Annales Henri Lebesgue
8 (2025) 1023-1035

ANNALES
HENRI LEBESGUE

RENAN GROSS

THE DISTANCE PROBLEM VIA SUBADDITIVITY

LE PROBLÈME DE LA DISTANCE VIA LA SOUS-ADDITIVITÉ

Abstract. — In a recent paper, Aldous, Blanc and Curien asked which distributions can be expressed as the distance between two independent random variables on some separable measured metric space. We show that every nonnegative discrete distribution whose support contains 0 arises in this way, as well as a class of compactly supported distributions with density.

Résumé. — Dans un article récent, Aldous, Blanc et Curien ont demandé quelles distributions peuvent s'écrire comme la distance entre deux variables aléatoires indépendantes sur un espace métrique mesuré séparable. Nous montrons que toute distribution discrète positive ou nulle dont le support contient 0 apparaît de cette manière, ainsi qu'une classe de distributions à densité à support compact.

1. Introduction

1.1. The problem

The *distance problem* asks the following. Given a distribution θ on $[0,\infty)$, is it possible to find a complete separable metric space (S,d) and a Borel probability measure μ on S, so that when $X,Y\sim\mu$ are independent, their distance satisfies $d(X,Y)\sim\theta$? If so, we say that (S,d,μ) *achieves* θ, and that θ is *feasible*. This very natural and quite inviting problem was first proposed by Aldous, Blanc and Curien in [ABC24], where they also give some statistical motivation.

Keywords: Measured metric space, trees, subadditivity, metric preserving function.
2020 *Mathematics Subject Classification:* 28C99, 51F99, 51K99, 60D05.
DOI: `https://doi.org/10.5802/ahl.254`

1.2. Examples

While any (S, d, μ) gives some distribution θ which can in principle be calculated, the inverse problem, of finding an explicit space which achieves a given distribution, is harder. The following examples show how to achieve some common distributions, and may give some intuition about the types of constructions one can use.

(1) *The uniform measure:* let $S = S^1 = \mathbb{R} \mod \mathbb{Z}$ be the unit circle, μ the uniform measure on S^1, and $d(x, y)$ the arc length between x and y. Then $d(X, Y) \sim U[0, \frac{1}{2}]$.

(2) *The exponential distribution:* let $S = \mathbb{R}$, let $d(x, y) = |x - y|$, and let $\mu = (1)$ be the exponential distribution. By the memoryless property, $|X - Y|$ is also exponentially distributed.

(3) *Bernoulli distribution:* let $0 < p_0 < 1$. In order to achieve the Bernoulli distribution $\theta = p_0\delta_0 + (1 - p_0)\delta_1$, let $m \geqslant 2$ be an integer and let $\alpha \in [0, 1]$ to be chosen later. Let $S = \{0, \ldots, m - 1\}$, equipped with the discrete metric

$$d(x, y) = \begin{cases} 0 & x = y \\ 1 & x \neq y. \end{cases}$$

Let μ be the distribution of the random variable on S whose value is 0 with probability α, and uniform among $\{1, \ldots, m - 1\}$ with probability $1 - \alpha$. Then

$$\mathbb{P}[d(X, Y) = 0] = \alpha^2 + \left(\frac{1 - \alpha}{m - 1}\right)^2. \tag{1.1}$$

Choosing m so that the quadratic equation $p_0 = \alpha^2 + (\frac{1-\alpha}{m-1})^2$ has a solution and solving for α gives us the desired probability space.

(4) *The uniform measure, revisited:* let $S = \{0, 1\}^{\mathbb{N}}$, let μ be the measure on S where all entries are i.i.d. 0-1 Bernoulli random variables with success probability $1/2$, and let $d(x, y) = \sum_{k=1}^{\infty} 2^{-k}|x_i - y_i|$. Then $d(X, Y) \sim U[0, 1]$. This is essentially a more explicit version of the first example.

(5) *The exponential distribution, revisited:* let $S = (0, 1]$, μ the uniform measure on S, and $d(x, y) = |\int_x^y \frac{1}{t} dt| = |\log \frac{y}{x}|$. A short calculation shows that $\mathbb{P}[d(X, Y) \leqslant t] = 1 - e^{-t}$, and so $d(X, Y) \sim \mathrm{Exp}(1)$.

(6) *Non-example:* let θ be any distribution with density on $[0, \infty)$, and denote its cumulative distribution function (CDF) by F. Let $S = [0, \infty)$, let

$$d(x, y) = \begin{cases} 0 & x = y \\ \max(x, y) & x \neq y, \end{cases}$$

and let μ be the distribution whose CDF is $\sqrt{F}$. Since θ has density, the event $\{x = y\}$ has measure 0 under μ, and so $d(X, Y) = \max(X, Y)$ with probability 1. The CDF of the maximum of two i.i.d. random variables is the square of their individual CDF, and so $d(X, Y) \sim \theta$. However, while μ is a Borel measure with respect to the Euclidean topology, it is not a Borel measure with respect to the topology induced by d. Note also that (S, d) is not separable (for small enough ε, the only element ε-close to x is x itself).

1.3. Results

In [ABC24], Aldous, Blanc and Curien showed some basic constraints on the support of the distribution (for example, it must contain 0) and gave several constructions and approximations. In their main constructions, the metric space S was a weighted tree and the measure μ a distribution on its leaves/rays. Choosing the weights and distribution in a clever way, they showed that every finite discrete distribution with 0 in its support is feasible. However, when applied to non-finite distributions via finite approximations, their approach yields non-separable spaces.

Our first result is a new construction which does not involve taking limits of finite distributions, avoiding limiting questions of separability. This result resolves the question of discrete distributions.

Theorem 1.1. — *Every nonnegative discrete distribution θ supported on 0 is feasible. (proof ↗)*

Rather than building a metric space (S, d, μ) which directly achieves θ from scratch, our proof combines together several simpler spaces (S_n, d_n, μ_n), allowing us to sample from different θ_ns with different probabilities. The main challenge is to decompose θ into subdistributions so that combining the spaces preserves the triangle inequality. As will be described in the next section, this construction can also be interpreted as a metric on the rays of a tree, where two rays are close if they travel together for a long time, generalizing a commonly used metric on the space of ends of a tree.

Our second result concerns distributions with density.

Theorem 1.2. — *Let g be a probability density function on $[0, 1]$, and suppose that there exist constants $c, C > 0$ such that*

$$c < g(t) < C \tag{1.2}$$

for all $t \in [0, 1]$. Then the distribution θ with density g is feasible. (proof ↗)

The theorem is derived by altering the metric of a known-to-be-feasible distribution; here, subadditivity is used in order to preserve the metric. The theorem represents one step towards answering the open problem in [ABC24].

Problem 1.3. — *Prove that for every probability density function f on $[0, \infty)$ whose support contains 0, the distribution with density f is feasible.*

Acknowledgments

We thank Wendelin Werner and Guillaume Blanc for useful comments, and the anonymous referee for their corrections and suggestions.

2. Proof of Theorem 1.1

Our proof uses the following construction, which, given several feasible distributions, returns a mixture of these distributions. The resultant distance can be thought of as being computed by a "selection" mechanism: we linearly order the distributions, and iteratively flip a coin for each one to see whether or not it is the one that we sample from; if none of the distributions are chosen, the default 0 is returned.

2.1. Selection space

Denote the integers by $\mathbb{Z}$ and the positive integers by $\mathbb{N} = \{1, 2, \ldots\}$. We say that a triplet (S, d, μ) is a measured metric space if (S, d) is a metric space and μ is a Borel probability measure on S with respect to d.

Let $I \subseteq \mathbb{Z}$, and for each $n \in I$, let (S_n, d_n, μ_n) be a complete countable measured metric space with more than 1 point. Assume that for every $n \in I \cap \mathbb{N}$ there exists a special $s_n \in S_n$ such that

$$\sum_{n \in I \cap \mathbb{N}} 1 - \mu_n(\{s_n\}) < \infty. \tag{2.1}$$

We now define a new space (S, d, μ) out of these spaces; later, in Proposition 2.2, we give conditions under which it is a complete separable measured metric space.

(1) *The space*: we say that a vector $x \in \prod_{n \in I} S_n$ is *eventually constant* if there exists an $N > 0$ such that $x_n = s_n$ for all $n \geqslant N$ (we still index the vector x by the set I). Note that when $\sup I < \infty$, every vector is eventually constant. We set

$$S = \left\{ x \in \prod_{n \in I} S_n \,\middle|\, x \text{ is eventually constant} \right\}.$$

(2) *The function d*: for $x \neq y \in S$, let $n(x, y) = \sup\{n \mid x_n \neq y_n\}$. This is always finite since x and y are eventually constant. We then define

$$d(x, y) = \begin{cases} 0 & x = y \\ d_{n(x,y)}\Big(x_{n(x,y)}, y_{n(x,y)}\Big) & x \neq y. \end{cases}$$

(3) *The measure*: let $\widetilde{\mu} = \otimes_{n \in I} \mu_n$ be the product measure on $\prod_{n \in I} S_n$. By (2.1) the sum $\sum_{n \in I \cap \mathbb{N}} 1 - \mu_n(\{s_n\})$ is finite, and the Borel–Cantelli lemma implies that $X \sim \widetilde{\mu}$ is eventually constant with probability 1 under $\widetilde{\mu}$. The measure $\widetilde{\mu}$ can therefore be restricted to a measure μ on S.

We call the tuple (S, d, μ) the *selection space* of $\{(S_n, d_n, \mu_n)\}_{n \in I}$. Let us find the distribution of $d(X, Y)$ when $X, Y \sim \mu$ are independent. The probability of obtaining 0 is

$$\mathbb{P}[d(X, Y) = 0] = \prod_{n \in I} \mathbb{P}[X_n = Y_n] = \prod_{n \in I} \mathbb{P}[d_n(X_n, Y_n) = 0], \tag{2.2}$$

while for $r > 0$,

$$\begin{aligned}(2.3)\quad \mathbb{P}[d(X,Y)=r] &= \sum_{n\in I}\mathbb{P}[n(X,Y)=n]\cdot\mathbb{P}[d_n(X_n,Y_n)=r \mid n(X,Y)=n] \\ &= \sum_{n\in I}\left(\prod_{k\in I,\,k>n}\mathbb{P}[X_k=Y_k]\right)\mathbb{P}[X_n\neq Y_n]\cdot\mathbb{P}[d_n(X_n,Y_n)=r \mid X_n\neq Y_n] \\ &= \sum_{n\in I}\left(\prod_{k\in I,\,k>n}\mathbb{P}[d(X_k,Y_k)=0]\right)\mathbb{P}[d_n(X_n,Y_n)=r],\end{aligned}$$

and we have indeed obtained a mixture of positive elements of the distributions achieved by (S_n, d_n, μ_n).

Remark 2.1. — As a space of vectors indexed by $I \subseteq \mathbb{Z}$, the selection space can also be interpreted as the steps of a bi-directional walk, where at time-step n, the walker may go in the "direction" $x_n \in S_n$. When $\sup I$ is finite and we view time as starting at $\sup I$ and directed down towards $\inf I$, the index $n(x,y)$ is the first time that two walkers x and y split up when starting at a common origin. In this case the elements of S can be seen as the space of ends on some rooted tree, and the function d is a generalization of the metric $e^{-h(x,y)}$ on the space of ends, where $h(x,y)$ is the depth of the last common vertex of the rays x and y. See Figure 2.1. When $\sup I = \infty$, however, there is no such common root.

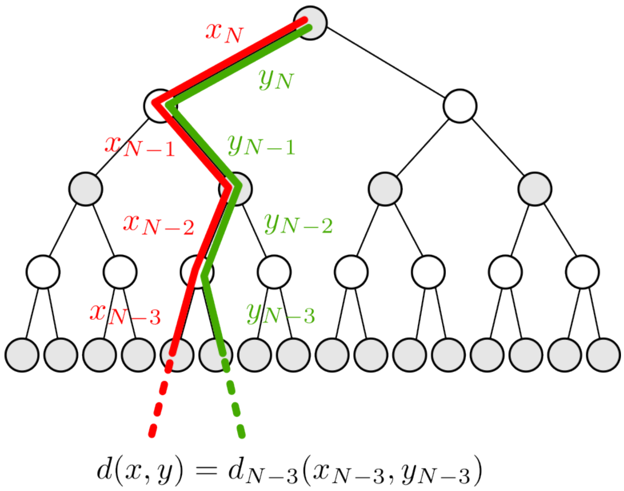

Figure 2.1. When $\sup I = N < \infty$, the function d is a generalization of a class of metrics on the space of ends of a tree rooted, where the distance between two rays from the origin depends on the depth of their last common vertex.

PROPOSITION 2.2. — *Let $I \subseteq \mathbb{Z}$ and $\{(S_n, d_n, \mu_n)\}_{n\in I}$ be as above. Assume that the following two properties hold.*

(1) *For every $m < n \in I$, every $x_1 \neq x_2 \in S_m$ and every $y_1 \neq y_2 \in S_n$,*

$$d_m(x_1,x_2) \leqslant 2d_n(y_1,y_2). \tag{2.4}$$

(2) *We have*

$$\text{either } \inf I > -\infty \quad \text{or} \quad \lim_{n\to-\infty} \sup_{z,w\in S_n} d_n(z,w) = 0. \tag{2.5}$$

Then the selection space (S,d,μ) is a complete separable measured metric space.

Proof.

Metric. — It is clear that $d(x,y)=0$ if and only if $x=y$, so we only need to verify the triangle inequality. Let $x,y,z\in S$ be distinct, and denote $b=n(x,y)$.

(1) If either $n(y,z)>b$ or $n(x,z)>b$, then since x and y agree on indices greater than b, we must have $n(x,z)=n(y,z)>b$, and also $d(x,z)=d(y,z)$. Then by (2.4),

$$d(x,y)=d_b(x_b,y_b)\leqslant 2d_{n(x,z)}\left(x_{n(x,z)},z_{n(x,z)}\right)=d(x,z)+d(y,z).$$

(2) If both $n(y,z)=b$ and $n(x,z)=b$, then the triangle inequality is satisfied because d_b is a metric.

(3) Otherwise, we have $n(y,z),n(x,z)\leqslant b$, and at least one of $n(y,z)$ or $n(x,z)$ is strictly smaller than b. But it is impossible for both to be strictly smaller, since that would mean that $x_b=z_b=y_b$, contradicting the fact that $x_b\neq y_b$. Supposing w.l.o.g. that $n(x,z)=b$, we then trivially have

$$d(x,y)=d_b(x_b,y_b)=d_b(x_b,z_b)=d(x,z)\leqslant d(x,z)+d(y,z).$$

Completeness. — Let $N\in I$, and let $a\neq b\in S_N$. Then by (2.4), for every $n>N$ and every $w,z\in S_n$, $d_n(w,z)\geqslant \frac{1}{2}d_N(a,b)$. Thus, for every $N\in I$,

$$\inf_{n>N}\inf_{w,z\in S_n} d_n(w,z)>0. \tag{2.6}$$

Let $(x^{(k)})_{k\in\mathbb{N}}$ be a Cauchy sequence in S, and let $L=\inf I$. When $L=-\infty$, (2.6) implies that if $d(x^{(k_1)},x^{(k_2)})<\varepsilon$ for some k_1 and all $k_2\geqslant k_1$ and small enough ε, then $x_n^{(k_2)}=x_n^{(k_1)}$ for all $n>N$ for some $N\in I$, with $N\to-\infty$ as $\varepsilon\to 0$. Thus $x^{(k)}$ converges to an element in S. When $L>-\infty$, (2.6) implies that for small enough ε, $x_n^{(k_2)}=x_n^{(k_1)}$ for all $n>L$, while the first elements $(x_L^{(k)})_{k\in\mathbb{N}}$ form a Cauchy sequence in S_L. Since S_L is itself complete, we again get that $x^{(k)}$ converges to an element in S.

Separability. — If $\inf I>-\infty$, then S is a countable union of countable sets, and so is countable itself. Otherwise, by (2.5), $\lim_{n\to-\infty}\sup_{z,w\in S_n} d(z,w)=0$. For each negative $n\in I$, let $a^{(n)}\in S_n$, and set

$$A_n=\left\{x\in S\,\middle|\, x_i=a^{(i)}\ \forall\, i\leqslant n\right\}.$$

Since all vectors in S are eventually constant, each A_n is a countable union of countable sets and is therefore countable, and so the union $\cup_{n\in I}A$ is also countable. It is also dense in S: for any $x\in S$ and any $n\in I$, the set A_n contains an element y such that $y_k=x_k$ for all $k>n$, and so $n(x,yt)\leqslant n$; denseness follows since $\sup_{z,w\in S_n} d_n(z,w)\to 0$ as $n\to-\infty$.

Borel measure. — Under the metric d, an open ball around $x \in S$ of radius r is given by

$$B(x,r) = \bigcup_{N \in I} \bigcup_{s \in S_N \mid d_N(x_N,s)<r, s\neq x_N} \{y \in S \mid y_N = s, y_n = x_n \text{ for all } n > N\} \cup \{x\}.$$

Each set of the form $\{y \in S \mid y_N = s, y_n = x_n \text{ for all } n > N\}$ is measurable under the product measure $\widetilde{\mu}$, and therefore so is the countable union of such sets. The restriction measure μ is therefore defined on the σ-algebra generated by the open balls in S. Since (S,d) is separable, μ is Borel. ☐

Proposition 2.3. — *Let $a > 0$, let $p_0 \in (0,1)$ and let λ be a discrete probability distribution with* $\mathrm{supp}(\lambda) \subseteq [a, 2a]$. *Then the distribution $\theta = p_0\delta_0 + (1-p_0)\lambda$ is feasible.*

Proof. — Let $a_0 = 0$, and let $(a_n)_{n\in I}$ be the positive atoms of λ arranged in some arbitrary order; here, $I = \mathbb{N}$ if there is an infinite number of atoms and $I = [N]$ for some N otherwise. For every $n \in I$, denote $p_n = \theta(\{a_n\})$, and define q_n by

$$q_n = \frac{p_n}{p_0 + p_1 + \ldots + p_n}.$$

For $n \in I$, let (S_n, d_n, μ_n) be the metric space corresponding to the Bernoulli measure $\theta_n = (1-q_n)\delta_0 + q_n\delta_{a_n}$, obtained by multiplying the metric described in the examples section by a_n. Then the assumption in the construction of the selection space and the two conditions in Proposition 2.2 hold for $\{(S_n, d_n, \mu_n)\}_{n\in I}$:

(1) When constructing a Bernoulli random variable $p\delta_0 + (1-p)\delta_1$with probability p of obtaining 0 as in Example 3, the parameters m and α may be chosen so that $\alpha \geqslant p$. Indeed, setting $f_m(x) = x^2 + (\frac{1-x}{m-1})^2$, we have $p = f_m(\alpha)$. Since $\lim_{m\to\infty} f_m(p) = p^2 < p$, there exists an m large enough so that $f_m(p) < p$, and by continuity of f_m and the fact that $f_m(1) = 1$, there exists $\alpha \in [p, 1]$ with $f_m(\alpha) = p$ as required. This shows that the probability to obtain 0 under μ_n is bounded by

$$\mu_n(\{0\}) \geqslant 1 - q_n. \tag{2.7}$$

Thus

$$\sum_{n\in I} 1 - \mu_n(\{0\}) \leqslant \sum_{n\in I} q_n = \sum_{n\in I} \frac{p_n}{p_0 + p_1 + \ldots + p_n} \leqslant \sum_{n\in I} \frac{p_n}{p_0} = \frac{1-p_0}{p_0}, \tag{2.8}$$

and the sum in (2.1) is finite.

(2) Since $a_n \in [a, 2a]$, the requirement (2.4) is immediate: the left-hand side of (2.4) is at most $2a$, while the right-hand side is at least $2a$.

(3) $\inf I = 1 > -\infty$ as required in (2.5).

Thus, the selection space (S, d, μ) constructed from $\{(S_n, d_n, \mu_n)\}_{n\in I}$ is a complete separable measured metric space. Since the positive atoms of the μ_ns are all distinct, equations (2.2) and (2.3) imply that

$$\mathbb{P}[d(X,Y) = 0] = \prod_{n\in I} \mathbb{P}[d_n(X_n, Y_n) = 0] = \prod_{n\in I} (1 - q_n) = p_0, \tag{2.9}$$

and

$$\mathbb{P}[d(X,Y) = a_n] = q_n \prod_{i \in I, i > n} (1 - q_i) = p_n$$

as needed.

We remark that by (2.7) and (2.9),

$$\mathbb{P}[X = (0,0,\ldots)] = \prod_{n \in I} \mu_n(\{0\}) \geqslant \prod_{n \in I} (1 - q_n) = p_0. \tag{2.10}$$ □

Proof of Theorem 1.1. — The measure θ can be written as

$$\theta = p_\infty \delta_0 + \sum_{n \in I} p_n \lambda_n,$$

where $I \subseteq \mathbb{Z}$, and for every $n \in I$, $p_n > 0$ and the measure λ_n is discrete and supported on $(2^n, 2^{n+1}]$. Note that if $p_\infty = 0$, then necessarily $\inf I = -\infty$. Define the numbers

$$\beta_n = \frac{p_n}{1 - \sum_{i \in I, i > n} p_i},$$

and let (S_n, d_n, μ_n) be the metric space defined in the proof of Proposition 2.3 corresponding to the measure $\theta_n = (1 - \beta_n)\delta_0 + \beta_n \lambda_n$. Note that each S_n is countable. Then the assumption in the construction of the selection space and the two conditions in Proposition 2.2 hold for $\{(S_n, d_n, \mu_n)\}_{n \in I}$:

(1) If $\sup I < \infty$ then of course the sum in (2.1) is finite. Otherwise, by (2.10), if $X_n \sim \mu_n$, then

$$\mathbb{P}[X_n = (0,0,\ldots)] \geqslant 1 - \beta_n.$$

Thus

$$\begin{aligned}\sum_{n \in I \cap \mathbb{N}} \mathbb{P}[X_n \neq (0,0,\ldots)] &\leqslant \sum_{n \in I \cap \mathbb{N}} 1 - (1 - \beta_n) = \sum_{n \in I \cap \mathbb{N}} \beta_n \\ &\leqslant \sum_{n \in I \cap \mathbb{N}} \frac{p_n}{1 - \sum_{i \in I, i > n} p_i} \leqslant \frac{1}{p_\infty + \sum_{i \in I, i < 1} p_i}.\end{aligned}$$

(2) Since the range of d_n is $(2^n, 2^{n+1}]$, the requirement (2.4) is immediate; in fact, the left hand side of (2.4) is bounded by just half of the right hand side.

(3) $\lim_{n \to -\infty} \sup_{z,w \in S_n} d_n(z,w) = \lim_{n \to -\infty} 2^{n+1} = 0$ as required in (2.5).

Thus, the selection space (S, d, μ) constructed from $\{(S_n, d_n, \mu_n)\}_{n \in I}$ is a complete separable measured metric space. Since the supports of λ_n are disjoint, equations (2.2) and (2.3) imply that

$$\mathbb{P}[d(X,Y) = 0] = \prod_{n \in I} \mathbb{P}[X_n = Y_n] = \prod_{n \in I} (1 - \beta_n) = p_\infty,$$

while the probability to sample from λ_n is exactly $\beta_n \prod_{i \in I, i > n} (1 - \beta_i) = p_n$, as needed. □

3. Proof of Theorem 1.2

Our proof first finds an easy-to-construct feasible "starter" distribution, and then changes its metric in order to obtain the distribution θ. The idea is as follows. Let W be a random variable on $[0,1]$ whose cumulative distribution function F_W is continuous and strictly increasing on $[0,1]$, and let $\varphi : [0,1] \to [0,1]$ be a strictly increasing bijection. The cumulative distribution function of the random variable $\varphi(W)$ is given by

$$F_{\varphi(W)}(t) = \mathbb{P}[\varphi(W) \leqslant t] = \mathbb{P}\Big[W \leqslant \varphi^{-1}(t)\Big] = F_W\Big(\varphi^{-1}(t)\Big).$$

Denoting by G the cumulative distribution function of θ and choosing $\varphi(t) = G^{-1}(F_W(t))$ gives us, for $t \in [0,1]$, $F_W(\varphi^{-1}(t)) = G(t)$, and so the random variable $\varphi(W)$ has distribution θ. If W is achievable by some (S, d, μ), then when $X, Y \sim \mu$ are independent, we have $\varphi(d(X,Y)) \sim \theta$. However, this does not immediately mean that $(S, \varphi \circ d, \mu)$ is a measured metric space that achieves θ, since there is no guarantee that $\varphi \circ d$ is still a metric (or that μ is still Borel under the topology induced by $\varphi \circ d$). The essence of the proof is to find a suitable W which allows φ to preserve the metric d.

Definition 3.1. — *A function $f : [0,\infty) \to [0,\infty)$ is called* metric preserving *if for every metric space (S,d), the function $f \circ d$ is also a metric on S. It is called* strongly metric preserving *if it is metric preserving and (S,d) is topologically equivalent to $(S, f \circ d)$.*

Metric preserving functions have been well studied; see [Cor99] for a brief introduction. We will require only the following standard result.

Theorem 3.2 ([Kel75, p. 131]). — *Let $f : [0,\infty) \to [0,\infty)$ be continuous, nondecreasing, and satisfying the following two conditions:*

(1) $f(x) = 0 \iff x = 0$.
(2) *(Subadditivity)* $f(x+y) \leqslant f(x) + f(y)$ *for all* $x, y \in [0,\infty)$.

Then f is strongly metric preserving.

Theorem 3.2 implies that it suffices to find a continuous random variable W such that $G^{-1}(F_W(t))$ is subadditive. If G^{-1} is itself already subadditive, then no modification by F_W is actually needed; in this case, we can take W to be the uniform measure on $[0,1]$, so that $F_W(t) = t$ for $t \in [0,1]$ and $G^{-1}(F_W(t)) = G^{-1}(t)$. Obtaining θ is then just an instance of the inverse sampling theorem (as we saw in the examples section, the uniform measure is indeed feasible). However, when G^{-1} is not subadditive, the idea is that if F_W itself is in some sense very strongly subadditive, then its subadditivity is enough to counter the non-subadditivity of G^{-1}. This is made precise in the following simple proposition.

Proposition 3.3. — *Let $\Psi : [0,1] \to [0,1]$ be an absolutely continuous nondecreasing function, and assume that there exist $m, M > 0$ such that*

$$m < \Psi'(t) < M. \tag{3.1}$$

Let $f : [0,\infty) \to [0,1]$ be such that for every $x, y \in [0,\infty)$ with $x \leqslant y$,

$$f(x+y) \leqslant \frac{m}{M} f(x) + f(y). \tag{3.2}$$

Then $\Psi \circ f$ is subadditive.

Proof. — By absolute continuity, for all $x \leqslant y$,

$$\begin{aligned}\Psi(f(x+y)) &= \Psi(f(y)) + \int_{f(y)}^{f(x+y)} \Psi'(t)\,dt \\ &\overset{(3.1)}{\leqslant} \Psi(f(y)) + M(f(x+y) - f(y)) \\ &\overset{(3.2)}{\leqslant} \Psi(f(y)) + mf(x) \\ &\overset{(3.1)}{\leqslant} \Psi(f(y)) + \Psi(f(x)).\end{aligned}$$

□

The next proposition shows that it is indeed possible to obtain such f.

Proposition 3.4. — *For every $\varepsilon \in (0,1)$, there exists a feasible continuous random variable W whose cumulative distribution function F_W satisfies*

$$F_W(x+y) \leqslant \varepsilon F_W(x) + F_W(y)$$

for all $x \leqslant y$.

Given the above, Theorem 1.2 quickly follows.

Proof of Theorem 1.2. — By (1.2), G^{-1} is absolutely continuous, strictly increasing, and has upper and lower bounds on its derivative $(G^{-1})(t)' = \frac{1}{g(G^{-1}(t))}$. Then by Propositions 3.3 and 3.4, there is a feasible W achieved by (S, d, μ) such that $\varphi = G^{-1} \circ F_W$ is subadditive. Since both G^{-1} and F_W are continuous, by Theorem 3.2, φ is strongly metric preserving, and so $(S, \varphi \circ d, \mu)$ achieves θ. □

Proof of Proposition 3.4. — Let $\alpha \in (0, \frac{1}{2})$ to be chosen later, define $H : \mathbb{R} \to [0,1]$ by

$$H(t) = \begin{cases} 0 & t < 0 \\ t^\alpha & 0 \leqslant t \leqslant 1 \\ 1 & t > 1, \end{cases}$$

and let μ be the distribution whose cumulative distribution function is H. This distribution has density $h(t) = \alpha t^{\alpha-1}$ for $t \in [0,1]$. Let $W = |X - Y|$, where $X, Y \sim \mu$ are i.i.d. By definition, W is feasible. For $t \in [0,1]$, the cumulative distribution function F_W is given by

$$\begin{aligned}F_W(t) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)h(y)\mathbf{1}_{|x-y|\leqslant t}\,dydx \\ &= 2\int_0^1\int_x^{x+t} h(x)h(y)\,dydx \\ &= 2\int_0^1 h(x)(H(x+t) - H(x))\,dx \\ &= 2\int_0^{1-t} \alpha x^{2\alpha-1}\left(\left(1+\frac{t}{x}\right)^\alpha - 1\right)dx + 2\int_{1-t}^1 \alpha x^{\alpha-1}(1 - x^\alpha)\,dx.\end{aligned}$$

The function F_W can readily be shown to be concave on $[0,\infty)$ by calculating the second derivative of $2\int_0^1\int_x^{x+t} h(x)h(y)dydx$. Thus, for all $z,w\in[0,\infty)$,

$$F_W(z)\leqslant F_W(w)+(z-w)F_W'(w).$$

Let $x,y\in(0,\infty)$ with $x\leqslant y$. Choosing $z=x+y$ and $w=y$, we get

$$\begin{aligned}F_W(x+y)&\leqslant F_W(y)+xF_W'(y)\\&\leqslant F_W(y)+xF_W'(x)\\&=F_W(y)+F_W(x)\cdot\frac{xF_W'(x)}{F_W(x)}.\end{aligned}$$

It therefore suffices to show that for all $t\in[0,\infty)$,

$$\frac{tF_W'(t)}{F_W(t)}\leqslant\varepsilon.$$

As F_W is concave and increasing, it is enough to show a bound of $\frac{1}{2}\varepsilon$ for all $t\leqslant\frac{1}{2}$, since for all $t>\frac{1}{2}$, $F_W'(t)\leqslant F_W'(\frac{1}{2})$ and $F_W(t)\geqslant F_W(\frac{1}{2})$. Assuming that $t\leqslant\frac{1}{2}$, the denominator is bounded by

$$\begin{aligned}F_W(t)&=2\alpha\Big(\int_0^{1-t}x^{2\alpha-1}\Big(\Big(1+\frac{t}{x}\Big)^{\alpha}-1\Big)\,dx+\int_{1-t}^1 x^{\alpha-1}(1-x^{\alpha})\,dx\Big)\\&\geqslant 2\alpha\int_0^t x^{2\alpha-1}\Big(\Big(\frac{t}{x}\Big)^{\alpha}-1\Big)\,dx=2\alpha\Big(\frac{1}{\alpha}t^{2\alpha}-\frac{1}{2\alpha}t^{2\alpha}\Big)=t^{2\alpha}.\end{aligned}$$

Recalling that $\alpha<\frac{1}{2}$, the numerator is bounded by

$$\begin{aligned}F_W'(t)&=2\int_0^1 h(x)h(x+t)\,dx\\&=2\alpha^2\Big(\int_0^t x^{2\alpha-2}\Big(1+\frac{t}{x}\Big)^{\alpha-1}dx+\int_t^{1-t}x^{2\alpha-2}\Big(1+\frac{t}{x}\Big)^{\alpha-1}dx\Big)\\&\leqslant 2\alpha^2\Big(\int_0^t x^{2\alpha-2}\Big(\frac{t}{x}\Big)^{\alpha-1}dx+\int_t^{1-t}x^{2\alpha-2}\,dx\Big)\\&=2\alpha t^{2\alpha-1}-2\alpha^2\frac{1}{1-2\alpha}\Big(\frac{1}{(1-t)^{1-2\alpha}}-\frac{1}{t^{1-2\alpha}}\Big)\\&\leqslant 2\alpha\frac{1}{t^{1-2\alpha}}\Big(1+\frac{\alpha}{1-2\alpha}\Big).\end{aligned}$$

Choosing $\alpha\leqslant\frac{1}{4}$ then gives $F_W'(t)\leqslant 4\alpha t^{2\alpha-1}$. We thus have

$$\frac{tF_W'(t)}{F_W(t)}\leqslant\frac{4\alpha t^{2\alpha}}{t^{2\alpha}}=4\alpha,$$

and the proposition follows for $\alpha\leqslant\frac{1}{8}\varepsilon$. □

Remark 3.5. — The function F_W may be expressed in a slightly more closed form using hypergeometric functions; a short calculation reveals that for $t\in[0,1]$,

$$F_W(t)=1+2(1-t)^{\alpha}({}_2F_1(-\alpha,1;1+\alpha;1-t)-1),$$

where ${}_2F_1(a,b;c;z)$ is the hypergeometric function

$$ {}_2F_1(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 \frac{x^{b-1}(1-x)^{c-b-1}}{(1-xz)^a}\, dx. $$

A plot of F_W for some values of α is given in Figure 3.1, where the asymptotic behavior of $t^{2\alpha}$ can be seen near the origin.

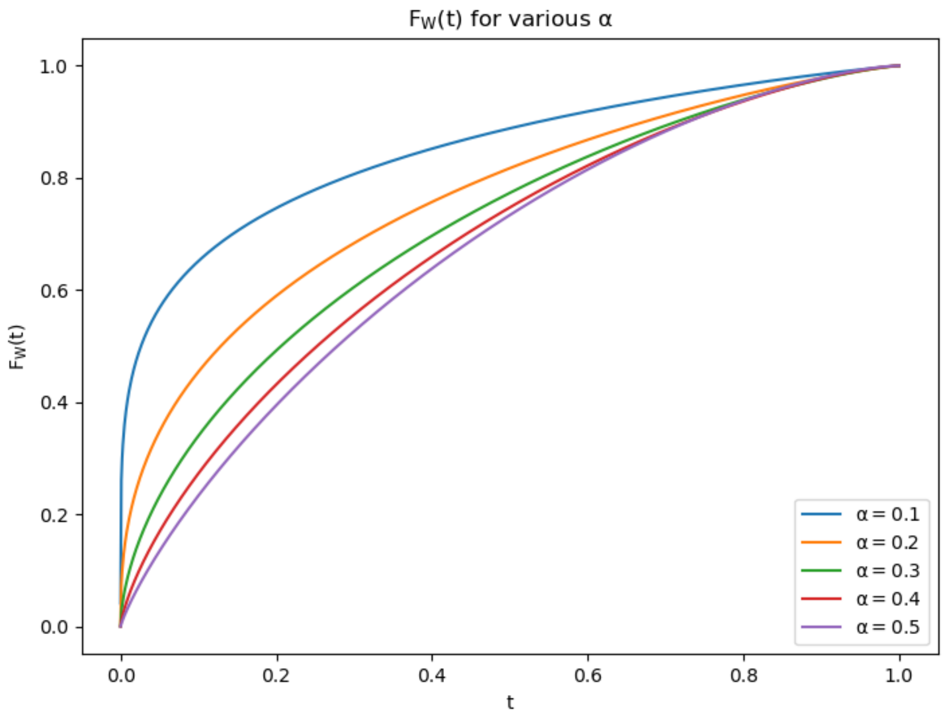

Figure 3.1. The function F_W for various values of α.

Remark 3.6. — It might perhaps be possible to extend the proof method of Theorem 1.2 to distributions on $[0,\infty)$ where the density is always positive. Indeed, if we compose G^{-1} with another function f, we still have

$$ \begin{aligned} G^{-1}(f(x+y)) &= G^{-1}(f(y)) + \int_{f(y)}^{f(x+y)} \frac{d}{dt} G^{-1}(t)\, dt \\ &\leqslant G^{-1}(f(y)) + (f(x+y)-f(y)) \sup_{t\in[f(y),\, f(x+y)]} \frac{d}{dt} G^{-1}(t). \end{aligned} $$

An analogue of Theorem 1.2 would follow if we could find, for each $\varepsilon > 0$, an unbounded feasible random variable W such that for all $x \leqslant y$

$$ F_W(x+y) \leqslant F_W(y) + \left(\varepsilon \sup_{t\in[F_W(y),\, F_W(x+y)]} \frac{d}{dt} G^{-1}(t) \right) F_W(x). $$

Unlike the random variable from Proposition 3.4, such a W would have to rely directly on G in some way. However, the method cannot be easily extended to distributions whose density is 0 in an interval: this would imply that $G^{-1} \circ F_W$ is discontinuous, which would prevent it from preserving the metric.

BIBLIOGRAPHY

[ABC24] David J. Aldous, Guillaume Blanc, and Nicolas Curien, *The distance problem on measured metric spaces*, 2024, https://arxiv.org/abs/2403.10926. ↑1023, 1025

[Cor99] Paul Corazza, *Introduction to metric-preserving functions*, Am. Math. Mon. **106** (1999), no. 4, 309–323. ↑1031

[Kel75] John L. Kelley, *General topology*, Graduate Texts in Mathematics, vol. 27, Springer, 1975, reprint of the 1955 original published by van Nostrand, Toronto. ↑1031

Manuscript received on 22nd January 2025,
revised on 11th July 2025,
accepted on 18th July 2025.

Recommended by Editors S. Gouëzel and D. Aldous.

eISSN: 2644-9463
This journal is a member of Centre Mersenne.

Renan GROSS
University of Cambridge,
Faculty of Mathematics,
Wilberforce Road,
Cambridge CB3 0WA (United Kingdom)
rg751@cam.ac.uk