Exercise 3 Solutions

THURSDAY, JANUARY 29

1. (i) For each $n \in \mathbb{N} \cup \{0\}$, set

$$\mu_n := \sum_{\mathtt{i} \in \mathcal{I}^n} p_{\mathtt{i}} \delta_{f_{\mathtt{i}}(z)}.$$

Let $\Psi\colon \mathcal{P}(X) \to \mathcal{P}(X)$ denote the map

$$\Psi(\mu) = \sum_{i \in \mathcal{I}} p_i \mu \circ f_i^{-1}.$$

Also, note that if $g\colon X \to X$ is a function and $x \in X$, then $\delta_x \circ g^{-1} = \delta_{g(x)}$. Therefore, we may compute

$$\Psi(\mu_n) = \sum_{j \in \mathcal{I}} \sum_{\mathtt{i} \in \mathcal{I}^n} p_j p_{\mathtt{i}} \delta_{f_{\mathtt{i}}(z)} \circ f_j^{-1} = \sum_{\mathtt{j} \in \mathcal{I}^{n+1}} p_{\mathtt{j}} \delta_{f_{\mathtt{j}}(z)} = \mu_{n+1}.$$

Thus by the Banach contraction mapping principle,

$$\lim_{n \to \infty} \mu_n = \lim_{n \to \infty} \Psi^n(\mu) = \mu_{\boldsymbol{p}}$$

as required.

(ii) First, let f be any 1-Lipschitz function on F. Let $a = \min_{x \in F} f(x)$ and $b = \max_{x \in F} f(x)$. Since f is 1-Lipschitz, $0 \leq b - a \leq \operatorname{diam} F$. Moreover, if $\mu \in \mathcal{P}(X)$, then

$$a \leq \int_X f \,\mathrm{d}\mu \leq b.$$

In particular, if $\mu, \nu \in \mathcal{P}(X)$ are arbitrary and f is any 1-Lipschitz function,

$$\left| \int_X f \,\mathrm{d}\mu - \int_X f \,\mathrm{d}\nu \right| \leq b - a \leq \operatorname{diam} F.$$

Thus $\operatorname{diam} \mathcal{P}(F) \leq \operatorname{diam} F$.
Conversely, since F is compact, get $y, z \in F$ such that $d(y, z) = \operatorname{diam} F$. Consider the measures δ_y and δ_z and the function $g(x) = d(x, y)$. Note that g is 1-Lipschitz by the triangle inequality; and moreover

$$\left| \int_X g \,\mathrm{d}\delta_z - \int_X g \,\mathrm{d}\delta_y \right| = d(y, z) = \operatorname{diam} F.$$

Therefore $\operatorname{diam}\mathcal{P}(F) \geq d_{\mathcal{P}}(\delta_x, \delta_y) \geq \operatorname{diam} F$, as required.
We actually showed that the embedding $x \mapsto \delta_x$ is an isometry. The situation is the same for the Hausdorff metric, with the map $x \mapsto \{x\}$.

(iii) Let $z \in Q$ be fixed and let $0 \leq \lambda < 1$ denote the maximal contraction ratio of any function f_i. By (ii), if $\mathtt{i} \in \mathcal{I}^n$, then

$$d_{\mathcal{P}}(\delta_{f_{\mathtt{i}}}(z), \nu_{\mathtt{i}}) \leq \operatorname{diam} f_{\mathtt{i}}(Q) \leq (\operatorname{diam} Q)\lambda^n$$

so that

$$d_{\mathcal{P}}\left(\sum_{\mathtt{i}\in\mathcal{I}^n} p_{\mathtt{i}}\delta_{f_{\mathtt{i}}}(z), \sum_{\mathtt{i}\in\mathcal{I}^n} p_{\mathtt{i}}\nu_{\mathtt{i}}\right) \leq \sum_{\mathtt{i}\in\mathcal{I}^n} p_{\mathtt{i}} d_{\mathcal{P}}(\delta_{f_{\mathtt{i}}(z)}, \nu_{\mathtt{i}}) \leq (\operatorname{diam} Q)\lambda^n.$$

Therefore we conclude by (i) that

$$\mu_{\boldsymbol{p}} = \lim_{n\to\infty} \mu_n = \lim_{n\to\infty} \sum_{\mathtt{i}\in\mathcal{I}^n} p_{\mathtt{i}}\nu_{\mathtt{i}}$$

as required.

2. (i) Since K is not a singleton, get points $x, y \in K$ with $d(x,y) > 0$. Set $\delta = d(x,y)/4$ and let $m \in \mathbb{N}$ be sufficiently large so that $\operatorname{diam} f_{\mathtt{i}}(K) \leq \delta$ for all $\mathtt{i} \in \mathcal{I}^m$. Now, since $x \in K$, there is some $\mathtt{i}_x \in \mathcal{I}^m$ so that $x \in f_{\mathtt{i}_x}(K) \subset B(x,\delta)$. The same is true for $y \in K$.
Now if $z \in \mathbb{R}$ is arbitrary, the choice of δ guarantees that $B(z,\delta)$ intersects at most one of $B(x,\delta)$ or $B(y,\delta)$. If $B(z,\delta) \cap B(x,\delta) = \varnothing$, then $f_{\mathtt{i}_x}(K) \cap B(z,\delta) = \varnothing$, and similarly with x and y swapped.

(ii) Set

$$\xi = 1 - \min_{\mathtt{i}\in\mathcal{I}^m} p_{\mathtt{i}}.$$

Note that $\xi \in (0,1)$ since $p_i > 0$ for all $i \in \mathcal{I}$. We prove by induction for all $n \in \mathbb{N}\cup\{0\}$ that, for all $z \in \mathbb{R}$,

$$\mu_{\boldsymbol{p}} B(z, \delta r^{(n-1)m}) \leq \xi^n.$$

The case $n = 0$ is trivial since $\mu_{\boldsymbol{p}}$ is a probability measure. Thus assume that the claim holds for $n \in \mathbb{N}\cup\{0\}$. Let $z \in \mathbb{R}$ be arbitrary. Iterating the self-similarity relationship,

$$\mu_{\boldsymbol{p}} B(z, \delta r^{nm}) = \sum_{\mathtt{i}\in\mathcal{I}^m} p_{\mathtt{i}}\mu_{\boldsymbol{p}}\big(f_{\mathtt{i}}^{-1}(B(z,\delta r^{nm}))\big).$$

Let us make two observations about the sum on the right. Firstly, by (i), there is some $\mathtt{j} \in \mathcal{I}^m$ (depending on z) such that $f_{\mathtt{j}}^{-1}(B(z,\delta r^{nm})) \cap K = \varnothing$, and in particular

$$\mu_{\boldsymbol{p}}\big(f_{\mathtt{j}}^{-1}(B(z,\delta r^{nm}))\big) = 0.$$

Secondly, for $\mathtt{i} \in \mathcal{I}^m$, since $f_{\mathtt{i}}$ is a similarity map with similarity ratio r^m,

$$f_{\mathtt{i}}^{-1}(B(z, \delta r^{nm})) = B(f_{\mathtt{i}}^{-1}(z), \delta r^{(n-1)m}).$$

Applying these above observations and then the inductive hypothesis,

$$\begin{aligned}\mu_{\boldsymbol{p}} B(z, \delta r^{nm}) &= \sum_{\mathtt{i} \in \mathcal{I}^m \setminus \{\mathtt{j}\}} p_{\mathtt{i}} \mu_{\boldsymbol{p}}\big(B(f_{\mathtt{i}}^{-1}(z), \delta r^{(n-1)m})\big) \\ &\leq \sum_{\mathtt{i} \in \mathcal{I}^m \setminus \{\mathtt{j}\}} p_{\mathtt{i}} \xi^n \\ &\leq \xi^{n+1}\end{aligned}$$

by the definition of ξ, as required.

(iii) Let t be such that $r^{mt} = \xi$ and observe that $t > 0$ since $\xi < 1$. Then, let $C = \delta^{-t}$. First, if $\operatorname{diam} A \geq \delta$, then

$$\mu(A) \leq 1 \leq \left(\frac{\operatorname{diam} A}{\delta}\right)^t \leq C(\operatorname{diam} A)^t.$$

Otherwise, let $n \in \mathbb{N}$ be such that $\delta r^{nm} \leq \operatorname{diam} A < \delta r^{(n-1)m}$. Then by (ii), since $A \subset B(z, \delta r^{(n-1)m})$ for any $z \in A$,

$$\mu(A) \leq \mu(B(z, \delta r^{(n-1)m})) \leq \xi^n = \delta^{-t}(\delta r^{nm})^t \leq C(\operatorname{diam} A)^t$$

as required.

3. Set

$$E = \{x \in \mathbb{R}^d : \underline{\dim}_{\text{loc}}(\mu, x) < s\}.$$

First, observe that $x \in E$ if and only if there is an $n \in \mathbb{N}$ so that for all arbitrarily small $r > 0$ such that $\mu(B(x,r)) \geq (r/2)^{s-1/n}$. Moreover, for such r, if $k \in \mathbb{Z}$ is such that $2^{-(k+1)} < r \leq 2^{-k}$, then

$$\mu(B(x, 2^{-k})) \geq \mu(B(x,r)) \geq (r/2)^s \geq (2^{-k})^{s-1/n}.$$

Therefore, $x \in E$ if and only if there exists $n \in \mathbb{N}$ and arbitrarily large $k \in \mathbb{N}$ such that $\mu(B(x, 2^{-k})) \geq 2^{-k(s-1/n)}$. Therefore,

$$E = \bigcup_{n=1}^{\infty} \bigcap_{j=1}^{\infty} \bigcup_{k=j}^{\infty} E(2^{-k}, 2^{-k(s-1/n)}).$$

where for $r > 0$ and $\alpha > 0$, we set

$$E(r, \alpha) = \big\{x \in \mathbb{R}^d : \mu(B(x,r)) \geq \alpha\big\}.$$

Thus it remains to show that the sets $E(r,\alpha)$ are Borel sets. Actually, we will show that they are closed sets. First, note that

$$\mu(B(x,r)) = \mu\left(\bigcap_{n=1}^{\infty} B(x, r+1/n)\right) = \lim_{n\to\infty} \mu(B(x, r+1/n).$$

Now let $x \notin E(r,\alpha)$ be arbitrary. The measure property above shows that there is an $n \in \mathbb{N}$ such that $\mu(B(x, r-1/n)) < \alpha$. Thus if $|y-x| \le 1/n$, then $\mu(B(y,r)) < \alpha$ as well, so $y \notin E(r,\alpha)$. Therefore $E(r,\alpha)$ is closed, as claimed.

4. (i) First, instead of considering the action of the IFS, let's replace it with a new operator which acts on left endpoints and rescales: for a number $x \ge 0$, define

$$\psi(x) = \{3x, 3x+2/3, 3x+2\}$$

and as usual extend the definition to sets. A direct computation shows for $n \in \mathbb{N} \cup \{0\}$ that

$$\psi^n(\{0\}) = \{3^n f_{\mathtt{i}}(0) : \mathtt{i} \in \mathcal{I}^n\}.$$

Moreover, an easy induction argument shows that $\psi^n(\{0\}) \subset 3^{-n}\mathbb{Z}$. This implies the claim.

(ii) Note that the count in question is exactly equal to $\#\psi^n(\{0\})$, since $f_{\mathtt{i}} = f_{\mathtt{j}}$ if and only if $f_{\mathtt{i}}(0) = f_{\mathtt{j}}(0)$. Our main concern is to keep track of cases in which $\psi(x) \cap \psi(y) \ne \varnothing$. Indeed, $\#\psi^2(\{0\}) = 8$ since $f_1 \circ f_3 = f_2 \circ f_1$.

Our count will be powered by the following elementary combinatorial observation.

Lemma 1. *For all $n \in \mathbb{N} \cup \{0\}$, the following hold:*

(a) $\psi^n(\{0\}) \subset \bigcup_{k=0}^{\infty} \{2k, 2k+2/3\}$.

(b) The sets $\psi(\{2k, 2k+2/3\})$ are disjoint for distinct $k \in \mathbb{N} \cup \{0\}$.

Proof. The proof of (a) follows by a simple induction argument.

To see (b), observe that if $x > y + 2/3$, then $\psi(\{x\}) \cap \psi(\{y\}) = \varnothing$. Since the sets $\{2k, 2k+2/3\}$ are pairwise separated by distance at least $4/3$, their images must be disjoint. □

Since we don't care about the values of k, we can represent $\psi^n(\{0\})$ as a sequence of digits in $\{1,2\}$, where 1 refers to a component $\{2k\}$ and 2 refers to a component $\{2k, 2k+2/3\}$. (In principle, we might also need to track a component like $\{2k+2/3\}$ but we will see below that it cannot appear.)

We can compute that

$$\psi(\{2k\}) = \{6k, 6k+2/3, 6k+2\}$$
$$\psi(\{2k, 2k+2/3\}) = \{6k, 6k+2/3, 6k+2, 6k+2+2/3, 6k+4\}$$

so a component of type 1 is replaced with a sequence of components $(2,1)$, and a component of type 2 is replaced with a sequence of components $(2,2,1)$. Moreover, the singleton $\{0\}$ is a component of type 1.

For example, iterating the above rules:

- $\psi^0(\{0\}) = (1)$.
- $\psi^1(\{0\}) = (2,1)$.
- $\psi^2(\{0\}) = (2,2,1,2,1)$.

Since we only care about the number of times that component 1 appears or that component 2 appears, we can keep track of the counts in ψ^n using the matrix and vector

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \qquad v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

where $A_{i,j}$ is the number of components of type i which appear in the replacement of a component of type j, and v is the initial component count. In particular, writing

$$(a_n, b_n) := A^n v,$$

a_n is the number of times type 1 occurs and b_n is the number of times type 2 occurs. Since a component of type 1 corresponds to a single element and a component of type 2 corresponds to two elements, the claim follows (taking $c = 2$).

(iii) First, note that $\|(1,0)\cdot A^n\|_1 \le \|A^n\|_1$, and conversely since $(1,0)\cdot A$ is a positive vector, for $n \ge 1$, if c is the smallest element of $(1,0)\cdot A$,

$$\|(1,0)\cdot A^n\|_1 = \left\|(1,0)\cdot A \cdot A^{n-1}\right\|_1 \ge c\left\|A^{n-1}\right\|_1.$$

Therefore, by Gelfand's formula,

$$\lim_{n\to\infty} \frac{\log \#\{f_{\mathtt{i}} : \mathtt{i} \in \mathcal{I}^n\}}{n} = \lim_{n\to\infty} \log \|A^n\|_1^{1/n} = \log \rho$$

where $\rho = (3+\sqrt{5})/2$ is the maximal eigenvalue of A.

To complete the proof, we show that

$$N_{3^{-n}}(K) \le \#\{f_{\mathtt{i}} : \mathtt{i} \in \mathcal{I}^n\} \le 7 N_{3^{-n}}(K).$$

The lower inequality is immediate, since $K \subset \bigcup_{\mathtt{i}\in\mathcal{I}^n} f_{\mathtt{i}}(K)$ and, since $K \subset [0,1]$, $\operatorname{diam} f_{\mathtt{i}}(K) \le 3^{-n}$. For the upper inequality, fix an optimal cover of K using balls of radius 3^{-n}. Then by the pigeonhole principle and (ii), any ball $B(x, 3^{-n})$ can intersect at most 7 distinct images $f_{\mathtt{i}}(K)$. Therefore we conclude that

$$\begin{aligned} \dim_{\mathrm{H}} K = \dim_{\mathrm{B}} K &= \lim_{n\to\infty} \frac{\log N_{3^{-n}}(K)}{n \log 3} \\ &= \lim_{n\to\infty} \frac{\log \#\{f_{\mathtt{i}} : \mathtt{i} \in \mathcal{I}^n\}}{n \log 3} \end{aligned}$$

$$= \frac{\log \rho}{\log 3}$$

as claimed.