Annales de la Faculté des Sciences de Toulouse
MATHÉMATIQUES

DAVID ALDOUS, GUILLAUME BLANC AND NICOLAS CURIEN

The distance problem on measured metric spaces

Tome XXXIV, nº 5 (2025), p. 1345–1365.

https://doi.org/10.5802/afst.1835

Publication membre du centre Mersenne pour l'édition scientifique ouverte
http://www.centre-mersenne.org/
e-ISSN : 2258-7519

Annales de la faculté des sciences de Toulouse Volume XXXIV, n° 5, 2025
pp. 1345-1365

The distance problem on measured metric spaces (*)

David Aldous (1), Guillaume Blanc (2) and Nicolas Curien (3)

Abstract. — What distributions arise as the distribution of the distance between two typical points in some measured metric space? This seems to be a surprisingly subtle problem. We conjecture that every distribution with a density function whose support contains 0 does arise in this way, and give some partial results in that direction.

Résumé. — Quelles distributions apparaissent comme lois de la distance entre deux points typiques dans un espace métrique mesuré ? Cette question naturelle semble étonnamment subtile. Nous conjecturons que toute distribution possédant une densité dont le support contient 0 peut apparaître de cette maniere. Nous étayons cette conjecture par des résultats partiels.

1. Introduction and main results

The problem below has apparently not previously been studied. It seems quite natural in itself, and some statistical motivation is given towards the end of the introduction. Let μ be a Borel probability measure on a complete *separable* metric space (S, d), and let ξ_1 and ξ_2 be independent random variables with distribution μ. The random variable $D := d(\xi_1, \xi_2)$ has some distribution, say θ, on $\mathbb{R}_+$. We call a distribution θ on $\mathbb{R}_+$ *feasible* if it can arise in this way, and we say that (S, d, μ) *achieves* θ.

(*) Reçu le 27 avril 2024, accepté le 20 décembre 2024.

Keywords: Characterization of distribution, construction of distribution, distance, measured metric space, tree.

2020 *Mathematics Subject Classification:* 60B05, 60E99.

(1) Department of Statistics, 367 Evans Hall # 3860, U.C. Berkeley, CA 94720, USA — aldousdj@berkeley.edu

(2) EPFL, MA A2 397, (Bâtiment MA) Station 8 1015 Lausanne, Switzerland — guillaume.blanc@epfl.ch

(3) Université Paris-Saclay, BP74, 38402 SMH Cedex, France — nicolas.curien@gmail.com

The second and third author were supported by the ERC Consolidator Grant "SuPerGRandMa" 101087572.

Article proposé par Xavier Tolsa.

PROBLEM 1.1. — *Describe the set of feasible distributions.*

The results stated in this introduction, and proved in later sections, are essentially all we know about this problem. Before coming to them, let us stress that the separability requirement is fundamental. For example, it guarantees that D is a random variable, since in this setting the Borel σ-algebra $\mathcal{B}(S \times S)$ agrees with $\mathcal{B}(S) \otimes \mathcal{B}(S)$. Besides, the construction of interesting Borel measures on non-separable metric spaces is problematic: for instance, it is known that (S, d) must be separable as soon as it supports a boundedly finite measure μ of full support [7, Theorem 4.1].

Let us mention a related problem (summarized in Section 5): when do two different measured metric spaces have the same distribution of D? We do not have any new answers.

In our problem, a first consequence of the separability assumption is that θ must have 0 in its support (Proposition 1.2). But Proposition 2.2 will show that the converse is false: there are distributions on $\mathbb{R}_+$ with 0 in their support that are not feasible.

PROPOSITION 1.2. — *Any feasible distribution must have* 0 *in its support.*

Informally, in nice homogeneous settings, we expect that

$$\theta[0, \varepsilon] = \mathbb{P}(D \leqslant \varepsilon) \approx \varepsilon^{\dim} \quad \text{as } \varepsilon \longrightarrow 0^+,$$

where dim is the "dimension" of (S, d). See Remark 2.1 for a precise statement in this direction. This suggests that if θ assigns small mass around 0, then (S, d) must be "large", and indeed Proposition 1.2 is a caricature of this fact: if θ assigns no mass around 0, then (S, d) must be so large that it cannot be separable. There is however no obstruction in general:

PROPOSITION 1.3. — *For every non-decreasing function $F : \mathbb{R}_+^* \to]0, 1]$, it is possible to construct a compact measured metric space (S, d, μ) such that $\mathbb{P}(D \leqslant \varepsilon) \leqslant F(\varepsilon)$ for all sufficiently small ε.*

To clarify notation:

$$\mathbb{R}_+ := [0, \infty[\quad \mathbb{R}_+^* :=]0, \infty[\quad \mathbb{N} := \{0, 1, 2, \ldots\} \quad \mathbb{N}^* := \{1, 2, \ldots\}.$$

Moving away from the behavior at 0^+, a common first thought on the problem is that the triangle inequality should put some constraint on the general structure of a feasible θ. Perhaps surprisingly, we first show that any distribution on $\mathbb{R}_+$ whose support is a finite set containing 0 is feasible:

THEOREM 1.4. — *For every distribution of the form* $\theta = p_0 \cdot \delta_0 + p_1 \cdot \delta_{d_1} + \cdots + p_k \cdot \delta_{d_k}$, *where* $k \in \mathbb{N}$ *and* $0 = d_0 < d_1 < \cdots < d_k$, *and*[(1)] *where* $p_0, p_1, \ldots, p_k \in]0,1[$ *are such that* $p_0 + p_1 + \cdots + p_k = 1$, *there exists a compact measured metric space* (S, d, μ) *that achieves* θ.

Our proof of Theorem 1.4 relies on a tree construction. Although it might be appealing to approximate a general target distribution on $\mathbb{R}_+$ by finitely supported distributions θ_n, each of which is achieved by some (S_n, d_n, μ_n), and hope to get something in the limit $n \to \infty$, we discuss in Remark 3.2 why this is not possible in general. On the other hand, the tree constructions of Theorem 1.4 and Proposition 1.3 can be used to prove the following result:

PROPOSITION 1.5. — *Let* f *be a continuous probability density function on* $\mathbb{R}_+$, *and assume that there exists* $\eta > 0$ *such that* $f(x) > 0$ *for all* $x \in]0, \eta]$. *Then, for every* $\zeta > 0$, *there exists a compact measured metric space* (S, d, μ) *such that the distribution* θ *achieved by* (S, d, μ) *has a continuous density* g *on* $\mathbb{R}_+$, *for some* $g \leqslant (1+\zeta) \cdot f$.

A weaker result is that, for such f, there exist feasible densities g_n such that $\int |g_n(x) - g(x)| \mathrm{d}x \to 0$. This weaker form could be derived quite easily from Theorem 1.4, by replacing each leaf of the tree by a short interval. The specific stronger form in Proposition 1.5 requires a more elaborate construction. The form is motivated by the following observation: if there were a space (S', d, μ') achieving the distribution θ with density f, then take a compact set K such that $\mu'(K) \geqslant 1 - \varepsilon$, and then the *compact* space $(K, d, \mu'(\cdot|K))$ achieves a distribution θ^* with $\theta^* \leqslant (1-\varepsilon)^{-2}\theta$.

A straightforward corollary of Proposition 1.5 is that every distribution with a suitable density is achieved by some *random* compact measured metric space (Corollary 1.6). In this setting, we first sample the random compact measured metric space (S, d, μ), then conditionally on (S, d, μ), we sample two independent random variables ξ_1 and ξ_2 with distribution μ, and we consider the annealed distribution of $d(\xi_1, \xi_2)$. See Section 4 for details.

COROLLARY 1.6. — *Let* f *be a continuous probability density function on* $\mathbb{R}_+$, *and assume that there exists* $\eta > 0$ *such that* $f(x) > 0$ *for all* $x \in]0, \eta]$. *Then, there exists a* random *compact measured metric space* (S, d, μ) *that achieves the distribution with density* f.

We highlight that the following problem remains open.

PROBLEM 1.7. — *Prove that for every probability density function* f *on* $\mathbb{R}_+$ *whose support contains* 0, *the distribution with density* f *is feasible.*

(1) In the trivial case $k = 0$ we have $p_0 = 1$.

We have not studied the case where θ is a singular non-atomic distribution.

Background and motivation

We finish this introduction with some statistical motivation for the problem. Consider a Borel probability measure μ on a complete separable metric space (S, d), and let $\xi_1, \xi_2, \ldots$ be a sequence of independent random variables with distribution μ. In this setting, it is known that the distribution of the infinite array $A_\infty = (d(\xi_i, \xi_j)\,;\, i, j \geqslant 1)$ determines (S, d, μ) up to measure-preserving isometry (we refer to [13, 14], see also [8, Chapter $3\frac{1}{2}_+$]), although there is no known explicit characterization[(2)] of the possible distributions of A_∞. See Remark 3.2 for the connections with Theorem 1.4. For each $n \geqslant 2$, the distribution of

$$A_n := (d(\xi_i, \xi_j)\,;\, 1 \leqslant i, j \leqslant n), \tag{1.1}$$

is usually called the *n-point function/distribution*. Describing the possible distributions of A_∞ is essentially equivalent to understanding, for each $n \geqslant 2$, what constraints on the $(n+1)$-point function are given by the n-point function. The fact that we do not know which distributions are 2-point functions (Problem 1.1) makes it hard to proceed, and our work is a modest start. After having described the set of feasible distributions, one could continue to ask for the possible 3-point functions, where now the triangle inequality would come into play.

Here is the statistical modelling context. Suppose that we have a large database of different objects of the same type,[(3)] and we want to decide whether a new object is significantly similar to some object in the database (more similar than would be expected "by chance"). A natural model in this general context is that there is a space (S, d) of possible objects with distances, and that our database objects and the new object are i.i.d. samples $(\xi_i,\, i \geqslant 1)$ from a probability measure μ on S. However, we do not observe S or μ, we observe only the distances $A_n = (d(\xi_i, \xi_j)\,;\, 1 \leqslant i, j \leqslant n)$ between these objects. One would like to devise an algorithm that, given A_n and $(d(\xi_{n+1}, \xi_i)\,;\, 1 \leqslant i \leqslant n)$, decides whether ξ_{n+1} is "too close to one of

[(2)] This setting contrasts with the case where d ranges over all *measurable* functions, in which case the analogous infinite arrays are characterized by exchangeability properties: see [2, 9, 10] and a more recent nice account by Tim Austin [4].

[(3)] E.g., fingerprints, human DNA (in the forensic context), facial recognition, musical tunes or lyrics (there are more than 100 million songs online [5]), plots of murder mystery novels (in the copyright context), ...

$\{\xi_1, \dots, \xi_n\}$ to just be chance", which would then suggest some causal relationship. In this context, we seek to make inferences which are "universal", i.e, do not depend on (S, d, μ), and this motivates the study of relationships between n-point functions.

2. Properties of the support of feasible distributions

As mentioned in the introduction, a first consequence of the separability assumption is that any feasible distribution must have 0 in its support (Proposition 1.2). Let us prove this now.

Proof of Proposition 1.2. — Let $\{s_n,\, n \geqslant 1\}$ be a dense countable subset of S. For each $\varepsilon > 0$, since the balls $(B(s_n, \varepsilon/2)\,;\, n \geqslant 1)$ cover S, there exists $n \in \mathbb{N}^*$ such that $\mu(B(s_n, \varepsilon/2)) > 0$. It follows that

$$\mathbb{P}(D \leqslant \varepsilon) \geqslant \mu^2(B(s_n, \varepsilon/2)) > 0. \qquad \square$$

Remark 2.1. — If (S, d) is compact, then the previous argument can be sharpened to show that

$$\mathbb{P}(D \leqslant \varepsilon) \geqslant 1/M(\varepsilon/2), \tag{2.1}$$

where $M(\varepsilon/2)$ is the minimal number of balls with radius $\varepsilon/2$ needed to cover S. Indeed, fix a covering $(B(s_i, \varepsilon/2)\,;\, i \in [\![1, k]\!])$ of S by some k balls[(4)] of radius $\varepsilon/2$. Then, let

$$B_i = B(s_i, \varepsilon/2) \setminus (B(s_1, \varepsilon/2) \cup \ldots \cup B(s_{i-1}, \varepsilon/2)) \quad \text{for all } i \in [\![1, k]\!],$$

so that $B_1 \cup \ldots \cup B_k = S$, where the B_i have diameter at most ε. On the one hand, we have

$$\mathbb{P}(D \leqslant \varepsilon) \geqslant \mathbb{P}\left(\bigsqcup_{i=1}^{k} (\xi_1, \xi_2 \in B_i)\right) = \sum_{i=1}^{k} \mu^2(B_i).$$

On the other hand, using the Cauchy–Schwarz inequality, we have

$$1 = \mu\left(\bigsqcup_{i=1}^{k} B_i\right)^2 = \left(\sum_{i=1}^{k} \mu(B_i)\right)^2 \leqslant k \cdot \sum_{i=1}^{k} \mu^2(B_i).$$

Thus, we get

$$\mathbb{P}(D \leqslant \varepsilon) \geqslant k^{-1}.$$

Taking k to be minimal proves (2.1).

(4) Notation $[\![1, k]\!]$ indicates the interval of integers.

The next example[5] shows that having 0 in its support is not a *sufficient* condition for a distribution on $\mathbb{R}_+$ to be feasible.

PROPOSITION 2.2. — *Any distribution of the form $\theta = p\cdot\delta_0 + (1-p)\cdot\theta'$, where $p \in]0,1[$, and where θ' is a non-atomic distribution on $\mathbb{R}_+$ whose support does not contain 0, is not feasible.*

Proof. — To prove this by contradiction, suppose θ is achieved by some (S,d,μ). If μ is non-atomic then $\mathbb{P}(D=0)=0$, whereas if μ has (at least) two atoms s_1, s_2 then $\mathbb{P}(D = d(s_1,s_2)) > 0$, each case contradicting the assumption. The remaining case is that μ has exactly one atom s_0, for which $0 < q := \mu\{s_0\} < 1$. In this case, μ has a decomposition $\mu = q\delta_{s_0} + (1-q)\mu^*$ and so

$$\theta > q^2\delta_0 + (1-q)^2\theta^* \tag{2.2}$$

where θ^* is the distribution achieved by (S,d,μ^*). By Proposition 1.2, the support of θ^* contains 0, so rewriting (2.2) to define θ' via

$$\theta = q^2\delta_0 + (1-q^2)\theta'$$

we find that θ' is non-atomic with support containing 0, contradicting the assumption. □

3. The tree constructions

In this section we use tree constructions to prove our main results, Theorem 1.4 and Proposition 1.3. For the proof of Theorem 1.4, the measured metric spaces (S,d,μ) we construct to achieve finitely supported distributions are finite rooted trees (T,ρ) with edge-lengths, equipped with the natural metric d induced by the edge-lengths, and endowed with a probability measure ν supported on the leaves. We call the tuple (T,ρ,d,ν) a *tree structure* for short. Also, recall that we use "(T,ρ,d,ν) achieves θ" as shorthand for "if ξ_1 and ξ_2 are independent random variables with distribution ν, then $d(\xi_1,\xi_2)$ has distribution θ". Theorem 1.4 is implied by:

PROPOSITION 3.1. — *For any distribution of the form $\theta = p_0\cdot\delta_0 + p_1\cdot\delta_{d_1} + \cdots + p_k\cdot\delta_{d_k}$, where $k\in\mathbb{N}$ and $0 = d_0 < d_1 < \cdots < d_k$, and where $p_0,p_1,\ldots,p_k \in]0,1[$ are such that $p_0+p_1+\cdots+p_k = 1$, there exists a tree structure (T,ρ,d,ν) that achieves θ.*

[5] See *Acknowledgements* for the origin of this example.

Proof. — Let us prove, by induction on $k \in \mathbb{N}$, the more detailed assertion

H_k: For any distribution of the form $\theta = p_0 \cdot \delta_0 + p_1 \cdot \delta_{d_1} + \cdots + p_k \cdot \delta_{d_k}$, where $0 = d_0 < d_1 < \cdots < d_k$, and where $p_0, p_1, \ldots, p_k \in]0,1[$ are such that $p_0 + p_1 + \cdots + p_k = 1$, there exists a tree structure (T, ρ, d, ν) that achieves θ, in which every leaf is at distance $d_k/2$ from the root vertex ρ'.

The construction is illustrated in Figure 3.1. Note that it is "backwards", in that the length of the edges emanating from ρ is $(d_k - d_{k-1})/2$, not $d_1/2$.

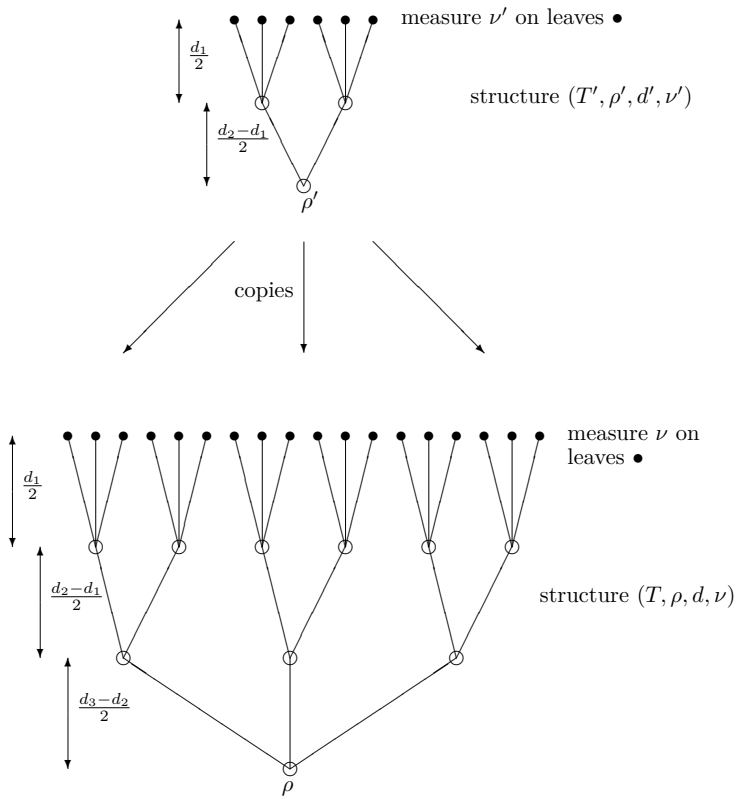

Figure 3.1. Illustrating construction of the tree structure $T = (T, \rho, d, \nu)$ for $k = 3$.

To prove H_0, consider the tree structure (T, ρ, d, ν) consisting of the single root vertex ρ. Then, fix $k \in \mathbb{N}$, assume that H_k holds, and let us prove H_{k+1}. Let θ be a distribution of the form $\theta = p_0 \cdot \delta_0 + p_1 \cdot \delta_{d_1} + \cdots + p_{k+1} \cdot \delta_{d_{k+1}}$, where $0 = d_0 < d_1 < \cdots < d_{k+1}$, and where $p_0, p_1, \ldots, p_{k+1} \in]0,1[$ are such that $p_0 + p_1 + \cdots + p_{k+1} = 1$. We can write

$$\theta = (1 - p_{k+1}) \cdot \theta' + p_{k+1} \cdot \delta_{d_{k+1}},$$

where

$$\theta' = \frac{p_0}{p_0 + \cdots + p_k} \cdot \delta_0 + \frac{p_1}{p_0 + \cdots + p_k} \cdot \delta_{d_1} + \cdots + \frac{p_k}{p_0 + \cdots + p_k} \cdot \delta_{d_k}.$$

Now, by H_k, there exists a tree structure (T', ρ', d', ν') that achieves θ', in which every leaf is at distance $d_k/2$ from the root vertex ρ'. Then, let $j \in \mathbb{N}^*$ be large enough so that we can choose $m_1, \ldots, m_j \in]0,1[$ with $m_1 + \cdots + m_j = 1$, such that $m_1^2 + \cdots + m_j^2 = 1 - p_{k+1}$. Take j copies $(T_1, \rho_1, d_1, \nu_1), \ldots, (T_j, \rho_j, d_j, \nu_j)$ of (T', ρ', d', ν'), and construct a tree structure (T, ρ, d, ν) by first drawing j edges of length $(d_{k+1} - d_k)/2$ emanating from a root vertex ρ, then grafting the $(T_i, \rho_i, d_i, \nu_i)$ onto those edges, identifying ρ_i with the end-vertex of the corresponding edge, and letting $\nu = m_1 \cdot \nu_1 + \cdots + m_j \cdot \nu_j$.

By construction, the probability measure ν is supported on the leaves of T, and every leaf of T is at distance $d_k/2 + (d_{k+1} - d_k)/2 = d_{k+1}/2$ from the root vertex ρ. Now, let us check that (T, ρ, d, ν) achieves θ. Let ξ_1 and ξ_2 be independent random variables with distribution ν. For every Borel function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$, we calculate

$$\begin{aligned}
&\mathbb{E}[\varphi(d(\xi_1, \xi_2))] \\
&\qquad = \sum_{u_1,\, u_2 \text{ leaves of } T} \varphi(d(u_1, u_2)) \cdot \nu\{u_1\} \cdot \nu\{u_2\} \\
&\qquad = \sum_{1 \leqslant i_1, i_2 \leqslant j} m_{i_1} \cdot m_{i_2} \cdot \sum_{\substack{u_1 \text{ leaf of } T_{i_1} \\ u_2 \text{ leaf of } T_{i_2}}} \varphi(d(u_1, u_2)) \cdot \nu_{i_1}\{u_1\} \cdot \nu_{i_2}\{u_2\}.
\end{aligned}$$

In the sum above, if $i_1 = i_2$, then

$$\begin{aligned}
&\sum_{\substack{u_1 \text{ leaf of } T_{i_1} \\ u_2 \text{ leaf of } T_{i_2}}} \varphi(d(u_1, u_2)) \cdot \nu_{i_1}\{u_1\} \cdot \nu_{i_2}\{u_2\} \\
&\qquad\qquad = \sum_{u_1,\, u_2 \text{ leaves of } T'} \varphi(d(u_1, u_2)) \cdot \nu'\{u_1\} \cdot \nu'\{u_2\} = \theta'(\varphi),
\end{aligned}$$

by the definition of (T', ρ', d', ν'). On the other hand, if $i_1 \neq i_2$, then the distance between any leaf u_1 of T_{i_1} and any leaf u_2 of T_{i_2} in the tree structure

(T, ρ, d, ν) is

$$d_k/2 + (d_{k+1} - d_k)/2 + (d_{k+1} - d_k)/2 + d_k/2 = d_{k+1},$$

hence

$$\sum_{\substack{u_1 \text{ leaf of } T_{i_1} \\ u_2 \text{ leaf of } T_{i_2}}} \varphi(d(u_1, u_2)) \cdot \nu_{i_1}\{u_1\} \cdot \nu_{i_2}\{u_2\}$$
$$= \sum_{\substack{u_1 \text{ leaf of } T_{i_1} \\ u_2 \text{ leaf of } T_{i_2}}} \varphi(d_{k+1}) \cdot \nu_{i_1}\{u_1\} \cdot \nu_{i_2}\{u_2\} = \varphi(d_{k+1}).$$

It follows that

$$\mathbb{E}[\varphi(d(\xi_1, \xi_2))] = \sum_{i=1}^{j} m_i^2 \cdot \theta'(\varphi) + \sum_{1 \leqslant i_1 \neq i_2 \leqslant j} m_{i_1} \cdot m_{i_2} \cdot \varphi(d_{k+1})$$
$$= (1 - p_{k+1}) \cdot \theta'(\varphi) + p_{k+1} \cdot \varphi(d_{k+1}) = \theta(\varphi). \qquad \square$$

Remark 3.2. — It is natural, for a given target distribution θ, say with a smooth and compactly supported density, to approximate θ by distributions θ_n whose support is a finite set containing 0, in such a way that $\theta_n \Rightarrow \theta$ as $n \to \infty$. By Theorem 1.4, each θ_n is achieved by some tree structure $(T_n, \rho_n, d_n, \nu_n)$, and it is natural to seek for some (sub-)sequential limits of $(T_n, \rho_n, d_n, \nu_n)_{n \in \mathbb{N}^*}$ to achieve θ. Alas, with the construction presented above, in general the sequence $(T_n, \rho_n, d_n, \nu_n)_{n \in \mathbb{N}^*}$ is not tight for the Gromov–Hausdorff–Prokhorov topology (see Section 4 for a brief reminder of that topology). Intuitively, the spaces we would end up with "in the limit" are non-separable. A cartoon of the phenomenon is the non-convergence of n-star graphs (n vertices, each connected to the same root vertex by an edge of length 1, with the uniform probability measure on the $(n+1)$ vertices) for the Gromov–Hausdorff–Prokhorov topology, although for each $k \geqslant 2$, their k-point functions converge in distribution as $n \to \infty$.

We now turn to the proof of Proposition 1.3, which is based on a different tree construction, illustrated in Figure 3.2.

Proof of Proposition 1.3. — Let $(\kappa_n)_{n \in \mathbb{N}}$ be a sequence of positive integers to be adjusted, and let T be the infinite spherically symmetric plane tree in which every node u at height n has $2\kappa_n$ children, the nodes $u1, \ldots, u(2\kappa_n)$ in the Ulam–Harris (or Neveu) notation.[6] For each $i \in [\![1, 2\kappa_n]\!]$, we set the length of the edge between u and ui to be $2^{-(n+1)}$ if i is odd, and $2 \cdot 2^{-(n+1)}$

[6] A vertex is labeled as a string $i_1 i_2 \ldots i_j$, meaning it is the i_j'th child of vertex $i_1 i_2 \ldots i_{j-1}$.

if i is even, and we denote by d the metric on T induced by these edge-lengths. If we complete T into $\overline{T} = T \sqcup \partial T$ by adding its boundary ∂T, which consists of rays $u = (u_k)_{k\in\mathbb{N}}$ emanating from the root $\varnothing$, then $(\overline{T}, d)$ is compact. Now, the distribution of the uniform non-backtracking random walk on T starting at the root yields a natural Borel probability measure π on $\overline{T}$, supported on ∂T. Let ξ_1 and ξ_2 be independent random variables with distribution π, that is independent random rays, and let us consider the distribution of $d(\varnothing, \xi_1)$ and $d(\varnothing, \xi_2)$, and of $d(\xi_1, \xi_2)$. Recall the following well known fact: if $B_1, B_2, \dots$ are independent Bernoulli random variables with success probability $1/2$, then the random variable $\sum_{n\geqslant 1} B_n \cdot 2^{-n}$ has uniform distribution on $[0, 1]$. Now the random variables $d(\varnothing, \xi_1)$ and $d(\varnothing, \xi_2)$ have the same distribution as

$$\sum_{n\geqslant 0} (1 + \mathbf{1}(I_n \text{ is even})) \cdot 2^{-(n+1)},$$

where $(I_n)_{n\in\mathbb{N}}$ is a sequence of independent random variables such that for each $n \in \mathbb{N}$, the random variable I_n has uniform distribution on $[\![1, 2\kappa_n]\!]$. Therefore, the random variables $d(\varnothing, \xi_1)$ and $d(\varnothing, \xi_2)$ have the same distribution as $1 + U$, where U has uniform distribution on $[0, 1]$.

We can continue this calculation (details deferred) to show the next lemma.

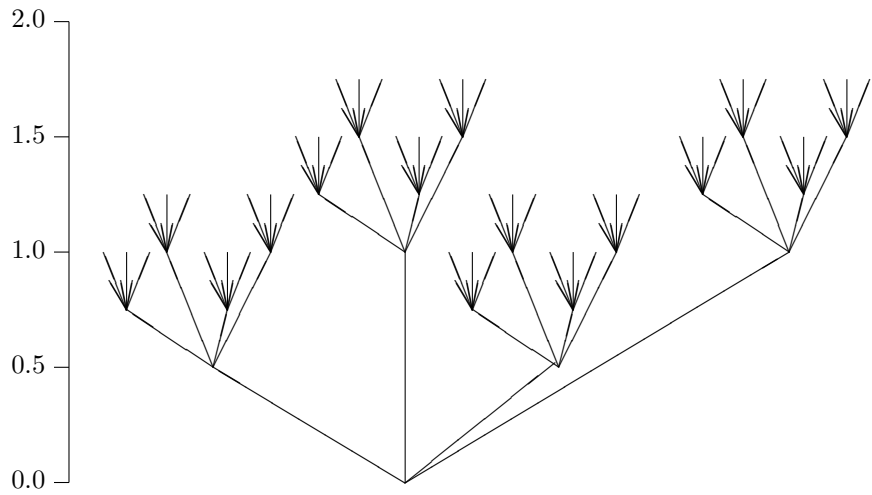

Figure 3.2. Initial stages of the tree construction for Proposition 1.3.

LEMMA 3.3. — *The random variable $d(\xi_1, \xi_2)$ has the continuous probability density function Ψ given by*

$$\Psi(x) = \sum_{n \geqslant 0} \frac{2\kappa_n - 1}{(2\kappa_0) \cdot \ldots \cdot (2\kappa_n)} \cdot \psi_n(x), \tag{3.1}$$

where for each $n \in \mathbb{N}$, the function ψ_n is a continuous probability density function supported in $[4 \cdot 2^{-(n+1)}, 4 \cdot 2^{-n}]$, and bounded by 2^{n+1} over this interval.

Granted Lemma 3.3, we complete the proof of Proposition 1.3 as follows. Choose $(\kappa_n)_{n \in \mathbb{N}}$ so that

$$(2\kappa_0) \cdot \cdots \cdot (2\kappa_n) \geqslant 1/F\Big(4 \cdot 2^{-(n+2)}\Big) \quad \text{for all } n \in \mathbb{N}.$$

This is certainly possible, because it suffices to be very crude and choose

$$\kappa_n \geqslant 1/F\Big(4 \cdot 2^{-(n+2)}\Big) \quad \text{for all } n \in \mathbb{N}.$$

This way, for each $n \in \mathbb{N}^*$, we have, for all $\varepsilon \in [4 \cdot 2^{-(n+1)}, 4 \cdot 2^{-n}]$,

$$\begin{aligned}
\int_0^\varepsilon \Psi(x)\mathrm{d}x &\leqslant \sum_{p \geqslant n} \frac{2\kappa_p - 1}{(2\kappa_0) \cdot \cdots \cdot (2\kappa_p)} \\
&= \sum_{p \geqslant n} \left(\frac{1}{(2\kappa_0) \cdot \cdots \cdot (2\kappa_{p-1})} - \frac{1}{(2\kappa_0) \cdot \cdots \cdot (2\kappa_p)} \right) \\
&= \frac{1}{(2\kappa_0) \cdot \cdots \cdot (2\kappa_{n-1})} \\
&\leqslant F\Big(4 \cdot 2^{-(n+1)}\Big) \leqslant F(\varepsilon),
\end{aligned}$$

hence $\int_0^\varepsilon \Psi(x)\mathrm{d}x \leqslant F(\varepsilon)$ for all $\varepsilon \in]0, 2]$. S o this compact measured metric space $(\overline{T}, d, \pi)$ has the properties stated in Proposition 1.3. □

Proof of Lemma 3.3. — For each $n \in \mathbb{N}$, the height $|\xi_1 \wedge \xi_2|$ of the branchpoint between two independent rays has distribution

$$\mathbb{P}(|\xi_1 \wedge \xi_2| = n) = \frac{2\kappa_n - 1}{(2\kappa_0) \cdot \cdots \cdot (2\kappa_n)},$$

and conditionally on $(|\xi_1 \wedge \xi_2| = n)$, the random variable $d(\xi_1, \xi_2)$ has the same distribution as

$$\begin{aligned}
&\sum_{p \geqslant n+1} \big(1 + \mathbf{1}\big(I_p^1 \text{ is even}\big)\big) \cdot 2^{-(p+1)} + \big(1 + \mathbf{1}\big(I_n^1 \text{ is even}\big)\big) \cdot 2^{-(n+1)} \\
&\quad + \big(1 + \mathbf{1}\big(I_n^2 \text{ is even}\big)\big) \cdot 2^{-(n+1)} + \sum_{p \geqslant n+1} \big(1 + \mathbf{1}\big(I_p^2 \text{ is even}\big)\big) \cdot 2^{-(p+1)}
\end{aligned}$$

conditioned on $I_n^1 \neq I_n^2$, where $(I_p^1)_{p\in\mathbb{N}}$ and $(I_p^2)_{p\in\mathbb{N}}$ are independent sequences of independent random variables such that for each $p \in \mathbb{N}$, the random variables I_p^1 and I_p^2 have uniform distribution on $[\![1, 2\kappa_p]\!]$. We simplify this into: conditionally on $(|\xi_1 \wedge \xi_2| = n)$, the random variable $d(\xi_1, \xi_2)$ has the same distribution as

$$\begin{aligned}(1+U_1)\cdot 2^{-(n+1)} \\ + \big(2 + \mathbf{1}\big(I_n^1 \text{ is even}\big) + \mathbf{1}\big(I_n^2 \text{ is even}\big)\big)\cdot 2^{-(n+1)} + (1+U_2)\cdot 2^{-(n+1)} \\ = \big(4 + \mathbf{1}\big(I_n^1 \text{ is even}\big) + \mathbf{1}\big(I_n^2 \text{ is even}\big) + U_1 + U_2\big)\cdot 2^{-(n+1)}\end{aligned}$$

conditioned $I_n^1 \neq I_n^2$, where U_1 and U_2 are independent random variables with uniform distribution on $[0, 1]$, independent of I_n^1 and I_n^2. Finally, we check that the last distribution has density ψ_n given by

$$\begin{aligned}\psi_n(x) = \frac{\kappa_n(\kappa_n-1)}{2\kappa_n(2\kappa_n-1)}\cdot 2^{n+1}\cdot\phi\big(2^{n+1}\cdot x - 4\big) \\ + \frac{2\kappa_n^2}{2\kappa_n(2\kappa_n-1)}\cdot 2^{n+1}\cdot\phi\big(2^{n+1}\cdot x - 5\big) \\ + \frac{\kappa_n(\kappa_n-1)}{2\kappa_n(2\kappa_n-1)}\cdot 2^{n+1}\cdot\phi\big(2^{n+1}\cdot x - 6\big),\end{aligned}$$

for all $x \in \mathbb{R}_+$, where ϕ is the probability density function of $U_1 + U_2$. The first term accounts for the case where neither I_n^1 nor I_n^2 is even, the second term for the case where exactly one of them is even, and the last term for the case where both are even. Note that we have

$$\phi(x) = \begin{cases} x & \text{if } x \in [0,1] \\ 2-x & \text{if } x \in [1,2] \\ 0 & \text{otherwise.}\end{cases} \quad \text{for all } x \in \mathbb{R}_+$$

Equation (3.1) readily follows. □

4. Consequences

In this section, we prove Proposition 1.5 and Corollary 1.6. We start with the proof of Proposition 1.5, which relies on the tree constructions of the previous section.

Proof of Proposition 1.5. — Let f be a continuous probability density function on $\mathbb{R}_+$, and assume that there exists $\eta > 0$ such that $f(x) > 0$ for all $x \in]0, \eta]$. First, consider the following lemma, which is very similar to Proposition 1.3.

LEMMA 4.1. — *There exists a pointed compact measured metric space (X, d, π, x_0) such that the following holds, where ξ_1 and ξ_2 are independent random variables with distribution π.*

- *The random variable $d(\xi_1, \xi_2)$ has a continuous probability density function Ψ supported on $[0, 4]$ such that for every $\varepsilon > 0$, we have $\Psi(x) = o(f(\varepsilon x))$ as $x \to 0^+$.*
- *The random variables $d(x_0, \xi_1)$ and $d(x_0, \xi_2)$ have the same distribution as $1 + U$, where U has uniform distribution on $[0, 1]$.*

Proof of the lemma. — Keeping the notation introduced in the proof of Proposition 1.3, recall the construction of the compact measured metric space $\big(\overline{T}, d, \pi\big)$, where the $(\kappa_n)_{n\in\mathbb{N}}$ are to be chosen later. The random variable $d(\xi_1, \xi_2)$ has a continuous probability density function Ψ given by (3.1), hence such that for each $n \in \mathbb{N}$, we have

$$\Psi(x) \leqslant \frac{2\kappa_n - 1}{(2\kappa_0) \cdot \cdots \cdot (2\kappa_n)} \cdot 2^{n+1} \quad \text{for all } x \in \Big[4 \cdot 2^{-(n+1)}, 4 \cdot 2^{-n}\Big].$$

We need to adjust the $(\kappa_n)_{n\in\mathbb{N}}$ so that for each $\varepsilon > 0$, we have $\Psi(x) = o(f(\varepsilon x))$ as $x \to 0^+$. This reduces to simply requiring

$$\Psi(x) = o(g(x)) \text{ as } x \longrightarrow 0^+, \text{ where } g(x) := \min_{[x^2,\eta]} f \tag{4.1}$$

because, for each $\varepsilon > 0$, as $x \mapsto \min_{[x,\eta]} f$ is non-decreasing, so we have

$$g(x) = \min_{[x^2,\eta]} f \leqslant f(\varepsilon x) \quad \text{for all sufficiently small } x.$$

To prove (4.1), let us choose the $(\kappa_n)_{n\in\mathbb{N}}$ so that

$$\kappa_0 \cdot \cdots \cdot \kappa_n \geqslant \frac{2^{n+1}}{g\big(4 \cdot 2^{-(n+2)}\big)} \quad \text{for all sufficiently large } n.$$

This is certainly possible, because it again suffices to be very crude and take

$$\kappa_n \geqslant \frac{2^{n+1}}{g\big(4 \cdot 2^{-(n+2)}\big)} \quad \text{for all sufficiently large } n.$$

This way, for all sufficiently large n, we have, for all $x \in [4 \cdot 2^{-(n+1)}, 4 \cdot 2^{-n}]$,

$$\begin{aligned}
\Psi(x) &\leqslant \frac{2\kappa_n - 1}{(2\kappa_0) \cdot \cdots \cdot (2\kappa_n)} \cdot 2^{n+1} \\
&\leqslant \frac{2\kappa_n}{(2\kappa_0) \cdot \cdots \cdot (2\kappa_n)} \cdot 2^{n+1} \\
&= \frac{2}{\kappa_0 \cdot \cdots \cdot \kappa_{n-1}} \\
&\leqslant 4 \cdot 2^{-(n+1)} \cdot g\Big(4 \cdot 2^{-(n+1)}\Big) \leqslant x \cdot g(x),
\end{aligned}$$

hence $\Psi(x) \leqslant x \cdot g(x)$ for all sufficiently small x. So Lemma 4.1 holds for $(X, d, \pi, x_0) = (\overline{T}, d, \pi, \varnothing)$. □

Now with the (X, d, π, x_0) and Ψ of Lemma 4.1 at hand, we resume the proof of Proposition 1.5. Fix $\beta \in]0, 1/3[$, and fix an integer $n \geqslant 4$. Using the tree construction of Proposition 3.1 together with Lemma 4.1, we will construct a compact measured metric space (S, d, μ) such that the distribution θ achieved by (S, d, μ) has a continuous probability density function g on $\mathbb{R}_+$, with $g \leqslant (1+\zeta(\beta, n)) \cdot f$, where $\zeta(\beta, n) \to 0$ as $\beta \to 0$ and $n \to \infty$. The proof involves a careful choice of a finite discrete distribution θ' approximating f, so that Proposition 3.1 can be applied to construct a space achieving θ'. The definition of θ' will appear at (4.3) but requires some technical preliminaries.

First, fix $0 < \kappa < K$ such that $\int_\kappa^K f(x)\mathrm{d}x \geqslant 1 - \beta$. Then, let $B = \{x \in [\kappa, K] : f(x) \geqslant \beta/K\}$, and note that

$$\int_B f(x)\mathrm{d}x = \int_\kappa^K f(x)\mathrm{d}x - \int_{[\kappa,K]\setminus B} f(x)\mathrm{d}x \geqslant (1-\beta) - K \cdot \frac{\beta}{K} = 1 - 2\beta.$$

Since f is uniformly continuous on $[0, K]$, we can fix $\varepsilon \in]0, \eta \wedge \kappa]$ such that

$$|f(x) - f(y)| \leqslant \frac{\beta^2}{K} \quad \text{for all } x, y \in [0, K] \text{ with } |x - y| \leqslant \varepsilon.$$

Next, we claim that it is possible to fix $p_0 \in]0, 1[$ such that

$$p_0 \cdot \Psi(x) \leqslant \varepsilon/n \cdot f(\varepsilon/n \cdot x) \quad \text{for all } x \in [0, 4], \tag{4.2}$$

where Ψ is the continuous probability density function supported on $[0, 4]$ provided by Lemma 4.1. To prove this claim, since $\Psi(x) = o(f(\varepsilon/n \cdot x))$ as $x \to 0^+$, there exists $x_0 \in]0, 4]$ such that

$$\Psi(x) \leqslant \varepsilon/n \cdot f(\varepsilon/n \cdot x) \quad \text{for all } x \in [0, x_0].$$

Then, since

$$\min_{x \in [x_0, 4]} f(\varepsilon/n \cdot x) = \min_{[\varepsilon x_0/n, 4\varepsilon/n]} f > 0$$

(we use here that $4\varepsilon/n \leqslant \eta$), we can fix $p_0 \in]0, 1[$ such that

$$p_0 \cdot \max_{[0,4]} \Psi \leqslant \varepsilon/n \cdot \min_{[\varepsilon x_0/n, 4\varepsilon/n]} f.$$

Inequality (4.2) readily follows.

Next, let $d_1 < \cdots < d_k \in B$ be at least ε apart from each other, and such that $[d_1, d_1 + \varepsilon] \cup \cdots \cup [d_k, d_k + \varepsilon] \supset B$. Finding such points is elementary: it suffices to take $d_1 = \inf B$, and by induction, for every $i \in \mathbb{N}^*$ such that $d_1, \ldots, d_i$ have been constructed, proceed as follows.

If $[d_1, d_1+\varepsilon] \cup \cdots \cup [d_i, d_i+\varepsilon]$ already covers B, then terminate, otherwise, let

$$d_{i+1} = \inf(B \setminus \big([d_1, d_1+\varepsilon] \cup \cdots \cup [d_i, d_i+\varepsilon]\big) bigr).$$

At the end of the construction, the intervals $[d_1, d_1+\varepsilon[, \ldots, [d_k, d_k+\varepsilon[$ are disjoint and included in $[\kappa, K+\varepsilon[$, so we have $k \cdot \varepsilon \leqslant (K+\varepsilon) - \kappa \leqslant K$ (we use here that $\varepsilon \leqslant \kappa$). Now, by the definition of ε, for each $i \in [\![1, k]\!]$, we have $|f(d_i) - f(x)| \leqslant \beta^2/K$ for all $x \in [d_i, d_i+\varepsilon]$, hence

$$\begin{aligned}\sum_{i=1}^{k} \varepsilon \cdot f(d_i) &\geqslant \sum_{i=1}^{k} \left(\int_{d_i}^{d_i+\varepsilon} f(x)\mathrm{d}x - \varepsilon \cdot \frac{\beta^2}{K} \right) \\ &\geqslant \int_B f(x)\mathrm{d}x - K \cdot \frac{\beta^2}{K} \geqslant (1-2\beta) - \beta^2 \geqslant 1 - 3\beta.\end{aligned}$$

By the mean value theorem, we deduce that it is possible to fix $\alpha_1, \ldots, \alpha_k \in]0, 1/(1-3\beta)[$ such that

$$\alpha_1 \cdot \varepsilon \cdot f(d_1) + \cdots + \alpha_k \cdot \varepsilon \cdot f(d_k) = 1 - p_0,$$

i.e, such that

$$\theta = p_0 \cdot \delta_0 + \alpha_1 \cdot \varepsilon \cdot f(d_1) \cdot \delta_{d_1} + \cdots + \alpha_k \cdot \varepsilon \cdot f(d_k) \cdot \delta_{d_k}$$

is a probability measure. Now discretise further: we let $J_n = [\![0, n^2 - 4n]\!]$, and consider

$$\theta' = p_0 \cdot \delta_0 + \sum_{i=1}^{k} \alpha_i \cdot \varepsilon \cdot f(d_i) \cdot \frac{1}{\#J_n} \sum_{j \in J_n} \delta_{d_i + j \cdot \varepsilon \cdot n^{-2}}. \tag{4.3}$$

By Proposition 3.1, there exists a tree structure (T, ρ, d, ν) that achieves θ'. We use it to construct our compact measured metric space (S, d, μ) as follows. Onto each leaf u of T, we graft a copy (X^u, d^u, π^u, x_0^u) of the pointed compact measured metric space (X, d, π, x_0) provided by Lemma 4.1, scaling distances by ε/n, and identifying the marked point x_0^u with u. This completes the construction, illustrated in Figure 4.1.

We denote by (S, d, μ) the compact measured metric space obtained in this way, where $\mu = \sum_{u \text{ leaf of } T} \nu\{u\} \cdot \pi^u$. Now, we claim that, if ξ_1 and ξ_2 are two independent random variables with distribution μ, then the random

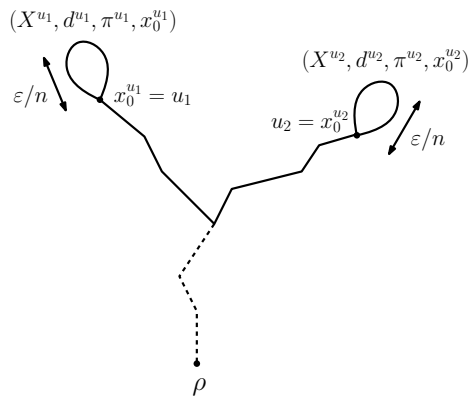

Figure 4.1. A part of the construction of (S, d, μ). Let ξ_1 and ξ_2 be independent random variables with distribution μ, and let $u_1 \neq u_2$ be leaves of T. Conditionally on $\xi_1 \in X^{u_1}$ and $\xi_2 \in X^{u_2}$, the random variable $d(\xi_1, \xi_2)$ has the same distribution as

$$\begin{aligned}\varepsilon/n \cdot (1 + U_1) + d(u_1, u_2) + \varepsilon/n \cdot (1 + U_2) \\ = d(u_1, u_2) + \varepsilon/n \cdot (2 + U_1 + U_2),\end{aligned}$$

where U_1 and U_2 are independent random variables with uniform distribution on $[0, 1]$. Moreover, by construction, we have $d(u_1, u_2) = d_i + j \cdot \varepsilon \cdot n^{-2}$ for some $i \in [\![1, k]\!]$ and $j \in J_n$.

variable $d(\xi_1, \xi_2)$ has probability density function g given by

$$\begin{aligned}g(x) &= p_0 \cdot \frac{\Psi\big(x \cdot (\varepsilon/n)^{-1}\big)}{\varepsilon/n} \\ &\quad + \sum_{i=1}^{k} \alpha_i \cdot \varepsilon \cdot f(d_i) \cdot \frac{1}{\#J_n} \cdot \sum_{j \in J_n} \frac{\phi((x - (d_i + j \cdot \varepsilon \cdot n^{-2})) \cdot (\varepsilon/n)^{-1} - 2)}{\varepsilon/n} \\ &= p_0 \cdot \frac{\Psi\big(x \cdot (\varepsilon/n)^{-1}\big)}{\varepsilon/n} + \sum_{i=1}^{k} \alpha_i \cdot f(d_i) \cdot \frac{n}{\#J_n} \cdot \sum_{j \in J_n} \phi\bigg(\frac{x - d_i}{\varepsilon/n} - \frac{j}{n} - 2\bigg)\end{aligned}$$

for all $x \in \mathbb{R}_+$, where ϕ is the probability density function of the sum of two independent random variables with uniform distribution on $[0, 1]$.

Now, recall that Ψ is supported in $[0, 4]$, hence

$$\Psi\bigg(\frac{x}{\varepsilon/n}\bigg) > 0 \implies x \in [0, 4\varepsilon/n].$$

Moreover, since ϕ is supported in $[0,2]$, we have, for each $i \in [\![1,k]\!]$, by the definition of J_n:

$$\sum_{j\in J_n} \phi\Big(\frac{x-d_i}{\varepsilon/n} - \frac{j}{n} - 2\Big) > 0 \implies x \in [d_i, d_i+\varepsilon].$$

Since $4\varepsilon/n \leqslant \kappa \leqslant d_1$, the interval $[0, 4\varepsilon/n[$ and the intervals $([d_i, d_i+\varepsilon[\,;$ $i \in [\![1,k]\!])$ are mutually disjoint: thus, we have, for all $x \in \mathbb{R}_+$,

$$g(x) = \begin{cases} p_0 \cdot \frac{\Psi\left(x\cdot(\varepsilon/n)^{-1}\right)}{\varepsilon/n} & \text{if } x \in [0, 4\varepsilon/n] \\ \alpha_i \cdot f(d_i) \cdot \frac{n}{\#J_n} \cdot \sum_{j\in J_n} \phi\Big(\frac{x-d_i}{\varepsilon/n} - \frac{j}{n} - 2\Big) & \text{if } x \in [d_i, d_i+\varepsilon] \\ 0 & \text{otherwise.} \end{cases}$$

To conclude the proof, we need to compare g with f over each one of these intervals. First, by (4.2), we have

$$g(x) = p_0 \cdot \frac{\Psi\left(x\cdot(\varepsilon/n)^{-1}\right)}{\varepsilon/n} \leqslant f(x) \quad \text{for all } x \in [0, 4\varepsilon/n].$$

Then, for each $i \in [\![1,k]\!]$, we have

$$\begin{aligned} g(x) &= \alpha_i \cdot f(d_i) \cdot \frac{n}{\#J_n} \cdot \sum_{j\in J_n} \phi\Big(\frac{x-d_i}{\varepsilon/n} - \frac{j}{n} - 2\Big) \\ &\leqslant \frac{1}{1-3\beta} \cdot \frac{f(x)}{1-\beta} \cdot \frac{n}{\#J_n} \cdot \max_{x\in[d_i,d_i+\varepsilon]} \sum_{j\in J_n} \phi\Big(\frac{x-d_i}{\varepsilon/n} - \frac{j}{n} - 2\Big) \end{aligned} \tag{4.4}$$

for all $x \in [d_i, d_i+\varepsilon]$, where the inequality for the second term comes from rearranging

$$|f(d_i) - f(x)| \leqslant \frac{\beta^2}{K} = \beta \cdot \frac{\beta}{K} \leqslant \beta \cdot f(d_i).$$

Finally, let us bound

$$\max_{x\in[d_i,d_i+\varepsilon]} \sum_{j\in J_n} \phi\Big(\frac{x-d_i}{\varepsilon/n} - \frac{j}{n} - 2\Big) = \max_{y\in[0,1]} \sum_{j\in J_n} \phi\Big(ny - \frac{j}{n} - 2\Big).$$

For every $y \in [0,1]$, we have

$$\begin{aligned} \sum_{j\in J_n} \phi\Big(ny - \frac{j}{n} - 2\Big) &= \sum_{j\in J_n} n \cdot \int_{j/n}^{(j+1)/n} \phi\Big(ny - \frac{\lfloor nz\rfloor}{n} - 2\Big)\mathrm{d}z \\ &\leqslant n \cdot \int_{-\infty}^{\infty} \phi\Big(ny - \frac{\lfloor nz\rfloor}{n} - 2\Big)\mathrm{d}z. \end{aligned}$$

Support considerations show that

$$\int_{-\infty}^{\infty} \phi\Big(ny - \frac{\lfloor nz\rfloor}{n} - 2\Big)\mathrm{d}z = \int_{ny-4}^{ny-2+1/n} \phi\Big(ny - \frac{\lfloor nz\rfloor}{n} - 2\Big)\mathrm{d}z.$$

Since $|\lfloor nz\rfloor/n - z| \leqslant 1/n$ for all z, using the 1-Lispchitz continuity of ϕ, we get

$$\int_{ny-4}^{ny-2+1/n} \phi\bigg(ny - \frac{\lfloor nz\rfloor}{n} - 2\bigg)\mathrm{d}z \leqslant \int_{ny-4}^{ny-2+1/n} \bigg(\phi(ny - z - 2) + \frac{1}{n}\bigg)\mathrm{d}z$$
$$= \int_{-1/n}^{2} \phi(t)\mathrm{d}t + \frac{2+1/n}{n} \leqslant 1 + \frac{3}{n}.$$

Thus, we obtain

$$\phi\bigg(ny - \frac{j}{n} - 2\bigg) \leqslant n \cdot \bigg(1 + \frac{3}{n}\bigg) \quad \text{uniformly in } y \in [0,1].$$

Plugging this into (4.4), we end up with

$$g(x) \leqslant \frac{1}{1-3\beta} \cdot \frac{f(x)}{1-\beta} \cdot \frac{n}{\#J_n} \cdot n \cdot \bigg(1 + \frac{3}{n}\bigg)$$
$$= \frac{1}{1-3\beta} \cdot \frac{1}{1-\beta} \cdot \frac{n^2}{n^2-4n+1} \cdot \bigg(1 + \frac{3}{n}\bigg) \cdot f(x)$$

for all $x \in [d_i, d_i + \varepsilon]$. Since this multiplicative constant depends only on β and n, and goes to 1 as $\beta \to 0$ and $n \to \infty$, the proof of Proposition 1.5 is complete. □

4.1. Random metric spaces

Finally, we show how Proposition 1.5 implies that every distribution with a suitable density is achieved by some *random* compact measured metric space (Corollary 1.6). Although we will not need sophisticated Gromov–Hausdorff–Prokhorov theory, let us take a brief paragraph to set the scene more rigorously. Let $\mathbb{M}$ be the space of compact measured metric spaces (S, d, μ) (i.e, compact metric spaces (S, d) endowed with a Borel probability measure μ), seen up to measure-preserving isometry. We denote by $[S, d, \mu]$ the equivalence class of (S, d, μ). The Gromov–Hausdorff–Prokhorov metric $D : \mathbb{M} \times \mathbb{M} \to \mathbb{R}_+$ is defined by

$$D([S_1, d_1, \mu_1], [S_2, d_2, \mu_2]) = \inf_{\substack{\phi_1 : S_1 \to S \\ \phi_2 : S_2 \to S}} d_{\mathrm{H}}(\phi_1(S_1), \phi_2(S_2)) \vee d_{\mathrm{P}}(\phi_1^*\mu_1, \phi_2^*\mu_2)$$

for all $[S_1, d_1, \mu_1], [S_2, d_2, \mu_2] \in \mathbb{M}$, where the infimum is over all isometric embeddings $\phi_1 : S_1 \to S$ and $\phi_2 : S_2 \to S$ into some common metric space (S, d), and where d_{H} denotes the Hausdorff distance between non-empty compact subsets of S, and d_{P} denotes the Prokhorov distance between Borel probability measures on S. This makes $(\mathbb{M}, D)$ into a separable and complete

metric space, that we equip with its Borel σ-algebra. Then, the fact that the mapping $\kappa : \mathbb{M} \times \mathcal{B}(\mathbb{R}_+) \to [0,1]$ defined by

$$\kappa([S,d,\mu],B) = \int_S \int_S \mathbf{1}(d(s_1,s_2) \in B) \mathrm{d}\mu(s_2) \mathrm{d}\mu(s_1) \qquad \text{for all } [S,d,\mu] \in \mathbb{M} \text{ and } B \in \mathcal{B}(\mathbb{R}_+)$$

is a Markov kernel allows to give a rigorous definition to sampling a random compact measured metric space $[S,d,\mu]$, and then conditionally on $[S,d,\mu]$, sampling a random variable distributed as $d(\xi_1,\xi_2)$, where ξ_1 and ξ_2 are independent random variables with distribution μ conditionally on $[S,d,\mu]$.

Without further ado, let us present the proof of Corollary 1.6.

Proof of Corollary 1.6. — The essential idea is that, using Proposition 1.5 iteratively, we can write

$$f = \sum_{n \geqslant 1} 2^{-n} \cdot g_n \quad \text{almost everywhere,} \tag{4.5}$$

where for each $n \in \mathbb{N}^*$, the function $g_n : \mathbb{R}_+ \to \mathbb{R}_+$ is the probability density function of a distribution which is achieved by some compact measured metric space (S_n, d_n, μ_n). Then, to achieve the distribution with density f, it suffices to take a mixture of the (S_n, d_n, μ_n), with weights 2^{-n}.

To complete the proof, we need to justify (4.5). Let $n \in \mathbb{N}$, and assume by induction that there are continuous probability density functions $g_1, \ldots, g_n$ and f_n on $\mathbb{R}_+$ such that

$$f = \sum_{k=1}^{n} 2^{-k} \cdot g_k + 2^{-n} \cdot f_n,$$

where for each $k \in [\![1,n]\!]$, the distribution with density g_k is achieved by some compact measured metric space (S_k, d_k, μ_k), and where there exists $\eta > 0$ such that $f_n(x) > 0$ for all $x \in]0,\eta]$. (Initially, consider the trivial decomposition $f = f_0$.) Then, by Proposition 1.5, there exists a compact measured metric space $(S_{n+1}, d_{n+1}, \mu_{n+1})$ such that the distribution θ achieved by $(S_{n+1}, d_{n+1}, \mu_{n+1})$ has a continuous density g_{n+1} on $\mathbb{R}_+$, where $g_{n+1} \leqslant 3/2 \cdot f_n$. Next, we can write

$$f_n = \frac{1}{2} \cdot g_{n+1} + \frac{1}{2} \cdot f_{n+1},$$

where

$$f_{n+1} = \frac{f_n - g_{n+1}/2}{1 - 1/2}.$$

By construction, we have

$$f = \sum_{k=1}^{n+1} 2^{-k} \cdot g_k + 2^{-(n+1)} \cdot f_{n+1},$$

and since

$$f_{n+1} \geqslant \frac{f_n/4}{1-1/2} = \frac{f_n}{2},$$

we have $f_{n+1}(x) > 0$ for all $x \in]0, \eta]$. By induction, we obtain in this way a sequence of continuous probability density functions $g_1, g_2, \dots$ on $\mathbb{R}_+$ such that for each $k \in \mathbb{N}^*$, the distribution with density g_k is achieved by some compact measured metric space (S_k, d_k, μ_k), and where

$$f \geqslant \sum_{k \geqslant 1} 2^{-k} \cdot g_k.$$

As both terms integrate to 1, they must be equal almost everywhere, which yields (4.5). □

5. Homometric structures.

If (S, d) is a finite metric space, the distance multiset of A is the multiset of pairwise distances between points of S. In metric geometry, two metric spaces having the same distance multisets are called *homometric*. When (S, d, μ) is no longer finite but instead a measured metric space, one can instead of the multiset of pairwise distances use the distribution of the distance between 2 independent points sampled according to μ. We shall by extension say that (S, d, μ) and (S', d', μ') are homometric if $d(\xi_1, \xi_2) = d'(\xi_1', \xi_2')$ in distribution where $\xi_1, \xi_2 \in S$ and $\xi_1', \xi_2' \in S'$ are independent and of distribution μ and μ' respectively. The study of homometric structures has significant applications, notably in crystallography [11], molecular biology—particularly in restriction site analysis of DNA [12], and algorithmic problems such as the turnpike problem [6]. Of course, homometric structures may not be isometric. A striking result in this direction is the recent extension [3] of the famous hexachord theorem, which was popularized in musicology by composers Milton Babbitt and David Lewin in the 1950s. This result states the following: *Suppose (S, d, μ) is a measured metric space where μ is preserved by a transitive group of isometries. If $A \subset S$ satisfies $\mu(A) = 1/2$, then A and its complement A^c are homometric.*

Acknowledgements

The counter-example in Proposition 2.2 was outlined by "Mike" in response to our question posted on MathOverflow [1], and after comments by Iosif Pinelis, the present version was given on MathOverflow by Yuval Peres.

Bibliography

[1] D. J. Aldous, "Distributions of distance between two random points in Hilbert space", MathOverflow, `https://mathoverflow.net/q/428539` (version: 2022-08-15).

[2] ———, "Representations for partially exchangeable arrays of random variables", *J. Multivariate Anal.* **11** (1981), no. 4, p. 581-598.

[3] M. Andreatta, C. Guichaoua & N. Juillet, "New hexachordal theorems in metric spaces with a probability measure", *Rend. Semin. Mat. Univ. Padova* **152** (2024), p. 45-58.

[4] T. Austin, "Exchangeable random arrays", 2013, Lecture notes, available at `https://www.stat.berkeley.edu/~aldous/206-Exch/Papers/austin_arrays.pdf`.

[5] B. Clark, "How Many Songs are There in the World? (2025)", Musicianwave, `https://www.musicianwave.com/how-many-songs-are-there-in-the-world` (version: 2023-07-18).

[6] T. Dakic, "On the Turnpike Problem", PhD Thesis, Simon Fraser University, BC, Canada, 2000.

[7] M. Gaczkowski & P. Górka, "Harmonic functions on metric measure spaces: convergence and compactness", *Potential Anal.* **31** (2009), no. 3, p. 203-214.

[8] M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Progress in Mathematics, vol. 152, Birkhäuser, 1999, based on the 1981 french original, with appendices by M. Katz, P. Pansu and S. Semmes, translated from the french by Sean M. Bates.

[9] D. N. Hoover, "Row-column exchangeability and a generalized model for probability", in *Exchangeability in probability and statistics (Rome, 1981)*, North-Holland, 1982, p. 281-291.

[10] O. Kallenberg, *Probabilistic symmetries and invariance principles*, Probability and Its Applications, Springer, 2005.

[11] A. L. Patterson, "Homometric structures", *Nature* **143** (1939), no. 3631, p. 939-940.

[12] S. S. Skiena & G. Sundaram, "A partial digest approach to restriction site mapping", *Bull. Math. Biol.* **56** (1994), p. 275-294.

[13] A. M. Vershik, "The universal Uryson space, Gromov's metric triples, and random metrics on the series of natural numbers", *Usp. Mat. Nauk* **53** (1998), no. 5, p. 57-64.

[14] ———, "Random and universal metric spaces", in *Fundamental mathematics today (Russian)*, Nezavis. Mosk. Univ., 2003, p. 54-88.