle dimension controls the other.

Solutions to Chapter I.5

Solution 1

Let $0 < r < R \le \sqrt{d}$ and $x \in [0,1]^d$. In each coordinate choose the direction from x in which at least length $1/2$ remains inside the cube. The corresponding orthant of the ball of radius $R/(2\sqrt(d))$ lies in the cube and in $B(x,R)$. Its Lebesgue volume is $2^{-d}\omega_d(R/(2\sqrt{d}))^d$, where ω_d is the volume of the Euclidean unit ball. Any cover of the local piece by M balls of radius r therefore has $M\omega_d r^d$ at least this large. For the upper bound, subdivide a cube enclosing $B(x,R)$ into cells of diameter at most r, and choose a centre in each cell meeting the local piece. There are at most a dimension-dependent constant times $(R/r)^d$ such cells. Thus

$$c_d(R/r)^d \le N_r([0,1]^d \cap B(x,R)) \le C_d(R/r)^d.$$

This includes all boundary points and the full admissible outer-radius range. Thus both Assouad and lower dimensions are d; modified lower dimension is squeezed between them. The global estimate $N_r([0,1]^d) \asymp r^{-d}$ gives both box dimensions. Restricting to $r = \mathbb{R}^{1/\theta}$ gives both spectra equal to d for every θ.

Solution 2

Let $0 < r < R < r_0/2$. First take any centred cover of $F \cap B(x,R)$ by M balls of radius r. Each centre belongs to F, so the upper mass bound and subadditivity give

$$cR^\alpha \le \mu(B(x,R)) \le MCr^\alpha.$$

Taking the least possible M proves $N_r(F \cap B(x,R)) \ge cC^{-1}(R/r)^\alpha$.

For the upper bound, take a maximal r-separated subset $\{x_1,\dots,x_N\}$ of $F \cap B(x,R)$. Its radius-r balls cover the local piece and its radius-$r/2$ balls are disjoint. They are contained in $B(x,2R)$, so the lower and upper mass bounds give

$$Nc(r/2)^\alpha \le \mu(B(x,2R)) \le C(2R)^\alpha,$$

Thus $N_r(F \cap B(x,R)) \le N \le C'(R/r)^\alpha$. For completeness, large outer scales do not create a gap. Fix $0 < R_1 < \min\{r_0/4, \operatorname{diam} F\}$ when F is nondegenerate. For $R \ge R_1$ and $r < R_1$, the mass lower bound in $B(x,R_1)$ gives

$$N_r(F \cap B(x,R)) \ge \frac{cR_1^\alpha}{Cr^\alpha} \ge \frac{c}{C}\left(\frac{R_1}{\operatorname{diam} F}\right)^\alpha (R/r)^\alpha.$$

For the upper bound, disjoint radius-$r/2$ balls centred in F have total mass at most one, so a maximal separated set has at most $c^{-1}(2/r)^{\alpha}$ points. This bounds the local covering number by a constant times $(R/r)^{\alpha}$ for $R \geq R_1$. When $r \geq R_1$, compactness bounds the covering number by $N_{R_1/2}(F)$ up to a fixed recentering factor, while $1 \leq R/r \leq \operatorname{diam}(F)/R_1$. Adjusting constants covers this last case. If F is a singleton, the displayed mass assumptions force $\alpha = 0$, and the conclusion follows directly. Consequently $\dim_{\mathrm{L}} F = \dim_{\mathrm{A}} F = \alpha$. T-I.5.03 squeezes all four dimensions between these two values. The same local estimates at $r = \mathbb{R}^{1/\theta}$ make both spectra equal to α.

Solution 3

Every local covering number of F is at most m, so no positive lower or box exponent can persist as r tends to zero. The Assouad upper estimate at every exponent $s > 0$ holds because

$$N_r(F \cap B(x, R)) \leq m \leq m(R/r)^s.$$

Thus all dimensions and spectra are zero, and the modified lower dimension is zero because every subset is finite.

Every strict miniset has at most m points. If a Hausdorff limit had $m+1$ distinct points, choose disjoint balls about them. Every sufficiently close approximating set would meet all $m+1$ balls and hence contain at least $m+1$ points, a contradiction. Every strict microset is therefore finite and has Hausdorff dimension zero. At least one strict microset exists: send one point to the centre of the window while the magnification tends to infinity and pass to a Hausdorff-convergent subsequence. Hence $\dim_* F = 0$.

Solution 4

At the isolated point 2, choose $R < 1/2$. Then

$$N_r(F \cap B(2, R)) = 1$$

for every $0 < r < R$, so $\dim_{\mathrm{L}} F = 0$. The subset $[0,1]$ has lower dimension one, hence $\dim_{\mathrm{ML}} F \geq 1$. The ambient Assouad bound gives $\dim_{\mathrm{ML}} F \leq \dim_{\mathrm{A}} F \leq 1$, while monotonicity and the interval subset give $\dim_{\mathrm{A}} F \geq 1$. Therefore

$$\dim_{\mathrm{L}} F = 0, \qquad \dim_{ML} F = \dim_{\mathrm{A}} F = 1.$$

Lower dimension is the nonmonotone quantity: it drops from one to zero under the enlargement. Assouad and modified lower dimensions remain one.

Solution 5

Singletons are immediate. Otherwise fix $0 < \theta < 1$, take R sufficiently small, put $r = \mathbb{R}^{1/\theta}$, and write

$$A' = T(F) \cap B(Tx, R), \qquad A = T^{-1}(A'), \qquad B_{\pm} = F \cap B(x, K^{\pm 1}R).$$

Thus $B_- \subset A \subset B_+$, and all these sets are nonempty. A centred r/K-cover of A maps to a centred r-cover of A'. Recentering from B_+ to A and then using (1.4b) gives

$$\begin{aligned} N_r(A') &\le N_{r/K}(A) \le N_{r/(2K)}(B_+) \\ &\le D_d(2K^{1+1/\theta}) N_{(KR)^{1/\theta}}(B_+). \end{aligned}$$

If s is an admissible upper-spectrum exponent for F, the last quantity is at most

$$CD_d(2K^{1+1/\theta})K^{s(1-1/\theta)}(R/r)^s.$$

For a lower bound, a centred r-cover of A' pulls back to a centred Kr-cover of A. Recentering in B_- yields $N_{2Kr}(B_-) \le N_r(A')$. Since $2Kr/(R/K)^{1/\theta} = 2K^{1+1/\theta}$, equation (1.4b) gives

$$N_r(A') \ge D_d(2K^{1+1/\theta})^{-1} N_{(R/K)^{1/\theta}}(B_-).$$

For every admissible lower-spectrum exponent s of F, this is at least

$$cD_d(2K^{1+1/\theta})^{-1}K^{s(1/\theta-1)}(R/r)^s.$$

The constants are independent of x, R. Both applications to F are valid when $KR < \min\{1, \operatorname{diam} F\}$. The omitted outer scales, bounded away from zero, cause no gap: their inner radii are bounded below, so compactness and (1.4a) give a uniform upper covering bound, while $N_r \ge 1$ gives a uniform positive lower bound after division by the bounded quantity $(R/r)^s$. Thus the same exponents are admissible for $T(F)$. Taking the infimum for upper exponents and the supremum for lower exponents gives

$$\dim_{\mathrm{A}}^{\theta} T(F) \le \dim_{\mathrm{A}}^{\theta} F, \qquad \dim_{\mathrm{L}}^{\theta} T(F) \ge \dim_{\mathrm{L}}^{\theta} F.$$

Apply the same argument to T^{-1}, with ambient dimension e in the covering constants, to obtain equality. Only the *inner* radius was changed within a fixed local piece. No bounded-ratio comparison of covering numbers under an arbitrary enlargement of the outer ball was used.

Solution 6

For the max metric,

$$B_{E\times F}((x, y), r) = B_E(x, r) \times B_F(y, r).$$

Therefore

$$(\mu \times \nu)(B((x, y), r)) = \mu(B(x, r))\nu(B(y, r)) \asymp r^{\alpha+\beta}.$$

The product is $(\alpha + \beta)$-Ahlfors regular for the max metric. The mass bounds and separated-set argument in Solution 2 use only the triangle inequality, so they give lower and Assouad dimensions and both spectra equal to $\alpha + \beta$ directly in this metric. The modified lower dimension is squeezed between lower and Assouad dimension.

For completeness, write $d_{\max}$ and d_2 for the max and Euclidean product metrics. Since $d_{\max} \le d_2 \le \sqrt{2}d_{\max}$,

$$B_{\max}((x,y), r/\sqrt{2}) \subset B_2((x,y),r) \subset B_{\max}((x,y),r).$$

The same measure is therefore $(\alpha+\beta)$-Ahlfors regular in the Euclidean product metric. Solution 2 gives its Hausdorff, packing, and both box dimensions, and the metric invariance proved in I.1–I.2 transfers these values to the max metric. This identifies all requested dimensions without treating a max-metric ball as a Euclidean ball.

Solution 7

The set E_N is

$$\{0\} \cup \{(N/n)^p \,:\, n \ge N\}.$$

Its positive points decrease from one to zero. By the mean-value theorem,

$$(N/n)^p - (N/(n+1))^p \le pN^p n^{-(p+1)} \le p/N$$

for $n \ge N$. Thus every point of $[0,1]$ lies within p/N of E_N, while $E_N \subset [0,1]$. Hence $d_H(E_N,[0,1]) \le p/N$. Since $N^p \to \infty$, the interval is a strict microset. Its Hausdorff dimension is one, agreeing with $\dim_{\mathrm{A}} F_p = 1$.

Solution 8

Every $x \in [0,1]$ lies between two consecutive dyadic grid points at distance at most 2^{-n} from one of them. Since $A_n \subset [0,1]$,

$$d_H(A_n,[0,1]) \le 2^{-n}.$$

Thus $A_n \to [0,1]$ in $K([0,1])$. Nevertheless each A_n has Hausdorff dimension zero while the limit has Hausdorff dimension one. Hausdorff dimension is therefore not continuous for d_H; the mass selection in T-I.5.12 is essential.

Solution 9

Since $T_k([0,1]) = [1,k+1]$, its intersection with $[0,1]$ is $\{1\}$ for every k. The similarity ratios tend to infinity and the Hausdorff limit is the boundary singleton. But

$$\dim_{\mathrm{L}}[0,1] = 1 > 0 = \dim_{\mathrm{H}}\{1\}.$$

Thus T-I.5.11 would be false if a limit confined to the boundary were admitted. The singleton is excluded by $E \cap (0,1) \ne \emptyset$.

Solution 10

If every Assouad spectrum equals α, the lower inequality becomes

$$\frac{1-\theta_2}{1-\theta_1}\alpha \le \alpha,$$

which holds because $\theta_1 < \theta_2$. The right side of the upper inequality is

$$\left(\frac{1-\theta_2}{1-\theta_1} + \frac{\theta_2-\theta_1}{1-\theta_1}\right)\alpha = \alpha.$$

For constant lower and Assouad spectra, the lower-spectrum upper coefficients sum to

$$\frac{\theta_1(1-\theta_2)+\theta_2-\theta_1}{\theta_2(1-\theta_1)} = 1.$$

The lower inequality is the same as before. All four comparisons hold.

Solution 11

Since $C \subset F$, taking $c = 1$ and $t = 0$ in D-I.5.08 shows that C is a Furstenberg miniset. On the other hand,

$$(\lambda[0,1]+t)\cap[0,1]$$

is empty, a singleton, or a compact interval for every $\lambda \ge 1$ and $t \in R$. The Cantor set contains 0 and 1 but omits $1/2$: its first ternary digit places it in $[0,1/3]\cup[2/3,1]$. It is therefore none of these sets, and is not a strict miniset. This exhibits the subset-versus-exact-intersection difference.

Solution 12

If E is a strict microset, then by definition

$$(\lambda_k F + t_k)\cap W \longrightarrow E, \qquad \lambda_k \longrightarrow \infty, \qquad E\cap W^\circ \ne \emptyset.$$

Each map $T_k(x) = \lambda_k x + t_k$ is a homothety of ratio λ_k. These are exactly the conditions of D-I.5.06, so E is a windowed weak tangent. T-I.5.10 and the hierarchy give

$$\dim_{\mathrm{H}} E \le \dim_{\mathrm{A}} E \le \dim_{\mathrm{A}} F.$$

Taking the supremum over strict microsets yields $\dim_* F \le \dim_{\mathrm{A}} F$ without any use of G-I.5.14.

Solutions to Chapter I.6

Solution 1

By definition,

$$\mathscr{H}^m = \frac{\omega_m}{2^m}\mathscr{H}^m.$$

Hence for every $r > 0$,

$$\frac{\mathscr{H}^m(E \cap \overline{B}(x,r))}{\omega_m r^m} = \frac{\mathscr{H}^m(E \cap \overline{B}(x,r))}{2^m r^m} = \frac{\mathscr{H}^m(E \cap \overline{B}(x,r))}{(2r)^m}.$$

Taking $\liminf$ gives the requested formula. Taking $\limsup$ gives the upper-density formula.

Solution 2

Let A be the set of admissible relative radii in (2.1). Since $\sup A \geq \alpha > \beta$, the number β is not an upper bound for A. Therefore some $a \in A$ satisfies $a > \beta$. If

$$\overline{B}(y, ar) \subset \overline{B}(x,r) \setminus E,$$

then the concentric ball $\overline{B}(y, \beta r)$ has the same containment.

For a geometric example, take $E = \{0\}$ in R, $x = 0$, and $r = 1$. Every $0 \leq a < 1/2$ is admissible: the interval $[1-2a, 1]$ is a closed ball of radius a disjoint from E. No closed ball of radius $1/2$ in $[-1, 1]$ avoids zero. Thus the admissible relative radii are exactly $[0, 1/2)$, and the porosity is $1/2$ without an attaining hole.

Solution 3

Fix $x \in \mathrm{cl}(E)$ and $0 < r \leq r_0/2$. Choose $x' \in E$ with $|x - x'| < r/4$. Uniform porosity at the radius $r/2$ gives

$$\mathrm{por}(E, x', r/2) \geq \alpha.$$

By Solution 2, with $\beta = 3\alpha/4$, there are $b > 3\alpha/4$ and y such that

$$\overline{B}(y, br/2) \subset \overline{B}(x', r/2) \setminus E.$$

Since

$$\frac{\alpha r}{4} < \frac{br}{2},$$

the ball $\overline{B}(y, \alpha r/4)$ lies a positive distance $br/2 - \alpha r/4$ from the complement of the larger hole. If it met $\mathrm{cl}(E)$, a sequence from E approaching the intersection point would eventually enter the larger hole, a contradiction. Finally,

$$\overline{B}(y, \alpha r/4) \subset \overline{B}(x', r/2) \subset \overline{B}(x, r).$$

Thus the closure has an $(\alpha/4)r$ hole at every required point and scale.

Solution 4

If $\mathscr{L}^d(F) > 0$, almost every point of F is a density point of E, because $F \subset E$. Fix such an x. At a scale with porosity at least α, Solution 2 supplies a hole of relative radius $\alpha/2$, so

$$\frac{\mathscr{L}^d(\overline{B}(x, r_n) \setminus E)}{\omega_d r_n^d} \geq (\alpha/2)^d.$$

This occurs along infinitely many $r_n \downarrow 0$, contradicting density one of E at x. Therefore no such positive-measure F exists.

Solution 5

Both balls lie inside $[0, 1]^d$, where normalized Lebesgue measure is a constant multiple of $\mathscr{L}^d$. Therefore

$$\frac{\lambda(\overline{B}(y, \alpha r))}{\lambda(\overline{B}(x, r))} = \frac{\omega_d(\alpha r)^d}{\omega_d r^d} = \alpha^d.$$

If $\varepsilon \geq \alpha^d$, every such inner ball is an epsilon-hole in the sense of (2.5). At interior points and all sufficiently small radii this occurs at every scale. Hence Lebesgue measure itself satisfies the corresponding weak mean-porosity condition. Requiring only one such large ε therefore cannot imply singularity.

Solution 6

Fix $p_1 = t$. Strict concavity of $-u \log u$ shows that the remaining entropy is maximized uniquely, apart from permutations among coordinates $2, \ldots, M$, when

$$p_2 = \cdots = p_M = \frac{1 - t}{M - 1}.$$

The maximal value at fixed t is

$$F(t) = h(t) + (1 - t) \log(M - 1).$$

Its derivative is

$$F'(t) = \log \frac{1 - t}{t(M - 1)} \geq 0 \quad (0 < t \leq 1/M).$$

Thus under $0 \le t \le \varepsilon \le 1/M$, the maximum occurs at $t = \varepsilon$. The maximizers are

$$p_1 = \varepsilon, \qquad p_2 = \cdots = p_M = \frac{1-\varepsilon}{M-1},$$

with the evident limiting interpretation at $\varepsilon = 0$. Their entropy is the right side of (3.15).

Solution 7

Put $b = 2^\ell$, $q = M - 1$, and

$$\mathscr{A} = \{0, \dots, b-1\}^d \setminus \{(b-1, \dots, b-1)\}.$$

If $q = 1$, then $d = \ell = 1$ and the sole allowed digit is zero. The measure $\mu = \delta_0$ has exactly the required masses and all the asserted properties. Assume henceforth that $q \ge 2$. Enumerate the q elements of $\mathscr{A}$ as $a_0, \dots, a_{q-1}$. On $[0, 1)$ with Lebesgue probability, define

$$j_n(u) = \lfloor q^n u \rfloor - q \lfloor q^{n-1} u \rfloor, \qquad \Phi(u) = \sum_{n=1}^{\infty} b^{-n} a_{j_n(u)}.$$

Each prescribed prefix of N digits j_n corresponds to one half-open interval of length q^{-N}. The map Φ is Borel, since its partial sums are Borel and converge coordinatewise. Let μ be the push-forward of Lebesgue probability under Φ.

We check the half-open convention rather than assume it. For any coordinate i, the proportion of allowed digit vectors with ith coordinate $b - 1$ is

$$\rho = \frac{b^{d-1} - 1}{b^d - 1} < 1.$$

The prefix probabilities show that the event of k consecutive such digits after any fixed position has probability ρ^k. By continuity from above, the event that this coordinate is eventually always $b - 1$ has probability zero. Take the countable union over starting positions and the finite union over coordinates. Off this null set, every coordinate tail satisfies

$$0 \le \sum_{n>N} b^{-n} (a_{j_n})_i < b^{-N}.$$

Consequently $\Phi(u)$ belongs to the half-open level-$N\ell$ cube specified by its first N vector digits, not to the next cube at an endpoint. All allowed cubes therefore have mass q^{-N}, all other cubes have mass zero, and $\mu([0, 1)^d) = 1$. This proves existence with the prescribed subdivision masses.

At every positive-mass block beginning, the designated descendant has zero mass. Hence (3.1) is at most ℓ at every block for every $0 < \varepsilon < 1$, and μ is dyadic mean $(\ell, 1)$-porous.

For every x outside the null union of zero-mass cubes, each block conditional mass equals $1/(M - 1)$. Therefore

$$\mu(Q_{N\ell}(x)) = \mu(Q_0(x))(M - 1)^{-N}.$$

As in (3.22),

$$\overline{d}_{\mu}(x) \leq \lim_{N\to\infty} \frac{-\log \mu(Q_{N\ell}(x))}{N\ell \log 2} = \frac{\log(M-1)}{\ell \log 2}.$$

This also realizes equality in the first bound of (3.18) when $\eta = 1$ at the level of the block entropy calculation.

Solution 8

If $f : \mathbb{R}^m \to \mathbb{R}^d$ is Lipschitz, then

$$f(\mathbb{R}^m) = \bigcup_{k=1}^{\infty} f(\overline{B}(0,k)).$$

Each image on the right is compact, so $f(\mathbb{R}^m)$ is sigma-compact and Borel. A countable union of such images is again Borel, and a countable union of countable unions is a countable union. Hence C_m has both asserted properties.

Let

$$s = \sup\{\mu(S) : S \in \mathscr{C}_m\}.$$

Choose $S_j \in C_m$ with $\mu(S_j) > s - 2^{-j}$ and put $S = \cup_j S_j$. Then $S \in C_m$ and $\mu(S) \geq s$; the definition of s gives the reverse inequality. Thus $\mu(S) = s$, so the supremum is attained.

Solution 9

For each j, the constant map

$$f_j : \mathbb{R}^m \longrightarrow \mathbb{R}^d, \qquad f_j(u) = x_j,$$

is Lipschitz and has image $\{x_j\}$. The countable union of these images carries μ, so μ is m-rectifiable for every $1 \leq m \leq d$, the range defined in this chapter. By uniqueness in Theorem 4.2,

$$\mu_{\mathrm{r}}^m = \mu, \qquad \mu_{\mathrm{u}}^m = 0.$$

This is why absolute continuity with respect to $\mathscr{H}^m$ must not be silently inserted into the chapter's measure convention.

Solution 10

Choose a countable union R of Lipschitz images carrying ν. Pure unrectifiability gives $\tau(R) = 0$. Since $\nu \leq \tau$,

$$\nu(\mathbb{R}^d) = \nu(R) \leq \tau(R) = 0.$$

Thus $\nu = 0$.

For uniqueness, suppose

$$\mu = \mu_r + \mu_u = \nu_r + \nu_u$$

are two decompositions. If R carries μ_r, then $\nu_u(R) = 0$, so $\mu_r = \mu|R \le \nu_r$. Hence $\nu_r - \mu_r$ is a positive measure. It is rectifiable because it is dominated by the rectifiable measure ν_r and is therefore carried by the same countable union of Lipschitz images. It is dominated by μ_u because

$$\nu_r - \mu_r = \mu_u - \nu_u \le \mu_u.$$

It is therefore purely unrectifiable: every Lipschitz image has zero μ_u-mass and hence zero $(\nu_r - \mu_r)$-mass. The first part forces $\nu_r - \mu_r = 0$. Therefore the rectifiable parts, and then the unrectifiable parts, agree.

Solution 11

Fix $u \in A$. The fiber of K over u is the interval

$$\{(u, t) \,:\, 0 \le t \le f(u)\}.$$

Its top point $(u, f(u))$ is the only point visible in the positive e direction. Hence

$$\mathrm{Vis}_e(K) = \{(u, f(u)) \,:\, u \in A\}.$$

If $f(u) > 0$, the full line meets the fiber in more than one point, so no point of that fiber is bi-visible. If $f(u) = 0$, the fiber is the singleton $\{(u, 0)\}$, which is bi-visible. Therefore

$$\mathrm{Vis}_e^2(K) = \{(u, 0) \,:\, u \in A,\ f(u) = 0\}.$$

Solution 12

The set

$$T_v = \{t > 0 \,:\, a + tv \in K\}$$

is nonempty because $v \in \pi_a(K)$. It is bounded because K is bounded, and it is bounded away from zero because $\mathrm{dist}(a, K) > 0$. It is closed in that bounded interval because K is closed. Thus T_v is compact and has a minimum $t(v)$. No point of K lies on the segment from a to $a + t(v)v$ before its endpoint, so the endpoint is visible from a.

For discontinuity, take $a = (0, 0)$ in $\mathbb{R}^2$ and

$$K = \{(1, 0)\} \cup \{(2\cos(1/n), 2\sin(1/n)) \,:\, n \ge 1\} \cup \{(2, 0)\}.$$

This set is compact and avoids a. For $v_n = (\cos(1/n), \sin(1/n))$, the first hit is $2v_n$, which tends to $(2, 0)$. In the limiting direction $v = (1, 0)$, the first hit is $(1, 0)$. Hence the first-hit map is discontinuous at v, even though every direction in $\pi_a(K)$ still has a first hit and (5.6) holds.

Solutions to Chapter I.7

Solution 1

The composition convention gives

$$K_{uv} = S_{uv}(K) = S_u(S_v(K)) = S_u(K_v).$$

Since $K_v \subset K$, applying S_u gives $K_{(uv)} \subset K_u$. For similarities,

$$\operatorname{diam}(K_{uv}) = r_{uv} \operatorname{diam}(K) = r_u r_v \operatorname{diam}(K).$$

Suppose u, v are incomparable. Remove their maximal common prefix, writing $u = qia$ and $v = qjb$ with $i \neq j$. Repeated use of $S_k(O) \subset O$ gives

$$S_{ia}(O) \subset S_i(O), \qquad S_{jb}(O) \subset S_j(O).$$

The two sets on the right are disjoint. Applying the injective map S_q therefore gives $S_u(O) \cap S_v(O) = \emptyset$.

Solution 2

Since $K = \mathscr{F}_\Phi(K)$, the triangle inequality and Lemma 1.1 give

$$\begin{aligned} d_H(A, K) &\leq d_H(A, \mathscr{F}_\Phi(A)) + d_H(\mathscr{F}_\Phi(A), \mathscr{F}_\Phi(K)) \\ &\leq d_H(A, \mathscr{F}_\Phi(A)) + c\, d_H(A, K). \end{aligned}$$

Because $c < 1$, rearrangement proves the first estimate. Theorem 1.2 and that estimate give

$$d_H(\mathscr{F}_\Phi^n(A), K) \leq c^n d_H(A, K) \leq \frac{c^n}{1-c} d_H(A, \mathscr{F}_\Phi(A)).$$

Solution 3

Induction on n gives, for every base point x_0,

$$C_{\omega|n}(x_0) = \sum_{k=1}^{n} \frac{\omega_k}{3^k} + \frac{x_0}{3^n}.$$

Letting n tend to infinity proves the coding formula. For distinct ω, τ, let $m = |\omega \wedge \tau|$ be their longest common prefix length. Then

$$|\pi_C(\omega) - \pi_C(\tau)| \le \sum_{k=m+1}^{\infty} \frac{2}{3^k} = 3^{-m}.$$

With $\gamma = \log 3/\log 2$, this is

$$|\pi_C(\omega) - \pi_C(\tau)| \le (2^{-m})^\gamma = d_\Sigma(\omega,\tau)^\gamma$$

for distinct words. The displayed series is precisely a ternary expansion whose digits lie in $\{0,2\}$. Conversely, every such expansion supplies an infinite word, so its value lies in the coding image. Surjectivity of the coding identifies this image with the attractor.

Solution 4

Since u is nonempty, the linear part A_u is a similarity of ratio $r_u < 1$; hence $\|A_u\|_{(op)} = r_u < 1$. Therefore the operator series

$$\sum_{n=0}^{\infty} A_u^n$$

converges in operator norm and satisfies

$$(I_d - A_u)\sum_{n=0}^{\infty} A_u^n = \sum_{n=0}^{\infty} A_u^n (I_d - A_u) = I_d.$$

Thus $I_d - A_u$ is invertible. A fixed point x of S_u obeys

$$(I_d - A_u)x = a_u,$$

so it is uniquely equal to $(I_d - A_u)^{-1}a_u$. Theorem 2.1 identifies this unique fixed point with $\pi_\Phi(u^\infty)$.

Solution 5

For $u = \omega|n$, contraction gives

$$|S_u(x) - S_u(x_0)| \le c^n |x - x_0|.$$

The telescoping estimate (2.7) also gives

$$|S_{\omega|n}(x_0) - \pi_\Phi(\omega)| \le \sum_{k=n}^{\infty} Mc^k = \frac{Mc^n}{1-c}.$$

Using (3.7), the product law $\beta_p \times \nu$, and the Lipschitz property of g, we obtain

$$\begin{aligned}
&\left|\int g\, d\mathscr{M}^n_{\Phi,p}(\nu) - \int g\, d\mu_{\Phi,p}\right| \\
&\le \int_\Sigma \int_{\mathbb{R}^d} L|S_{\omega|n}(x) - \pi_\Phi(\omega)|\, d\nu(x)\, d\beta_p(\omega) \\
&\le Lc^n \left(\int |x - x_0|\, d\nu(x) + \frac{M}{1-c} \right).
\end{aligned}$$

The finite first moment makes the right side finite.

Solution 6

The set Λ_+ is nonempty because the weights sum to one. Terms with zero weight vanish, so

$$\mathscr{M}_{\Phi,p}(\nu) = \sum_{i\in\Lambda_+} p_i (S_i)_\# \nu.$$

The restricted vector is a strict probability vector on Λ_+. Theorem 3.1 applied to the sub-IFS therefore gives a unique invariant probability and says that its support is K_+. Since the full operator is the same operator, this proves both assertions.

For example, take the dyadic interval IFS and weights $p_0 = 1$, $p_1 = 0$. The full attractor is $[0,1]$, whereas the positive-weight sub-IFS consists only of D_0; its attractor and the support of the invariant measure are $\{0\}$.

Solution 7

The singleton $\{0\}$ is invariant under every labelled map, so uniqueness of the attractor gives

$$K = \{0\}, \qquad \dim_{\mathrm{H}} K = 0.$$

The labelled similarity dimension solves

$$Nr^s = 1,$$

and hence

$$s(\Phi) = \frac{\log N}{\log(1/r)}.$$

For fixed r, this tends to infinity with N, while the attractor dimension remains zero. All distinct one-letter labels already form exact overlaps.

Solution 8

Induction gives

$$D_{\omega|n}(x_0) = \sum_{k=1}^{n} \frac{\omega_k}{2^k} + \frac{x_0}{2^n},$$

so the asserted formula follows on letting n tend to infinity.

Suppose two distinct sequences a, b have the same image, and let m be their first differing coordinate. After interchanging the sequences, assume $a_m = 0$ and $b_m = 1$. Equality gives

$$\sum_{k>m} \frac{a_k - b_k}{2^k} = \frac{1}{2^m}.$$

The left side is at most

$$\sum_{k>m} \frac{1}{2^k} = \frac{1}{2^m}.$$

Equality forces $a_k = 1$ and $b_k = 0$ for every $k > m$. Thus every double address has the form

$$w0\,111\ldots \quad \text{and} \quad w1\,000\ldots$$

for one finite word w. Their common value is dyadic and lies strictly between zero and one. Conversely, every dyadic point in $(0,1)$ has a finite binary expansion ending in 1; replacing that last 1 and the following zeros by 0 and an infinite tail of 1s gives the second address. The first-difference argument shows there are no others. A non-dyadic point therefore has one address. The only addresses of zero and one are, respectively, 000... and 111....

Solution 9

Let

$$O_\varepsilon = \{x \,:\, \mathrm{dist}(x, K) < \varepsilon\}.$$

If $x \in O_\varepsilon$, choose $y \in K$ with $|x - y| < \varepsilon$. Since $S_i(y) \in K$,

$$\mathrm{dist}(S_i(x), K) \le r_i|x - y| < r_i\varepsilon < \varepsilon.$$

Hence $S_i(O_\varepsilon) \subset O_\varepsilon$. Moreover, $S_i(O_\varepsilon)$ lies in the open $r_i\varepsilon$-neighborhood of $S_i(K)$. For $i \neq j$, the sum of the two neighborhood radii is at most

$$2r_{\max}\varepsilon < \delta.$$

The neighborhoods, and therefore $S_i(O_\varepsilon)$ and $S_j(O_\varepsilon)$, are disjoint. Thus O_ε is feasible.

Solution 10

For every word u, cancellation gives

$$\tilde{S}_u = Q \circ S_u \circ Q^{-1}.$$

Consequently

$$\bigcup_i \tilde{S}_i(Q(K)) = Q\left(\bigcup_i S_i(K)\right) = Q(K).$$

Attractor uniqueness gives $\widetilde{K} = Q(K)$. Taking the coding base point $Q(x_0)$ and passing to the limit gives

$$\pi_{\widetilde{\Phi}} = Q \circ \pi_\Phi.$$

The identities

$$\tilde{S}_i(\widetilde{K}) = Q(S_i(K)), \qquad \tilde{S}_i(Q(O)) = Q(S_i(O))$$

show that SSC and OSC are invariant. Equality $\tilde{S}_u = \tilde{S}_v$ is equivalent to $S_u = S_v$, so exact overlap is invariant.

For relative maps,

$$\tilde{S}_u^{-1} \circ \tilde{S}_v = Q \circ (S_u^{-1} \circ S_v) \circ Q^{-1}.$$

To check continuity in the coefficient metric explicitly, if $T(x) = Ax + a$, then

$$(Q \circ T \circ Q^{-1})(x) = RAR^{-1}x + qRa + b - RAR^{-1}b.$$

Both coefficients depend continuously on those of T, and the inverse operation is conjugation by Q^{-1}. Thus conjugation is a homeomorphism of the similarity group and fixes the identity. It preserves isolation from the nonidentity relative maps, proving WSC invariance.

Write $S_u(x) = A_u x + a_u$. If $A_u = A_v$, the conjugated maps have the same linear part and their translation-vector difference is

$$qR(a_u - a_v).$$

Its norm is $q|a_u - a_v|$. If the original linear parts differ, so do their conjugates. Therefore every pairwise cylinder distance, and then its minimum, obeys

$$\Delta_n(\widetilde{\Phi}) = q\Delta_n(\Phi).$$

If $\Delta_n(\Phi) \geq c^n$, then the conjugated bound holds with c when $q \geq 1$ and with qc when $0 < q < 1$, because $qc^n \geq q^n c^n$. This works either on all levels or on the same infinite set of levels. Applying the argument to Q^{-1} proves the converse, so both forms of exponential separation are invariant.

Solution 11

At level n, every dyadic cylinder has the form

$$D_u(x) = 2^{-n}x + t_u, \qquad 2^n t_u \in \mathbb{Z}.$$

Because all one-step ratios equal $1/2$, a word in $\Lambda_{(r_u)} = \Lambda_{(2^{-n})}$ also has length n. Hence every neighbor map has the form

$$D_u^{-1} \circ D_v(x) = x + m, \qquad m = 2^n(t_v - t_u) \in \mathbb{Z}.$$

The intersection $D_u(B) \cap D_v(B) \neq \emptyset$ is equivalent, after applying D_u^{-1}, to

$$B \cap (B + m) \neq \emptyset.$$

Since $B = [-1, 2]$, this implies $|m| \leq 3$. Thus every neighbor map is translation by an integer in $\{-3, -2, -1, 0, 1, 2, 3\}$. The union of neighbor families is finite, proving finite type. Lemma 5.4 then gives WSC.

Solution 12

If, for example, $v = uw$ with w nonempty, equality of the maps would imply

$$r_u = r_v = r_u r_w,$$

which is impossible because $0 < r_w < 1$. The reverse prefix relation is excluded in the same way.

The words uv and vu have the same length. They are distinct: if $|u| \le |v|$ and $uv = vu$, comparison of the first $|u|$ symbols makes u a prefix of v, contrary to the first paragraph. The other length ordering is symmetric. Since $S_u = S_v$,

$$S_{uv} = S_u \circ S_v = S_v \circ S_u = S_{vu}.$$

Put $\ell = |u| + |v|$. For $n \ge \ell$, choose any word w_n of length $n - \ell$ and set

$$a_n = uvw_n, \qquad b_n = vuw_n.$$

The words remain distinct and their maps are equal. Thus an exact overlap occurs at every level $n \ge \ell$.

Solution 13

Put $q_n = (-\log \delta_n)/n$. If the first condition holds, then

$$q_n \le -\log c$$

on infinitely many levels, so q_n cannot tend to $+\infty$.

Conversely, if q_n does not tend to $+\infty$, then some finite $M > 0$ and infinitely many n satisfy $q_n < M$. At those levels,

$$\delta_n > e^{-Mn} > \left(e^{-(M+1)}\right)^n.$$

Thus the first condition holds with $c = e^{-(M+1)}$. The argument includes the values zero and infinity under the stated logarithmic conventions.

For a nonuniform example, define

$$\delta_n = \begin{cases} e^{-n}, & n \text{ even}, \\ e^{-n^2}, & n \text{ odd}. \end{cases}$$

The lower bound with $c = e^{-1}$ holds on every even level. Given any $c = e^{-a} \in (0,1)$, however, every sufficiently large odd n satisfies $e^{-n^2} < e^{-an} = c^n$. Hence no exponential lower bound holds at every level.

Solution 14

For every ω, the ratios $r_{(\omega|n)}$ tend to zero. Hence there is a first $n \ge 1$ for which $r_{(\omega|n)} \le t$, and that prefix belongs to Λ_t. It is unique by minimality.

Choose m so large that $r_{\max}^m \le t$. Every first crossing occurs by level m, so Λ_t is finite. If u were a proper prefix of v and both belonged to Λ_t, then v^- would extend u and

$$r_{v^-} \le r_u \le t,$$

contradicting the stopping condition $t < r_{(v^-)}$. Thus Λ_t is prefix-free.

The cylinders $[u]$, $u \in \Lambda_t$, are consequently pairwise disjoint, and the unique-prefix statement says that their union is Σ. For the Bernoulli law with $p_i = r_i^s$, Lemma 4.1 gives

$$\beta_p([u]) = p_u = r_u^s.$$

Finite additivity over the cylinder partition now yields

$$1 = \beta_p(\Sigma) = \sum_{u \in \Lambda_t} \beta_p([u]) = \sum_{u \in \Lambda_t} r_u^s.$$

Solutions to Chapter I.8

Solution 1

Write $S_n = \sum_{k=1}^n X_k$ and expand S_n^4. Independence and centering annihilate every monomial containing an index exactly once. The surviving terms either use one index four times or two distinct indices twice each. There are six orderings of $X_j^2 X_k^2$ for each $j < k$. Hence

$$\begin{aligned}\mathbb{E}(S_n^4) &= \sum_{k=1}^{n} \mathbb{E}(X_k^4) + 6 \sum_{1\le j<k\le n} \mathbb{E}(X_j^2)\mathbb{E}(X_k^2) \\ &= nm_4 + 6\binom{n}{2}\sigma^4 = nm_4 + 3n(n-1)\sigma^4.\end{aligned}$$

If the variables are bounded, the last expression is at most Cn^2. Markov's inequality gives

$$\mathbb{P}\{|S_n| > \varepsilon n\} \le \frac{C}{\varepsilon^4 n^2}.$$

The probabilities are summable. Borel–Cantelli implies $|S_n| \le \varepsilon n$ eventually, for each $\varepsilon = 1/m$. Intersecting these countably many probability-one events gives $S_n/n \to 0$. Adding back the mean proves the bounded iid strong law.

Solution 2

The union of cylinders $[u]$, $u \in A_n(\varepsilon)$, is the event

$$\left|\frac{-\log p_{\omega|n}}{n} - h\right| \le \varepsilon.$$

Lemma 1.1 says that the probability of this event tends to one. Since the level-n cylinders are disjoint, its probability is exactly $\sum_{u\in A_n(\varepsilon)} p_u$.

For the upper count, every word in $A_n(\varepsilon)$ has mass at least $e^{-n(h+\varepsilon)}$, so

$$1 \ge \sum_{u\in A_n(\varepsilon)} p_u \ge \#A_n(\varepsilon)e^{-n(h+\varepsilon)}.$$

For all sufficiently large n, the total mass of the typical words is at least $1/2$. Since each such word has mass at most $e^{-n(h-\varepsilon)}$,

$$\frac{1}{2} \le \sum_{u \in A_n(\varepsilon)} p_u \le \#A_n(\varepsilon) e^{-n(h-\varepsilon)}.$$

Rearranging the two inequalities gives the asserted bounds.

Solution 3

Equicontractivity gives

$$\chi(\Phi, p) = -\sum_i p_i \log r = \log(1/r).$$

Apply the finite Gibbs inequality to p and the uniform vector $q_i = 1/N$:

$$0 \le \sum_i p_i \log \frac{p_i}{1/N} = -h(p) + \log N.$$

Thus $h(p) \le \log N$, and division by $\log(1/r)$ gives the claimed bound. Equality in Gibbs' inequality holds exactly when $p_i = 1/N$ for all i.

Solution 4

Since $0 < r_{\min} \le r_{\max} < 1$ and $\log(1/2) < 0$, division reverses the first logarithmic inequality and gives

$$\gamma_{\min} \ge \gamma_{\max} > 0.$$

If $n = |\omega \wedge \tau|$, then

$$r_{\min}^n \le r_{\omega\wedge\tau} \le r_{\max}^n.$$

Because $d_\Sigma(\omega, \tau) = 2^{-n}$, this is exactly

$$d_\Sigma(\omega, \tau)^{\gamma_{\min}} \le \rho_\Phi(\omega, \tau) \le d_\Sigma(\omega, \tau)^{\gamma_{\max}}.$$

Each identity map between the two metric spaces is Hölder, so the induced topologies agree.

If $\tau \in [\omega|n]$, its common prefix with ω has length at least n, and therefore $\rho_\Phi(\omega, \tau) \le r_{(\omega|n)}$. If τ is outside that cylinder, the common prefix has length less than n; its ratio is strictly larger than $r_{(\omega|n)}$. This proves the closed-ball identity.

Solution 5

The frequency hypothesis gives

$$\frac{-\log p_{\omega|n}}{n} = \sum_i f_{n,i}(\omega)(-\log p_i) \longrightarrow -\sum_i q_i \log p_i$$

and

$$\frac{-\log r_{\omega|n}}{n} = \sum_i f_{n,i}(\omega)(-\log r_i) \longrightarrow -\sum_i q_i \log r_i > 0.$$

Thus the ratio of cylinder log-mass to cylinder log-radius tends to the displayed quotient. Solution 4 identifies weighted balls with cylinders, and the bounded one-step ratios show that intermediate radii have the same limit. Under SSC, the bi-Lipschitz ball sandwich (2.14) transfers this limit to Euclidean balls at $x = \pi_\Phi(\omega)$. Hence the local dimension exists and has the asserted value.

Solution 6

The contraction denominator is always $\log 3$. If the frequency of digit 0 tends to θ, Solution 5 gives

$$d_\mu(\pi(\omega)) = \frac{-\theta \log t - (1-\theta)\log(1-t)}{\log 3}.$$

At a typical point, $\theta = t$, and this becomes

$$\frac{-t \log t - (1-t)\log(1-t)}{\log 3}.$$

For $(02)^\infty$, $\theta = 1/2$, so the local dimension is

$$\frac{-\log(t(1-t))}{2\log 3}.$$

Solution 7

Uniform weights have entropy $\log N$, while every contraction information is $\log(1/r)$. Therefore

$$\operatorname{sdim}(\mu) = \frac{\log N}{\log(1/r)}, \qquad \dim_{\mathrm{nat}}(\mu) = \min\left\{d, \frac{\log N}{\log(1/r)}\right\}.$$

The labelled similarity dimension tends to infinity with N, and the natural dimension is at most d by definition. All maps have the same fixed point zero, so the invariant probability is δ_0 and $\dim_{\mathrm{H}} \mu = 0$ for every N.

Solution 8

Stopping words are prefix-free. Let q be the longest common prefix of distinct $u, v \in \Lambda_t$, and write their next letters as $i \neq j$. Since q is a prefix of u^-,

$$r_q \geq r_{u^-} > t.$$

The two cylinders lie in the images under S_q of two different first-level pieces. Therefore

$$\operatorname{dist}(K_u, K_v) \geq r_q \operatorname{dist}(S_i(K), S_j(K)) \geq r_q \delta > \delta t.$$

If a radius-t ball meets m stopping pieces, select one point from each intersection. The selected points are more than δt apart. The balls of radius $\delta t/2$ about them have disjoint interiors and lie inside the concentric ball of radius $(1+\delta/2)t$. Comparing Euclidean volumes gives

$$m\left(\frac{\delta t}{2}\right)^d \le ((1+\delta/2)t)^d\,,$$

which is the claimed bound.

Solution 9

Take a maximal r-separated set $\{x_1, \ldots, x_m\}$ in E. The balls $\overline{B}(x_j, r/3)$ are pairwise disjoint and, for small r, each has mass at least $a(r/3)^s$. Hence

$$1 \ge ma(r/3)^s, \qquad m \le a^{-1}3^s r^{-s}.$$

Maximality makes the radius-r balls about these points cover E, proving the upper bound for $N_r(E)$.

Conversely, suppose E is covered by m balls of radius r. Discard empty balls and choose $x_j \in E$ from each remaining intersection. The portion of E in the original ball lies in $\overline{B}(x_j, 2r)$, so

$$1 \le \sum_{j=1}^{m} \nu(\overline{B}(x_j, 2r)) \le mb(2r)^s.$$

Thus $m \ge b^{-1}2^{-s}r^{-s}$. The two constants prove the result.

Solution 10

Let $D = \operatorname{diam}(K)$ and let $\delta > 0$ be the SSC gap. For all sufficiently long words, Ahlfors regularity and $K_u \subset \overline{B}(x, Dr_u)$ for $x \in K_u$ give

$$p_u = \mu(K_u) \le bD^t r_u^t. \tag{10.1}$$

If $v \ne u$ has the same length and $q = u \wedge v$, then

$$\operatorname{dist}(K_u, K_v) \ge \delta r_q \ge \delta r_{u^-} \ge \frac{\delta}{r_{\max}} r_u.$$

Thus, with $c = \delta/(2r_{\max})$, a ball $\overline{B}(x, cr_u)$ centered at a point of K_u meets no other level-$|u|$ piece. The lower Ahlfors estimate therefore gives

$$ac^t r_u^t \le \mu(\overline{B}(x, cr_u)) \le p_u. \tag{10.2}$$

Apply (10.1)–(10.2) to $u = i^n$. Constants independent of n bound

$$\left(\frac{p_i}{r_i^t}\right)^n$$

above and below away from zero. Taking nth roots gives $p_i = r_i^t$. Summing over i yields $\sum_i r_i^t = 1$, so uniqueness of the similarity dimension gives $t = s(\Phi)$ and p is canonical. Conversely, T-I.8.06 says that the canonical measure is $s(\Phi)$-Ahlfors regular.

Solution 11

Choose one map S_C from each class. All maps in C equal S_C, so

$$\sum_{i\in\Lambda} p_i (S_i)_\# \nu = \sum_C q_C (S_C)_\# \nu.$$

The two measure operators coincide. Their invariant probabilities are therefore equal by uniqueness. The reduced IFS is SSC by hypothesis, so

$$\dim_{\mathrm{H}} \mu = \frac{h(q)}{-\sum_C q_C \log r_C}.$$

Put $\alpha_{(i|C)} = p_i / q_C$. Since $p_i = q_C \alpha_{(i|C)}$,

$$\begin{aligned} h(p) &= -\sum_C \sum_{i\in C} q_C \alpha_{i|C} \log(q_C \alpha_{i|C}) \\ &= h(q) + \sum_C q_C h((\alpha_{i|C})_{i\in C}). \end{aligned}$$

Equal maps have equal contraction ratios, and hence

$$\chi(\Phi, p) = -\sum_C q_C \log r_C.$$

The labelled entropy/Lyapunov upper bound is $h(p)/\chi$, while the actual dimension is $h(q)/\chi$. Their difference is the nonnegative conditional entropy term divided by χ; with strict weights it is positive whenever a class contains more than one label.

Solution 12

Terms with $p_i = 0$ vanish from the measure operator. Thus the operator is exactly that of the sub-IFS on Λ_+ with its strict restricted weight vector. T-I.7.04 gives uniqueness and support equal to the sub-attractor K_+. A sub-IFS of an SSC system is again SSC. Applying T-I.8.05 to this subsystem gives

$$\dim_{\mathrm{H}} \mu = \frac{-\sum_{i\in\Lambda_+} p_i \log p_i}{-\sum_{i\in\Lambda_+} p_i \log r_i}.$$

If there is only one positive weight, the numerator is zero and the measure is the point mass at that map's fixed point, so the same formula remains valid.

Solution 13

Write $Q(x) = qRx + b$. In the conjugate $Q \circ S_i \circ Q^{-1}$, the scale q from Q cancels the scale q^{-1} from its inverse, leaving contraction ratio r_i. The probability vector is unchanged. Hence $h(p)$, $\chi(\Phi, p)$, the measure similarity dimension, and the natural dimension are unchanged.

The attractor of the conjugated system is $Q(K)$. Moreover,

$$\begin{aligned}\sum_i p_i(\tilde{S}_i)_\#(Q_\#\mu) &= \sum_i p_i(Q\circ S_i)_\#\mu \\ &= Q_\#\left(\sum_i p_i(S_i)_\#\mu\right) = Q_\#\mu.\end{aligned}$$

Invariant-measure uniqueness gives $\mu_{(\widetilde{\Phi},p)} = Q_\#\mu_{(\Phi,p)}$. Finally,

$$Q(\overline{B}(x,r/q)) = \overline{B}(Qx,r).$$

The corresponding ball masses are equal, and $\log(r/q)/\log r \to 1$. Thus local dimensions, exact dimensionality, and the common exact dimension are preserved. Hausdorff and packing dimensions are also preserved because Q is bi-Lipschitz.

Solutions to Chapter I.9

Solution I.9.1

For $u = u_1 ... u_n$, the chain rule gives

$$DS_u(x) = DS_{u_1}(S_{u_2...u_n}x) \cdots DS_{u_n}(x).$$

Each factor is a scalar multiple of an orthogonal map. Norms therefore multiply exactly:

$$a_u(x) = \prod_{j=1}^{n} a_{u_j}(S_{u_{j+1}...u_n}x).$$

If $\omega|_n = u$, then $S_{(u_{j+1}...u_n)}(\pi\sigma^n\omega) = \pi\sigma^j\omega$. Hence

$$\begin{aligned} S_n\phi(\omega) &= \sum_{j=0}^{n-1} \log a_{\omega_{j+1}}(\pi\sigma^{j+1}\omega) \\ &= \log a_u(\pi\sigma^n\omega). \end{aligned}$$

This is (I.9.25).

Solution I.9.2

For every cylinder $[u]$,

$$\left| \sup_{[u]} S_n\varphi - \sup_{[u]} S_n\psi \right| \le n\|\varphi - \psi\|_\infty.$$

Thus

$$e^{-n\|\varphi-\psi\|_\infty} Z_n(\psi) \le Z_n(\varphi) \le e^{n\|\varphi-\psi\|_\infty} Z_n(\psi).$$

Take logarithms, divide by n, and pass to the limit to obtain (I.9.36).
For $0 \le \lambda \le 1$,

$$\sup_{[u]} S_n(\lambda\varphi + (1-\lambda)\psi) \le \lambda \sup_{[u]} S_n\varphi + (1-\lambda) \sup_{[u]} S_n\psi.$$

Hölder's inequality for the finite sum gives

$$Z_n(\lambda\varphi + (1-\lambda)\psi) \le Z_n(\varphi)^\lambda Z_n(\psi)^{1-\lambda}.$$

After logarithms, division by n, and passage to the limit,

$$P(\lambda\varphi + (1-\lambda)\psi) \le \lambda P(\varphi) + (1-\lambda)P(\psi).$$

Taking $\varphi = t_1\psi_0$ and $\psi = t_2\psi_0$ proves convexity of $t \mapsto \mathscr{P}(t\psi_0)$.

Solution I.9.3

For a similarity system,

$$S_n\phi(\omega) = \log r_{\omega_1\ldots\omega_n}.$$

It is constant on each length-n cylinder. Therefore

$$Z_n(t\phi) = \sum_{u\in\Lambda^n} r_u^t = \left(\sum_{i\in\Lambda} r_i^t\right)^n.$$

Hence

$$P_\Phi(t) = \log \sum_i r_i^t.$$

The pressure vanishes exactly when $\sum_i r_i^s = 1$, which is the similarity equation from I.7.

Solution I.9.4

Let $a_i = |x_i - y|$, $a_j = |x_j - y|$, and let θ be the angle between $x_i - y$ and $x_j - y$. If a center equals the common point y, there can be no other ball: that other ball would contain this center. The bound is then immediate. Otherwise all these vectors are nonzero. Since neither center lies in the other ball and y lies in both,

$$|x_i - x_j|^2 > a_i^2, \qquad |x_i - x_j|^2 > a_j^2.$$

The cosine rule gives

$$a_i^2 + a_j^2 - 2a_ia_j\cos\theta > a_i^2,$$

so $\cos\theta < a_j/(2a_i)$. Interchanging i, j gives $\cos\theta < a_i/(2a_j)$. The smaller of these two bounds is at most $1/2$. Thus the associated unit vectors have mutual distance

$$|v_i - v_j| = (2 - 2\cos\theta)^{1/2} > 1.$$

The radius-$1/2$ balls around them are disjoint and contained in the radius-$3/2$ ball about the origin. Volume comparison gives at most 3^d such vectors. Together with the singleton case, this is the required bound.

Solution I.9.5

Let $\nu = \theta_* m$, $\nu_A = \theta_*(1_A m)$, and $f = d\nu_A/d\nu$. The function $f \circ \theta$ is measurable with respect to $\theta^{-1}B(\mathbb{R}^d)$. For a Borel set D,

$$\int_{\theta^{-1}D} f \circ \theta \, dm = \int_D f \, d\nu = \nu_A(D) = \int_{\theta^{-1}D} \mathbf{1}_A \, dm.$$

Inverse images of Borel sets generate $\theta^{-1}B(\mathbb{R}^d)$. The monotone-class theorem extends the equality to every set in that sigma-algebra. Hence $f \circ \theta$ satisfies the defining property of $E(1_A | \theta^{-1}B(\mathbb{R}^d))$.

Solution I.9.6

The two terms inside the minimum agree when $t_0 = \log(B/a) \geq 0$. Therefore

$$\begin{aligned} \int_0^\infty \min\{a, Be^{-t}\} \, dt &= \int_0^{t_0} a \, dt + \int_{t_0}^\infty Be^{-t} \, dt \\ &= a \log(B/a) + Be^{-t_0} \\ &= a\,(1 + \log(B/a)) \,. \end{aligned}$$

Apply this with $a = m(C)$ and $B = B_d$ to each cell C in the layer-cake calculation. Summing yields

$$\sum_C m(C) + \sum_C m(C) \log B_d - \sum_C m(C) \log m(C) = 1 + \log B_d + H_m(\xi),$$

which is (I.9.54).

Solution I.9.7

Let D be Borel. On $[i]$, the coding identity gives

$$[i] \cap \sigma^{-1}\pi^{-1}D = [i] \cap \pi^{-1}(S_i(D \cap K)).$$

The restriction $S_i : U \to S_i(U)$ is a homeomorphism, and $S_i(U)$ is open. It therefore carries the Borel subset $D \cap K$ of U to a Borel subset of $S_i(U)$, hence to a Borel subset of $\mathbb{R}^d$. This proves $\widehat{\mathscr{P}} \vee \sigma^{-1}A_\pi \subset \widehat{\mathscr{P}} \vee A_\pi$.

Conversely,

$$[i] \cap \pi^{-1}D = [i] \cap \sigma^{-1}\pi^{-1}(S_i^{-1}(D \cap S_i(K))).$$

The inverse image is Borel, which proves the reverse inclusion.

Solution I.9.8

Write $A_j = \sigma^{-j}A_\pi$. By the chain rule,

$$H(P \vee \sigma^{-1}P | A_0) = H(P|A_0) + H(\sigma^{-1}P | P\text{-hat} \vee A_0),$$

$$H(P \vee \sigma^{-1}P|A_1) = H(P|A_1) + H(\sigma^{-1}P|P\text{-hat} \vee A_1).$$

The second terms agree by (I.9.72). Hence

$$H(P \vee \sigma^{-1}P|A_1) - H(P \vee \sigma^{-1}P|A_0) = H(P|A_1) - H(P|A_0) = h_\pi.$$

Apply the same identity after one pull-back by σ. Invariance gives

$$H(P \vee \sigma^{-1}P|A_2) - H(P \vee \sigma^{-1}P|A_1) = h_\pi.$$

Adding proves the $n = 2$ case of (I.9.73).

For the requested information-level identity, put $J_j = I_m(P \vee \sigma^{-1}P|A_j)$. The same cancellation and the pull-back rule give, almost everywhere,

$$J_1 - J_0 = \Delta_\pi, \qquad J_2 - J_1 = \Delta_\pi \circ \sigma, \qquad J_2 - J_0 = \Delta_\pi + \Delta_\pi \circ \sigma.$$

The two summands are functions, not pointwise constants h_π. Invariance makes their integrals both equal to $\int \Delta_\pi dm = h_\pi$.

Solution I.9.9

Let the nonzero probabilities be $q_1, ..., q_k$, where $k \le M$. Gibbs' inequality relative to the uniform vector gives

$$0 \le \sum_{j=1}^{k} q_j \log \frac{q_j}{1/k} = \sum_j q_j \log q_j + \log k.$$

Thus $-\sum q_j log q_j \le log k \le log M$. Equality with $log M$ requires $k = M$ and equality in Gibbs' inequality, so every nonzero probability is $1/M$.

At a point z, the vector $(p_u(z))$ in (I.9.93) has at most $N_n(z)$ nonzero terms. Its entropy is therefore at most $\log N_n(z) \le log M_n$.

Solution I.9.10

For $\varepsilon > 0$,

$$\sum_{n=1}^{\infty} m\{|G| > \epsilon n\} \le \int \frac{|G|}{\epsilon} dm < \infty.$$

Invariance gives

$$m\{|G \circ T^n| > \epsilon n\} = m\{|G| > \epsilon n\}.$$

For every M, the measure of the union of these events over $n \ge M$ is at most the corresponding tail of the convergent series, hence tends to zero. Thus $|G(T^n x)| \le \varepsilon n$ eventually. Intersecting over rational $\varepsilon > 0$ gives $G(T^n x)/n \to 0$.

In Lemma 5.1, after the indices with $n - j \ge N$ are controlled by the tail envelope H_N, fewer than N terms remain. Each is of the form $(F_k - F)(T^{n-k}x)/n$ for one fixed $k < N$; the result just proved makes every such term vanish.

Solution I.9.11

If $r_{(n+1)} < r \le r_n$, then

$$\nu(B(x, r_{n+1})) \le \nu(B(x, r)) \le \nu(B(x, r_n)).$$

First multiply ν by a positive constant so that $\nu(\mathbb{R}^d) \le 1$; this does not affect a local dimension. All three logarithms of masses are then nonpositive, and all logarithms of radii are negative. The radius-ratio hypothesis gives bounded $\log r_{(n+1)} - \log r_n$, while $\log r_n \to -\infty$; hence

$$\frac{\log r_{n+1}}{\log r_n} \to 1, \qquad \frac{\log r}{\log r_n} \to 1.$$

Combining the monotonicity of the mass logarithms with that of the radius logarithms gives the exact inequalities

$$\frac{\log \nu(B(x, r_n))}{\log r_{n+1}} \le \frac{\log \nu(B(x, r))}{\log r} \le \frac{\log \nu(B(x, r_{n+1}))}{\log r_n}$$

The left endpoint is the level-n quotient multiplied by $\log r_n / \log r_{(n+1)}$; the right endpoint is the level-$n+1$ quotient multiplied by $\log r_{(n+1)} / \log r_n$. Both multipliers tend to one. Taking liminf and limsup gives the two equalities in Lemma 6.3.

Solution I.9.12

Let an effective symbol be L for original labels $0, 1$ and R for label 2. The effective IFS is the strongly separated middle-third Cantor system, so the projected tail determines the effective tail code. The original first symbol is independent of the tail under the Bernoulli law. Hence

$$H(\mathscr{P} | \sigma^{-1} \mathscr{A}_\pi) = -p_0 \log p_0 - p_1 \log p_1 - p_2 \log p_2.$$

The full projected point determines whether its first effective symbol is L or R at every point: the two first-level Cantor cylinders are disjoint, so there is no endpoint ambiguity in this effective coding. Suppose first that $q > 0$. Conditional on L, the probabilities of original labels $0, 1$ are $p_0/q, p_1/q$; conditional on R, label 2 is certain. Therefore

$$H(\mathscr{P} | \mathscr{A}_\pi) = -p_0 \log \frac{p_0}{q} - p_1 \log \frac{p_1}{q}.$$

Subtracting gives

$$h_\pi(\sigma, \beta_p) = -q \log q - p_2 \log p_2.$$

This is exactly the entropy of the effective two-symbol process.

If $q = 0$, then $p_2 = 1$; the symbolic measure is supported on the constant sequence with label 2. Both conditional entropies and h_π are zero. This agrees with the final formula using $0 \log 0 = 0$; the intermediate conditional probabilities p_i/q are not used in this case.

Solution I.9.13

A scale class is nonempty only if there is an integer k satisfying

$$2^{-(k+1)} < r_{\max}^n, \qquad r_{\min}^n \le 2^{-k}.$$

Equivalently,

$$n\log_2(1/r_{\max}) - 1 < k \le n\log_2(1/r_{\min}).$$

The number of integers in this interval is at most

$$2 + n\left(\log_2(1/r_{\min}) - \log_2(1/r_{\max})\right) = 2 + n\log_2(r_{\max}/r_{\min}).$$

If every $r_i = r$, all length-n words have ratio r^n; only one scale class is needed. The packing estimate (I.9.152) then gives $M_n \le C_1$ uniformly in n.

Solution I.9.14

Subexponential prefix multiplicity and (I.9.83) give

$$h_\pi(\sigma, m_t) = h_{m_t}(\sigma).$$

The exact-dimensional formula (I.9.142) therefore yields

$$\dim(\pi_* m_t) = \frac{h_{m_t}(\sigma)}{\chi_\Phi(m_t)}.$$

Rearrange (I.9.31):

$$h_{m_t}(\sigma) = P_\Phi(t) + t\chi_\Phi(m_t).$$

Division by $\chi_\Phi(m_t) > 0$ gives (I.9.157). At $t = s_\Phi$, pressure is zero, so the right side is s_Φ, proving (I.9.158).

Solutions to Chapter I.10

Solution 1

Convexity and $K(0) = 0$ imply that the secant slope $K(t)/t$ is nondecreasing in t, with its continuous value $K'(0)$ at zero. Put $t = q - 1$. Then

$$H_q(a) = -\frac{K(q-1)}{q-1}.$$

The negative of a nondecreasing function is nonincreasing, so H_q decreases with q. Since

$$K'(0) = \frac{\sum_i a_i \log a_i}{\sum_i a_i} = \sum_i a_i \log a_i,$$

one has

$$\lim_{q \to 1} H_q(a) = -K'(0) = -\sum_i a_i \log a_i.$$

Solution 2

Fix $\varepsilon > 0$ and restrict to points at which

$$r^{\alpha+\varepsilon} \leq \mu(B(x,r)) \leq r^{\alpha-\varepsilon}$$

for all small r. If $q_+ > 0$, then

$$\mu(B(x,r))^{q_+} \geq r^{q_+(\alpha+\varepsilon)}.$$

If $q_- < 0$, the upper mass bound and the decreasing map $y \mapsto y^{q_-}$ give

$$\mu(B(x,r))^{q_-} \geq r^{q_-(\alpha-\varepsilon)}.$$

Thus a maximal disjoint family of N_r balls satisfies, respectively,

$$N_r r^{q_+(\alpha+\varepsilon)} \leq \mathscr{M}_\mu(q_+, r), \qquad N_r r^{q_-(\alpha-\varepsilon)} \leq \mathscr{M}_\mu(q_-, r).$$

The reversal occurs exactly when the inequality $\mu(B) \leq r^{\alpha-\varepsilon}$ is raised to the negative power q_-. The covering calculation in Theorem 1.6, followed by $\varepsilon \downarrow 0$, gives both bounds and hence their minimum.

Solution 3

The pressure equation is

$$r^{-\tau(q)} \sum_{i=1}^{N} p_i^q = 1.$$

Therefore

$$\tau(q) = \frac{\log \sum_i p_i^q}{\log r}, \qquad \tau'(q) = \frac{\sum_i p_i^q \log p_i}{(\sum_i p_i^q) \log r}.$$

For $q > 0$, $q \neq 1$,

$$D_q^{\mathrm{dyad}}(\mu) = \frac{\log \sum_i p_i^q}{(q-1) \log r}.$$

At $q = 1$, l'Hôpital's rule or Solution 1 gives

$$D_1^{\mathrm{dyad}}(\mu) = \frac{\sum_i p_i \log p_i}{\log r} = \frac{-\sum_i p_i \log p_i}{-\log r}.$$

Solution 4

For $i = 1, 2$,

$$\frac{\log p_i}{\log r_i} = \frac{\log(1/8)}{\log(1/4)} = \frac{3}{2}.$$

For $i = 3$,

$$\frac{\log p_3}{\log r_3} = \frac{\log(3/4)}{\log(1/5)} = \frac{\log(4/3)}{\log 5} < 1 < \frac{3}{2}.$$

Thus $\alpha_{\max} = 3/2$, and its extremal face is the two-symbol face $\{1, 2\}$. Its dimension s_+ solves

$$2(1/4)^{s_+} = 1,$$

so $s_+ = 1/2$. Theorem 3.3 therefore gives $\dim_{\mathrm{H}} E_\mu(3/2) = 1/2$, not zero.

Solution 5

Differentiate

$$0 = \sum_i \theta_i(q)\big(\log p_i - \tau'(q) \log r_i\big).$$

Since

$$\theta_i'(q) = \theta_i(q)\big(\log p_i - \tau'(q) \log r_i\big),$$

a second differentiation gives

$$0 = \sum_i \theta_i(q)\big(\log p_i - \tau'(q)\log r_i\big)^2 - \tau''(q)\sum_i \theta_i(q)\log r_i.$$

This is the displayed formula. Its denominator is negative and its numerator nonnegative.

If τ is affine on an open interval, then $\tau''(q) = 0$ at any point of that interval. Every $\theta_i(q) > 0$, so every squared term vanishes:

$$\log p_i = \tau'(q)\log r_i.$$

Writing $s = \tau'(q)$ gives $p_i = r_i^s$; summing gives $\sum_i r_i^s = 1$. Conversely, if $p_i = r_i^s$, substitution in the pressure equation gives $\tau(q) = s(q-1)$, which is affine.

Solution 6

Set

$$A_s = \int \psi_0\, dm_s, \qquad B_s = \int \phi\, dm_s < 0, \qquad \Delta = T(q+h) - T(q).$$

Evaluate the pressure at $q+h$ using m_q:

$$0 \ge h_{m_q} + (q+h)A_q - T(q+h)B_q = hA_q - \Delta B_q.$$

Thus $\Delta B_q \ge hA_q$. Division by $hB_q < 0$ reverses the inequality and gives

$$\frac{\Delta}{h} \le \frac{A_q}{B_q}.$$

Evaluate the pressure at q using m_{q+h}:

$$0 \ge h_{m_{q+h}} + qA_{q+h} - T(q)B_{q+h} = -hA_{q+h} + \Delta B_{q+h}.$$

Hence $\Delta B_{q+h} \le hA_{q+h}$, and division by $hB_{q+h} < 0$ gives

$$\frac{\Delta}{h} \ge \frac{A_{q+h}}{B_{q+h}}.$$

The negative sign of the geometric average B_s causes both reversals.

Solution 7

Pressure zero at the tilted equilibrium state says

$$h_{m_q}(\sigma) + q\int \psi_0\, dm_q - T(q)\int \phi\, dm_q = 0.$$

Put $A = \int \psi_0\, dm_q$, $B = \int \phi\, dm_q < 0$. Since $\alpha = A/B$,

$$h_{m_q} = -qA + T(q)B.$$

The separated conformal entropy/Lyapunov formula gives

$$\dim_{\mathrm{H}}(\pi_* m_q) = \frac{h_{m_q}}{-B} = \frac{-qA + T(q)B}{-B} = q\frac{A}{B} - T(q) = q\alpha - T(q).$$

Solution 8

The ratio is

$$\varepsilon^q + (1-\varepsilon)^q.$$

When $q < 0$, $\varepsilon^q \to \infty$ as $\varepsilon \downarrow 0$, while $(1-\varepsilon)^q \to 1$. Thus the ratio tends to infinity. Bounded incidence controls how many pieces a cylinder meets, but it gives no positive lower bound on the mass of each piece. One arbitrarily small grid-cut piece can therefore make the negative-order grid moment arbitrarily larger than the cylinder moment.

Solution 9

For a tail cylinder $[v]$,

$$m([u] \cap \sigma^{-|u|}[v]) = m([uv]) \le C_+ m([u]) m([v]).$$

If $A = \bigsqcup_{j=1}^{k} [v_j]$ is a disjoint finite union, add these inequalities to get

$$m([u] \cap \sigma^{-|u|} A) \le C_+ m([u]) m(A).$$

The cylinder algebra generates the Borel sigma-algebra. Both sides are continuous under increasing unions and decreasing intersections, so the monotone-class theorem extends the inequality to every Borel A.

For a stopping antichain $\mathscr{G}$, decompose $\pi^{-1}E$ over the disjoint cylinders $[u]$. On $[u]$,

$$[u] \cap \pi^{-1} E = [u] \cap \sigma^{-|u|} \pi^{-1}(S_u^{-1} E).$$

Apply the tail inequality to each term and sum:

$$\mu(E) \le C_+ \sum_{u \in \mathscr{G}} m([u]) \mu(S_u^{-1} E).$$

Solution 10

Fix ℓ. Repeatedly use the inequality with the second block of length ℓ. For $n = k\ell + r$, $0 \le r < \ell$,

$$A_{k\ell+r} \le \begin{cases} A_r (c_\ell A_\ell)^k, & 1 \le r < \ell, \\ A_\ell (c_\ell A_\ell)^{k-1}, & r = 0, \end{cases}.$$

The second line removes any need to invent an undefined A_0. After absorbing the finitely many factors into $C_\ell > 0$,

$$\limsup_{n\to\infty} \frac{\log A_n}{n} \le \frac{\log c_\ell + \log A_\ell}{\ell}.$$

This holds for every ℓ. Let $\ell \to \infty$ along a subsequence realizing the liminf of $\ell^{-1} \log A_\ell$; the $\ell^{-1} \log c_\ell$ term vanishes. Hence

$$\limsup_n \frac{\log A_n}{n} \le \liminf_n \frac{\log A_n}{n},$$

so the limit exists. If all inequalities are reversed, use the same two remainder cases, absorb their finitely many positive factors into a positive lower constant, and reverse the limsup/liminf comparison.

Solution 11

The recursion makes t_n positive and decreasing. Suppose $t_n \to t > 0$. Then

$$\log b_n = \sum_{j=1}^{n} t_j \ge tn$$

for large n, while $\log(A_n e^{-an}) = o(n)$. Hence $A_n b_n e^{-an} \ge 1$ eventually, so the recursion halves t_n at every subsequent step. This contradicts convergence to a positive t. Therefore $t_n \to 0$.

There must be infinitely many halving steps. At every index that triggers a halving, the corresponding series term is at least one. Thus

$$\sum_n A_n b_n e^{-an} = \infty.$$

Monotonicity of t_n gives

$$\sum_{j=m+1}^{m+n} t_j \le \sum_{j=1}^{n} t_j,$$

so $b_{m+n} \le b_m b_n$. Finally $t_j \le t_1 = \varepsilon$ gives $b_n \le e^{\varepsilon n}$.

Solution 12

The sequence $x_k = \omega(\vartheta^k)$ tends to zero because $\vartheta^k \to 0$ and $\omega(\rho) \to 0$. Cesàro convergence gives

$$\frac{1}{n} \sum_{k=0}^{n-1} \omega(\vartheta^k) = \frac{1}{n} \sum_{k=0}^{n-1} x_k \longrightarrow 0.$$

Therefore the sum is $o(n)$. If the one-step derivative comparison has logarithmic error at most $C\omega(\vartheta^k)$, multiplication along a word gives a total multiplicative error bounded by

$$\exp\left(C \sum_{k=0}^{n-1} \omega(\vartheta^k)\right) = e^{o(n)}.$$

Solution 13

Let the side length be $\delta = 2^{-n}$. Color the cube with index $k = (k_1, \dots, k_d)$ by

$$(k_1 \bmod L, \dots, k_d \bmod L),$$

where

$$L = \lceil 8\sqrt{d} \rceil + 5.$$

There are L^d colors. If two distinct cubes have the same color, their indices differ by at least L in some coordinate. Points chosen in those cubes are therefore separated by at least $(L-1)\delta$, which exceeds $4\sqrt{d}\,\delta$. Thus balls of radius $2\sqrt{d}\,\delta$ about the chosen points are disjoint.

Each such ball contains its cube, so for a fixed color,

$$\sum_{Q \text{ of that color}} \mu(Q)^q \leq \sum_Q \mu(B(x_Q, 2\sqrt{d}\,\delta))^q \leq \mathscr{M}_\mu(q, 2\sqrt{d}\,\delta).$$

Summing over L^d colors proves the required comparison with a scale-independent constant.

Solution 14

The chain is:

1. Two-sided quasi-Bernoulli supplies upper quasi-Bernoulli for $q \geq 1$ and lower quasi-Bernoulli for $0 < q < 1$. Lemmas 6.5 and 7.2 give the two perturbed product inequalities. Lemmas 7.3 and 7.5 then give the finite dyadic limit for every $q > 0$, which is Theorem 8.3.
2. Lemma 8.1 uses $q > 0$ to compare dyadic moments with centered disjoint ball moments. It identifies the two spectra.
3. Since $t \geq 1$, the upper quasi-Bernoulli half of Lemma 7.5 gives the upper t-Gibbs ball estimate (I.10.193).
4. Lemma 8.2 supplies a finite concave all-real lower packing spectrum. At differentiability, Lemma 5.4 uses the upper Gibbs-ball estimate and the neighboring exponents $t \pm u$ to show that the auxiliary measure is carried by $E_\mu(\tau'(t))$.
5. The local dimension lower bound for that auxiliary measure gives $\dim_{\mathrm{H}} E_\mu(\tau'(t)) \geq t\tau'(t) - \tau(t)$. Theorem 1.6 gives the reverse inequality, and concavity identifies this value with the full Legendre infimum.

No step uses a separation condition or uniform bounded distortion; the geometric input is the $e^{o(n)}$ pull-back comparison of Lemma 6.5.

Solutions to Chapter I.11

Solution I.11.1

Put $g(z) = f(z) - z$. Since $\mathbb{E}\xi > 1$, $g(1) = 0$ and $g'(1) > 0$, so $g(z) < 0$ immediately to the left of 1. Also $g(0) = \mathbb{P}(\xi = 0) \geq 0$. If $g(0) = 0$, then zero is the second root. Otherwise continuity gives a root in $(0, 1)$.

The supercritical hypothesis forces $\mathbb{P}(\xi \geq 2) > 0$, so f, and hence g, is strictly convex on $(0, 1)$. A strictly convex function with $g(1) = 0$, negative values immediately to the left of 1, and a nonnegative value at zero has exactly one root in $[0, 1)$. Thus there are exactly two roots counting 1.

Let $q_0 = 0$ and $q_{n+1} = f(q_n)$. Since f is increasing and $q_1 \geq q_0$, induction gives $q_{n+1} \geq q_n$. The sequence is bounded by one, so $q_n \to q$, and continuity gives $f(q) = q$. If a is any fixed point in $[0, 1]$, then $q_0 \leq a$ and induction gives $q_n \leq a$. Hence $q \leq a$, so q is the smaller fixed point.

Solution I.11.2

Assume $j < k$. Since Δ_j is $\mathscr{F}_j$-measurable and $\mathscr{F}_j \subset \mathscr{F}_{k-1}$,

$$\mathbb{E}(\Delta_j \Delta_k) = \mathbb{E}\big(\Delta_j \mathbb{E}(\Delta_k \mid \mathscr{F}_{k-1})\big) = 0.$$

Thus the differences are pairwise orthogonal. Since $M_n - M_r = \sum_{k=r+1}^{n} \Delta_k$, Pythagoras in L^2 gives

$$\mathbb{E}|M_n - M_r|^2 = \sum_{k=r+1}^{n} \mathbb{E}|\Delta_k|^2.$$

Also

$$\mathbb{E}|M_n|^2 = \mathbb{E}|M_0|^2 + \sum_{k=1}^{n} \mathbb{E}|\Delta_k|^2.$$

If the left side is uniformly bounded, the series of nonnegative difference norms converges. Its tails tend to zero uniformly in $n > r$, which is exactly the L^2-Cauchy property.

Solution I.11.3

The generating function is

$$f(z) = 1 - p + pz^2.$$

Its fixed-point equation factors as

$$f(z) - z = (z-1)(pz - (1-p)).$$

Thus

$$q = \frac{1-p}{p} < 1.$$

Moreover,

$$m = \mathbb{E}\xi = 2p, \qquad \sigma^2 = \mathbb{E}\xi^2 - m^2 = 4p - 4p^2 = 4p(1-p).$$

Formula (I.11.20) gives

$$\lim_{n\to\infty} \mathbb{E}W_n^2 = 1 + \frac{4p(1-p)}{2p(2p-1)} = \frac{1}{2p-1}.$$

All denominators are positive because $p > 1/2$.

Solution I.11.4

Writing $q_n = \mathbb{E}X_n^2$, recurrence (I.11.45) and $q_0 = 1$ give

$$q_n = c^n + B\sum_{j=0}^{n-1} c^j = c^n + B\frac{1-c^n}{1-c}.$$

Because $0 \le c < 1$,

$$\mathbb{E}X_\varnothing^2 = \lim_{n\to\infty} q_n = \frac{B}{1-c}.$$

This is at least one, as it must be since $\mathbb{E}X_\varnothing = 1$; that inequality also follows by taking the limit in $q_n \ge (\mathbb{E}X_n)^2 = 1$.

Solution I.11.5

For a first-level label i, the definition of $\mathbb{Q}_T$ and the cylinder formula give

$$\mathbb{Q}_T(I_1 = i) = \mathbb{E}\big(\mathbf{1}_{\{i\in A_\varnothing\}} r_i^s X_i\big) = \mathbb{P}(i \in A_\varnothing) r_i^s = \theta_i.$$

For a word $u = u_1 \cdots u_n$ with positive cylinder probability and a label i,

$$\frac{\mathbb{Q}_T(I|_{n+1} = ui)}{\mathbb{Q}_T(I|_n = u)} = \frac{\prod_{j=1}^n \theta_{u_j}\theta_i}{\prod_{j=1}^n \theta_{u_j}} = \theta_i.$$

Thus the next label has the same law independently of the past.

Finally, for a Borel set C,

$$
\begin{aligned}
\mathbb{Q}_T(X_{I|_n} \in C) &= \sum_{|u|=n} \mathbb{E}\big(\mathbf{1}_{\{u\in T\}} r_u^s X_u \mathbf{1}_{\{X_u\in C\}}\big) \\
&= \mathbb{E}\big(X_\varnothing \mathbf{1}_{\{X_\varnothing\in C\}}\big) \sum_{|u|=n} \prod_{j=1}^{n} \theta_{u_j} \\
&= \mathbb{E}\big(X_\varnothing \mathbf{1}_{\{X_\varnothing\in C\}}\big).
\end{aligned}
$$

This is the claimed size-biased law.

Solution I.11.6

Let $D = \operatorname{diam} K_{\mathcal{S}}$. Independence along words gives

$$\mathbb{E} \sum_{u\in T_n} (Dr_u)^t = D^t \Phi(t)^n.$$

Since $t > s$, $\Phi(t) < 1$, and therefore

$$\mathbb{E} \sum_{n\geq 1} \sum_{u\in T_n} (Dr_u)^t = D^t \sum_{n\geq 1} \Phi(t)^n < \infty.$$

Tonelli implies that the nonnegative double series is finite almost surely. In particular, its generation-n term tends to zero. The sets $S_u(K_{\mathcal{S}})$, $u \in T_n$, cover K_T, and their maximum diameter is at most $D(\max_i r_i)^n \to 0$. These covers are admissible in the definition of $\mathscr{H}^t$, and their t-cost tends to zero. Hence $\mathscr{H}^t(K_T) = 0$.

Solution I.11.7

Given $\mathscr{F}_n$, the descendant environments of the Z_n retained cubes are independent copies of the root environment. Each survives with probability $\vartheta = 1 - q$. Therefore

$$V_n \mid \mathscr{F}_n \sim \operatorname{Bin}(Z_n, \vartheta).$$

Conditional Chebyshev gives

$$\mathbb{P}(|V_n - \vartheta Z_n| > \varepsilon Z_n \mid \mathscr{F}_n) \leq \frac{\vartheta(1-\vartheta)}{\varepsilon^2 Z_n}.$$

On $S_k = \{W \geq 1/k\}$, eventually $Z_n \geq m^n/(2k)$. Intersect the deviation event with this lower bound, integrate the conditional estimate, and sum:

$$\sum_n \mathbb{P}\left(|V_n - \vartheta Z_n| > \varepsilon Z_n,\ Z_n \geq \frac{m^n}{2k}\right) \leq C_{\varepsilon,k} \sum_n m^{-n} < \infty.$$

Borel–Cantelli proves convergence of the ratio on S_k; the union of the S_k is survival modulo null sets.

The count Z_n includes retained cubes whose descendants later become empty. Such cubes need not meet K, so selecting one point of K from each retained cube is impossible. The live count V_n is what supplies the separated covering lower bound.

Solution I.11.8

Let $\tilde{\lambda}_T = X_\varnothing \nu_T$ on the full symbolic space, set equal to zero on extinction. For a word u of length n,

$$\mathbb{E}\tilde{\lambda}_T([u]) = \mathbb{E}\big(\mathbf{1}_{\{u\in T\}} m^{-n} X_u\big) = p^n m^{-n} = b^{-n}.$$

Finite unions of symbolic cylinders form a generating algebra. Thus $\mathbb{E}\tilde{\lambda}_T$ is the uniform b-ary probability on all symbolic Borel sets. Push forward by π_M. Its image is Lebesgue measure, so $\mathbb{E}\lambda_T(A) = \mathscr{L}^d(A)$ for every Euclidean Borel set A. This order avoids assuming zero boundary mass before proving it.

If H is an M-adic boundary, then $\mathscr{L}^d(H) = 0$, hence $\mathbb{E}\lambda_T(H) = 0$. Nonnegativity implies $\lambda_T(H) = 0$ almost surely. A countable union over all generations has the same property.

Solution I.11.9

From

$$Y_{n+1} = b^{-1} \sum_{i=1}^{b} W_i Y_n^{(i)},$$

the diagonal terms contribute $b^{-1}\mathbb{E}W^2\, \mathbb{E}Y_n^2$, while the $b(b-1)$ ordered off-diagonal terms contribute $(b-1)/b$. Thus, with $a = \mathbb{E}W^2/b$,

$$q_{n+1} = aq_n + \frac{b-1}{b}, \qquad q_0 = 1.$$

Consequently,

$$q_n = a^n + \frac{b-1}{b}\frac{1-a^n}{1-a}$$

when $a \neq 1$, and $q_n = 1 + n(b-1)/b$ when $a = 1$. If $a < 1$, this sequence is bounded. If $a = 1$, it grows linearly; if $a > 1$, it grows exponentially. Since $b \geq 2$, this proves

$$\sup_n \mathbb{E}Y_n^2 < \infty \quad \Longleftrightarrow \quad \mathbb{E}W^2 < b.$$

Solution I.11.10

Partitioning the Peyrière integral over first-generation cubes gives

$$\begin{aligned}\mathbb{Q}(W_{I_1} \in C) &= \sum_{i=1}^{b} b^{-1}\mathbb{E}(W_i Y_i \mathbf{1}_{\{W_i\in C\}}) \\ &= \mathbb{E}(W\mathbf{1}_{\{W\in C\}}),\end{aligned}$$

because W_i and Y_i are independent and $\mathbb{E}Y_i = 1$. At generation n, the same calculation gives

$$\mathbb{Q}(Y_{I|_n} \in C) = \mathbb{E}(Y\mathbf{1}_{\{Y\in C\}}).$$

Taking $C = [0, e^{-\varepsilon n})$ yields

$$\mathbb{Q}(Y_{I|_n} < e^{-\varepsilon n}) \le e^{-\varepsilon n}.$$

Taking $C = (e^{\varepsilon n}, \infty)$ and using $Y\mathbf{1}_{\{Y>t\}} \le t^{-1}Y^2$ yields

$$\mathbb{Q}(Y_{I|_n} > e^{\varepsilon n}) \le e^{-\varepsilon n}\mathbb{E}Y^2.$$

Solution I.11.11

Take $1 \le \ell \le n$. If $A \ge M^\ell$, the desired bound is trivial with a constant at least A. Otherwise fix an offset a with $1 \le a \le A < M^\ell$. If adding a propagates through at least ℓ digits, then the residue of j modulo M^ℓ lies among the final a residues

$$M^\ell - a, \dots, M^\ell - 1.$$

There are at most $a\lceil M^n/M^\ell\rceil \le 2AM^{n-\ell}$ such j, when $\ell \le n$. For a negative offset, the residue must be among the first $|a|$ residues, giving the same bound. A union over the at most $2A$ nonzero offsets gives probability at most $C_{A,M}M^{-\ell}$. Indices leaving $\{0, \dots, M^n - 1\}$ can simply be discarded and only reduce the count.

In d dimensions, union over coordinates multiplies the constant by d. With $\ell = \lceil \varepsilon n\rceil$, the probabilities are summable. Borel–Cantelli and a countable intersection over rational $\varepsilon > 0$ give $R_n/n \to 0$.

Solution I.11.12

The variable Y_v is measurable with respect to weights strictly below the vertex v. If $u \ne v$ have the same length, that descendant weight family is disjoint from:

1. every weight on the path from the root to u; and
2. every weight strictly below u, which determines Y_u.

Tree-indexed independence therefore makes Y_v independent of

$$\mu(Q_u) = b^{-n}\left(\prod_{j=1}^{n} W_{u|_j}\right)Y_u,$$

even if the two paths share an arbitrarily long proper prefix. Consequently,

$$\mathbb{E}\big(\mu(Q_u)\mathbf{1}_{\{Y_v>t\}}\big) = \mathbb{E}\mu(Q_u)\,\mathbb{P}(Y > t) \le b^{-n}t^{-1}.$$

For $v = u$, summing gives at most $t^{-1}\mathbb{E}Y^2$. There are at most Jb^n neighboring ordered pairs, so

$$\mathbb{Q}\left(\max_{v\in\mathcal{N}_n(I|_n)} Y_v > t\right) \le \frac{\mathbb{E}Y^2 + J}{t}.$$

Solution I.11.13

At level n, a word contributes to the deletion cascade exactly when it is retained. Every retained path has weight product p^{-n}. Therefore

$$Y_n = b^{-n} p^{-n} Z_n = (pb)^{-n} Z_n.$$

The equal-ratio weighted branching sum has $r_u^s = M^{-ns} = (pb)^{-n}$, so

$$X_n = (pb)^{-n} Z_n = Y_n.$$

Apply the same computation in every descendant environment and pass to the L^2 limits to obtain $X_u = Y_u$ for every word on one probability-one event. Hence, for retained u,

$$\frac{\mu(Q_u)}{Y} = \frac{(pb)^{-|u|} Y_u}{Y} = \frac{r_u^s X_u}{X_\varnothing} = \nu_T([u]).$$

Grid boundaries have zero mass, so equality on cylinders proves equality of the two Borel probabilities.

Solution I.11.14

Statement (I.11.165) has the form

$$\mathbb{P}(\{\omega : \forall \ell,\ E(\omega, \ell)\} \mid K \neq \varnothing) = 1.$$

Statement (I.11.167) has the form

$$\forall a \in (\mathbb{R} \setminus \{0\})^k, \quad \mathbb{P}(\{\omega : E(\omega, a)\} \mid \text{all factors survive}) = 1.$$

The universal quantifier lies inside the probability in the first statement and outside it in the second.

For a concrete example, take $\Omega = [0,1]$ with Lebesgue measure and, for $a \in [0,1]$, let $E_a = \Omega \setminus \{a\}$. Every E_a has probability one, so $\forall a\, \mathbb{P}(E_a) = 1$. But

$$\bigcap_{a \in [0,1]} E_a = \varnothing,$$

and therefore $\mathbb{P}(\forall a\, E_a) = 0$. This is why a source's for-each-fixed-parameter statement cannot be silently upgraded to a common event over an uncountable parameter set.

Solutions to Chapter I.12

S-I.12.01 — A concrete weak metric

Every summand is nonnegative and symmetric, so the same is true of d. The triangle inequality follows term by term. If $d(\mu, \nu) = 0$, then

$$\int f_j \, d\mu = \int f_j \, d\nu \qquad (j \geq 1). \tag{S.I.12.1}$$

For $f \in C(K)$, put $a = \max\{1, \|f\|_\infty\}$. Choose f_j with $\|f/a - f_j\|_\infty < \varepsilon/a$. Then

$$\left| \int f \, d\mu - \int f \, d\nu \right| \leq 2\varepsilon + a \left| \int f_j \, d\mu - \int f_j \, d\nu \right| = 2\varepsilon. \tag{S.I.12.2}$$

Letting $\varepsilon \downarrow 0$ and using Riesz uniqueness gives $\mu = \nu$. Thus d is a metric.

If $\mu_n \to \mu$ weakly, each fixed summand tends to zero. Split the series at J: its finite head tends to zero, while its tail is at most $\sum_{j>J} 2^{-j}$. Hence $d(\mu_n, \mu) \to 0$.

Conversely, if $d(\mu_n, \mu) \to 0$, then the integral against every f_j converges. Approximate an arbitrary f/a by one fixed f_j as in (S.I.12.2). The two approximation errors are uniform in n, so $\int f d\mu_n \to \int f d\mu$. This is weak convergence.

S-I.12.02 — A periodic empirical distribution

Let the orbit have period dividing q and define

$$P = \frac{1}{q} \sum_{j=0}^{q-1} \delta_{T^j z}. \tag{S.I.12.3}$$

Write $N = kq + r$, $0 \leq r < q$. For $f \in C(K)$, complete blocks contribute $k \sum_{j<q} f(T^j z)$, while the remainder has absolute value at most $q\|f\|_\infty$. Therefore

$$\left| \int f \, dA_N^T(z) - \int f \, dP \right| \leq \frac{2q\|f\|_\infty}{N} \longrightarrow 0. \tag{S.I.12.4}$$

Also

$$T_* P = \frac{1}{q} \sum_{j=0}^{q-1} \delta_{T^{j+1} z} = P, \tag{S.I.12.5}$$

because $T^q z = z$ cyclically permutes the summands.

S-I.12.03 — Flow averages over shifted intervals

For $f \in C(K)$, the symmetric difference of the intervals $[0, R]$ and $[a_R, a_R + R]$ has length at most $2|a_R|$. Hence

$$\left|\int f\, dA_{R,a_R}(z) - \int f\, dA_R^{\mathsf{T}}(z)\right| \leq \frac{2|a_R|}{R}\|f\|_\infty \longrightarrow 0. \tag{S.I.12.6}$$

This is weak convergence of the signed difference to zero.

S-I.12.04 — Tangents of Lebesgue measure

Translation and scaling give

$$(T_{x,r})_* \mathscr{L}^d = r^d \mathscr{L}^d. \tag{S.I.12.7}$$

Thus every tangent sequence has the form

$$c_n r_n^d \mathscr{L}^d. \tag{S.I.12.8}$$

If it converges vaguely to a nonzero Radon measure, testing against one nonnegative $f \in C_c(\mathbb{R}^d)$ with positive Lebesgue integral shows that $c_n r_n^d \to c$ for some finite $c \geq 0$. Testing every other compactly supported continuous function then identifies the limit as $c\mathscr{L}^d$. Since the limit is nonzero, $c > 0$. Conversely, choosing $c_n = c r_n^{-d}$ realizes every $c\mathscr{L}^d$. Therefore

$$\operatorname{Tan}(\mathscr{L}^d, x) = \{c\mathscr{L}^d \,:\, c > 0\}. \tag{S.I.12.9}$$

S-I.12.05 — Existence under Ahlfors regularity

Choose any $r_n \downarrow 0$ and put $\lambda_n = \lambda_{r_n}$. If μ has Ahlfors constants c, C for all sufficiently small radii, then for fixed $R \geq 1$ and all large n,

$$\lambda_n(B(0,R)) = \frac{\mu(B(x, Rr_n))}{\mu(B(x, r_n))} \leq \frac{C}{c} R^s. \tag{S.I.12.10}$$

Use the countable test family with cutoffs from the proof of L-I.12.02. For each test supported in a fixed compact ball, (S.I.12.10), with a slightly larger radius, bounds its integrals uniformly for all sufficiently large n. Each of the finitely many earlier integrals is finite. Diagonal extraction thus makes the integrals of all countably many tests converge. Uniform approximation on a common compact support and the same local mass bounds make the integrals of every $f \in C_c(\mathbb{R}^d)$ converge along that subsequence. Their limits define a positive linear functional L with $|L(f)| \leq C_m \|f\|_\infty$ when $\operatorname{supp}(f) \subset B(0, m)$. The allowed locally compact Riesz theorem gives a Radon measure λ representing L. Hence $\lambda_{n_k} \to \lambda$ vaguely. Moreover,

$$\lambda_n(B(0, 1/2)) = \frac{\mu(B(x, r_n/2))}{\mu(B(x, r_n))} \geq \frac{c}{C} 2^{-s}. \tag{S.I.12.11}$$

Choose a nonnegative compactly supported function that is one on $B(0, 1/2)$. Equation (S.I.12.11) shows that the limit is nonzero. Hence $\lambda \in \operatorname{Tan}(\mu, x)$.

S-I.12.06 — Why the boundary-null clause is necessary

Vague convergence of Dirac masses gives

$$\lambda_n \longrightarrow \lambda = \delta_0 + \delta_1. \tag{S.I.12.12}$$

Because $1 + 1/n$ lies outside B,

$$\mathscr{N}_B(\lambda_n) = \delta_0. \tag{S.I.12.13}$$

But B is closed and contains the boundary point 1, so

$$\mathscr{N}_B(\lambda) = \frac{1}{2}(\delta_0 + \delta_1). \tag{S.I.12.14}$$

Thus normalized restriction is discontinuous here. Exactly the new mass on $\partial B = \{-1, 1\}$ causes the failure.

S-I.12.07 — The scenery of Lebesgue measure

For a Borel set $A \subset B$,

$$((S_t)_*(T_x)_*\mathscr{L}^d)(A) = e^{-dt}\mathscr{L}^d(A), \tag{S.I.12.15}$$

while

$$\mathscr{L}^d(x + e^{-t}B) = e^{-dt}\mathscr{L}^d(B). \tag{S.I.12.16}$$

Therefore, with $m_B = \mathscr{L}^d|_B/\mathscr{L}^d(B)$,

$$(\mathscr{L}^d)^{\square}_{x,t} = m_B, \qquad \langle \mathscr{L}^d \rangle_{x,R} = \delta_{m_B}. \tag{S.I.12.17}$$

The unique tangent distribution is δ_{m_B}.

S-I.12.08 — Support pairs are not closed

Since $\nu_n(\{\omega\}) = 1/n > 0$, every open neighborhood of ω has positive ν_n-mass. Hence $\omega \in \mathrm{supp}(\nu_n)$. Weakly,

$$\nu_n \longrightarrow \delta_\eta. \tag{S.I.12.18}$$

Because $\omega \neq \eta$, some cylinder contains ω and excludes η; that cylinder has zero δ_η-mass. Thus $\omega \notin \mathrm{supp}(\delta_\eta)$. The example invalidates any claim that the support-pair subset of $\mathscr{P}(\Sigma)x\Sigma$ is closed, and therefore invalidates compactness arguments based on that claim.

S-I.12.09 — Composition of components

For Borel $A \subset \Sigma$, use the definitions twice:

$$
\begin{aligned}
(\nu_u)_v(A) &= \frac{\nu_u(vA)}{\nu_u([v])} = \frac{\nu(uvA)/\nu([u])}{\nu([uv])/\nu([u])} \\
&= \frac{\nu(uvA)}{\nu([uv])} = \nu_{uv}(A).
\end{aligned}
\tag{S.I.12.19}
$$

For an admissible pair all prefix masses are positive. Apply (S.I.12.19) inductively while deleting one symbol at each application of M; this gives

$$
M^n(\nu, \omega) = (\nu_{\omega|_n}, \sigma^n \omega). \tag{S.I.12.20}
$$

S-I.12.10 — The exact small-cell identity

Adaptedness permits sampling ν first with law $\bar{Q}$ and then ω with law ν. For fixed ν, the selected first symbol is i with probability $\nu([i])$. Therefore

$$
Q\{q < a\} = \int \sum_{i\,:\,\nu([i])<a} \nu([i])\, d\bar{Q}(\nu). \tag{S.I.12.21}
$$

There are at most $\#\Lambda$ summands, and every displayed summand is less than a. Hence

$$
Q\{q < a\} \le (\#\Lambda)a. \tag{S.I.12.22}
$$

S-I.12.11 — Closedness of adaptedness

Suppose $\nu_k \to \nu$. Add and subtract the integral of $F(\nu, \eta)$ against ν_k:

$$
\begin{aligned}
|J_F(\nu_k) - J_F(\nu)| \le \sup_\eta |F(\nu_k, \eta) - F(\nu, \eta)| \\
+ \left| \int F(\nu, \eta)\, d\nu_k(\eta) - \int F(\nu, \eta)\, d\nu(\eta) \right|.
\end{aligned}
\tag{S.I.12.23}
$$

The first term tends to zero by uniform continuity on the compact product; the second tends to zero by weak convergence. Thus J_F is continuous.

If Q_k are adapted and $Q_k \to Q$, their first marginals converge. Passing to the limit in

$$
\int F\, dQ_k = \int J_F\, d\bar{Q}_k \tag{S.I.12.24}
$$

gives the same identity for Q, so Q is adapted. The adapted set is closed inside compact $\mathscr{P}(Y_\Sigma)$, hence compact.

S-I.12.12 — Bernoulli scenery and entropy

For words u, v,

$$
(\beta_p)_u([v]) = \frac{\beta_p([uv])}{\beta_p([u])} = \frac{p_u p_v}{p_u} = p_v = \beta_p([v]). \tag{S.I.12.25}
$$

Cylinders determine the probability, so every positive component equals β_p.

For a typical sequence, the strong law gives

$$\frac{1}{n}\log\beta_p([\omega|_n]) = \frac{1}{n}\sum_{j=1}^{n}\log p_{\omega_j} \longrightarrow \sum_i p_i \log p_i = -h(p). \tag{S.I.12.26}$$

A radius ρ^n ultrametric ball is a prefix cylinder up to one harmless index. Dividing (S.I.12.26) by $\log(\rho^n) = n\log\rho$ proves

$$\lim_{r\downarrow 0}\frac{\log\beta_p(B_{d_\rho}(\omega,r))}{\log r} = \frac{h(p)}{|\log\rho|} \tag{S.I.12.27}$$

almost surely.

S-I.12.13 — A boundary-null tangent computation

Lebesgue measure gives zero mass to ∂B and $c\mathscr{L}^d(B) > 0$. L-I.12.04 therefore gives

$$\mu^{\square}_{x,t_n} \longrightarrow \frac{(c\mathscr{L}^d)|_B}{c\mathscr{L}^d(B)} = \frac{\mathscr{L}^d|_B}{\mathscr{L}^d(B)}. \tag{S.I.12.28}$$

The tangent scalar c disappears under probability normalization.

S-I.12.14 — The canonical Moran energy threshold

The Frostman estimate is

$$\mu(B(x,r)) \le Cr^s. \tag{S.I.12.29}$$

For $0 < t < s$, layer cake gives

$$\begin{aligned} I_t(\mu) &= t\int\int_0^\infty r^{-t-1}\mu(B(x,r))\,dr\,d\mu(x) \\ &\le tC\int_0^1 r^{s-t-1}\,dr + t\int_1^\infty r^{-t-1}\,dr \\ &= \frac{tC}{s-t} + 1 < \infty. \end{aligned} \tag{S.I.12.30}$$

Finally, $p_i = r_i^s$ gives

$$h(p) = -\sum_i r_i^s \log(r_i^s) = -s\sum_i r_i^s \log r_i = s\chi(p). \tag{S.I.12.31}$$

No dimension theorem is used in this algebraic identity.

References

[Bes39] A. S. Besicovitch. “On the Fundamental Geometrical Properties of Linearly Measurable Plane Sets of Points III”. In: *Mathematische Annalen* 116 (1939), pp. 349–357.

[Fal03] K. Falconer. *Fractal Geometry: Mathematical Foundations and Applications.* 2nd ed. Wiley, 2003.

[Fen07] D.-J. Feng. “Gibbs Properties of Self-Conformal Measures and the Multifractal Formalism”. In: *Ergodic Theory and Dynamical Systems* 27 (2007), pp. 787–812. DOI: 10.1017/S0143385706000952.

[FH09] D.-J. Feng and H. Hu. “Dimension Theory of Iterated Function Systems”. In: *Communications on Pure and Applied Mathematics* 62.11 (2009), pp. 1435–1500. DOI: 10.1002/cpa.20276. arXiv: 1002.2036.

[Fra20] J. M. Fraser. *Assouad Dimension and Fractal Geometry.* Vol. 222. Cambridge Tracts in Mathematics. Cambridge University Press, 2020. DOI: 10.1017/9781108778459. arXiv: 2005.03763.

[Fur08] H. Furstenberg. “Ergodic Fractal Measures and Dimension Conservation”. In: *Ergodic Theory and Dynamical Systems* 28.2 (2008), pp. 405–422. DOI: 10.1017/S0143385708000084.

[Hau19] F. Hausdorff. “Dimension und äußeres Maß”. In: *Mathematische Annalen* 79 (1919), pp. 157–179. DOI: 10.1007/BF01457179.

[Hoc13] M. Hochman. *Dynamics on Fractals and Fractal Distributions.* Version 2, revised 30 July 2013; preprint. 2013. arXiv: 1008.3731.

[How96] J. D. Howroyd. “On Hausdorff and Packing Dimension of Product Spaces”. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 119 (1996), pp. 715–727. DOI: 10.1017/S0305004100074545.

[Hut81] J. E. Hutchinson. “Fractals and Self Similarity”. In: *Indiana University Mathematics Journal* 30.5 (1981), pp. 713–747. DOI: 10.1512/iumj.1981.30.30055.

[Kna05a] A. W. Knapp. *Advanced Real Analysis.* Birkhäuser, 2005. ISBN: 978-0-8176-4382-9.

[Kna05b] A. W. Knapp. *Basic Real Analysis.* Birkhäuser, 2005. ISBN: 978-0-8176-3250-2.

[Mat22] P. Mattila. *Rectifiability: A Survey.* Version 3, revised 1 March 2022. 2022. arXiv: 2112.00540.

[Mat75] P. Mattila. "Hausdorff m Regular and Rectifiable Sets in n-Space". In: *Transactions of the American Mathematical Society* 205 (1975), pp. 263–274. DOI: 10.1090/S0002-9947-1975-0357741-4.

[Mat95] P. Mattila. *Geometry of Sets and Measures in Euclidean Spaces.* Cambridge University Press, 1995.

[OR26] T. Orponen and A. Rutar. *Visibility Problem in the Plane.* Version 2, revised 25 August 2026; preprint. 2026. arXiv: 2606.06965.

[Orp23] T. Orponen. "On the Dimension of Visible Parts". In: *Journal of the European Mathematical Society* 25.5 (2023), pp. 1969–1983. DOI: 10.4171/JEMS/1230. arXiv: 1912.10898.

[Pre87] D. Preiss. "Geometry of Measures in R^{n}: Distribution, Rectifiability, and Densities". In: *Annals of Mathematics* 125.3 (1987), pp. 537–643. DOI: 10.2307/1971410.

[RS14] M. Rams and K. Simon. "The Dimension of Projections of Fractal Percolations". In: *Journal of Statistical Physics* 154 (2014), pp. 633–655. DOI: 10.1007/s10955-013-0886-6.

[Shm11] P. Shmerkin. "Porosity, Dimension, and Local Entropies: A Survey". In: *Revista de la Unión Matemática Argentina* 52.2 (2011), pp. 81–103. arXiv: 1110.5682.

[Shm12] P. Shmerkin. "The Dimension of Weakly Mean Porous Measures: A Probabilistic Approach". In: *International Mathematics Research Notices* 2012.9 (2012), pp. 2010–2033. DOI: 10.1093/imrn/rnr085. arXiv: 1010.1394.

[SS05] E. M. Stein and R. Shakarchi. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces.* Vol. 3. Princeton Lectures in Analysis. Princeton University Press, 2005. ISBN: 978-0-691-11386-9.

[SS18] P. Shmerkin and V. Suomala. "Spatially Independent Martingales, Intersections, and Applications". In: *Memoirs of the American Mathematical Society* 251.1195 (2018). DOI: 10.1090/memo/1195. arXiv: 1409.6707.

[Tes06] B. Testud. "Phase Transitions for the Multifractal Analysis of Self-Similar Measures". In: *Nonlinearity* 19.5 (2006), pp. 1201–1217. DOI: 10.1088/0951-7715/19/5/009. arXiv: math/0604035.

[WWW16] C. Wei, S. Wen, and Z. Wen. "Remarks on Dimensions of Cartesian Product Sets". In: *Fractals* 24.3 (2016), p. 1650031. DOI: 10.1142/S0218348X16500316. arXiv: 1501.01713.

Subject and Result Index

Results