Title

The semigroup of metric measure spaces and its infinitely divisible probability measures

Journal

Transactions of the American Mathematical Society, 369(3)

ISSN

0002-9947

Authors

Evans, Steven N

Molchanov, Ilya

Publication Date

2017

DOI

10.1090/tran/6714

Peer reviewed

THE SEMIGROUP OF METRIC MEASURE SPACES AND ITS INFINITELY DIVISIBLE PROBABILITY MEASURES

STEVEN N. EVANS AND ILYA MOLCHANOV

ABSTRACT. A metric measure space is a complete, separable metric space equipped with a probability measure that has full support. Two such spaces are equivalent if they are isometric as metric spaces via an isometry that maps the probability measure on the first space to the probability measure on the second. The resulting set of equivalence classes can be metrized with the Gromov–Prohorov metric of Greven, Pfaffelhuber and Winter. We consider the natural binary operation $\boxplus$ on this space that takes two metric measure spaces and forms their Cartesian product equipped with the sum of the two metrics and the product of the two probability measures. We show that the metric measure spaces equipped with this operation form a cancellative, commutative, Polish semigroup with a translation invariant metric. There is an explicit family of continuous semicharacters that is extremely useful for, *inter alia*, establishing that there are no infinitely divisible elements and that each element has a unique factorization into prime elements.

We investigate the interaction between the semigroup structure and the natural action of the positive real numbers on this space that arises from scaling the metric. For example, we show that for any given positive real numbers a, b, c the trivial space is the only space $\mathcal{X}$ that satisfies $a\mathcal{X} \boxplus b\mathcal{X} = c\mathcal{X}$.

We establish that there is no analogue of the law of large numbers: if $\mathbf{X}_1, \mathbf{X}_2, \dots$ is an identically distributed independent sequence of random spaces, then no subsequence of $\frac{1}{n}\boxplus_{k=1}^{n} \mathbf{X}_k$ converges in distribution unless each $\mathbf{X}_k$ is almost surely equal to the trivial space. We characterize the infinitely divisible probability measures and the Lévy processes on this semigroup, characterize the stable probability measures and establish a counterpart of the LePage representation for the latter class.

Date: February 23, 2015.

2010 *Mathematics Subject Classification.* 43A05, 60B15, 60E07, 60G51.

Key words and phrases. Gromov–Prohorov metric, cancellative semigroup, monoid, Delphic semigroup, semicharacter, irreducible, prime, unique factorization, Lévy-Hinčin formula, Itô representation, Lévy process, stable probability measure, LePage representation, law of large numbers.

SNE supported in part by NSF grant DMS-09-07630 and NIH grant 1R01GM109454-01. IM supported in part by Swiss National Science Foundation grant 200021-137527.

1. Introduction

The Cartesian product $G \square H$ of two finite graphs G and H with respective vertex sets $V(G)$ and $V(H)$ and respective edge sets $E(G)$ and $E(H)$ is the graph with vertex set $V(G \square H) := V(G) \times V(H)$ and edge set

$$\begin{aligned} E(G \square H) := &\{((g', h), (g'', h)) : (g', g'') \in E(G),\, h \in V(H)\} \\ &\cup \{((g, h'), (g, h'')) : g \in V(G),\, (h', h'') \in E(H)\}. \end{aligned}$$

This construction plays a role in many areas of graph theory. For example, it is shown in [Sab60] that any connected finite graph is isomorphic to a Cartesian product of graphs that are irreducible in the sense that they cannot be represented as Cartesian products and that this representation is unique up to the order of the factors (see, also, [Viz63, Mil70, Imr71, Wal87, AFDF00, Tar92]). The study of the problem of embedding a graph in a Cartesian product was initiated in [GW85, GW84]. A comprehensive review of factorization and embedding problems is [Win87].

If two connected finite graphs G and H are equipped with the usual shortest path metrics r_G and r_H, then the shortest path metric on the Cartesian product is given by $r_{G \times H} = r_G \oplus r_H$, where

$$\begin{aligned} (r_G \oplus r_H)((g', h'), (g'', h'')) := r_G(g', g'') + r_H(h', h''), \\ (g', h'),\, (g'', h'') \in G \times H. \end{aligned}$$

We use the notation $\oplus$ because if we think of the shortest path metric on a finite graph as a matrix, then the matrix for the shortest path metric on the Cartesian product of two graphs is the Kronecker sum of the matrices for the two graphs and the $\oplus$ notation is commonly used for the Kronecker sum [SH11].

It is natural to consider the obvious generalization of this construction to arbitrary metric spaces and there is a substantial literature in this direction. For example, a related binary operation on metric spaces is considered by Ulam [Mau81, Problem 77(b)] who constructs a metric on the Cartesian product of two metric spaces (Y, r_Y) and (Z, r_Z) via $((y', z'), (y'', z'')) \mapsto \sqrt{r_Y(y', y'')^2 + r_Z(z', z'')^2}$ and asks whether it is possible that there could be two nonisometric metric spaces U and V such that the metrics spaces $U \times U$ and $V \times V$ are isometric. An example of two such spaces is given in [Fou71]. However, it follows from the results of [Gru70, Mos92] that such an example is not possible if U and V are compact subsets of a Euclidean space.

On the other hand, a classical result of de Rahm [dR52] says that a complete, simply connected, Riemannian manifold has a product decomposition $M_0 \times M_1 \times \cdots \times M_k$, where the manifold M_0 is a Euclidean space (perhaps just a point) and M_i, $i = 1, \ldots, k$, are irreducible Riemannian manifolds that each have more than one point and are not isometric to the real line. By convention, the metric on a product of manifolds is the one appearing in Ulam's problem. This last result was extended to the setting of geodesic metric spaces of finite dimension in [FL08].

Ulam's problem is closely related to the question of *cancellativity* for this binary operation; that is, if $Y \times Z'$ and $Y \times Z''$ are isometric, then are Z' and Z'' isometric? This property clearly does not hold in general; for example, $\ell^2(\mathbb{N}) \times \ell^2(\mathbb{N})$ and $\ell^2(\mathbb{N})$ (where $\mathbb{N} := \{0, 1, 2, \ldots\}$) are isometric, but $\ell^2(\mathbb{N})$ and the trivial metric space are not isometric. Moreover, an example is given in [Her94] showing that it does not even hold for arbitrary subsets of $\mathbb{R}$. However, we note from [BP95] that there are many compact Hausdorff topological spaces K with the property that if L' and L'' are two compact Hausdorff spaces such that $K \times L'$ and $K \times L''$ are homeomorphic, then L' and L'' are homeomorphic (see also [Zer01]).

Returning to the binary operation that combines two metric spaces (Y, r_Y) and (Z, r_Z) into the metric space $(Y \times Z, r_Y \oplus r_Z)$, it is shown in [Tar92] that if a metric space is isometric to a product of finitely many irreducible metric spaces, then this factorization is unique up to the order of the factors. However, there are certainly metric spaces that are not isometric to a finite product of finitely many irreducible metric spaces and the study of this binary operation seems to be generally rather difficult.

In this paper we consider a closely-related binary operation on the class of *metric measure spaces*; that is, objects that consist of a complete, separable metric space (X, r_X) equipped with a probability measure μ_X that has full support. Following [Gro99] (see, also, [Ver98, Ver03, Ver04]), we regard two such spaces as being equivalent if they are isometric as metric spaces with an isometry that maps the probability measure on the first space to the probability measure on the second. Denote by $\mathbb{M}$ the set of such equivalence classes. With a slight abuse of notation, we will not distinguish between an equivalence class $\mathcal{X} \in \mathbb{M}$ and a representative triple (X, r_X, μ_X).

Gromov and Vershik show that a metric measure space (X, r_X, μ_X) is uniquely determined by the distribution of the infinite random matrix of distances

$$(r_X(\xi_i, \xi_j))_{(i,j) \in \mathbb{N} \times \mathbb{N}},$$

where $(\xi_k)_{k\in\mathbb{N}}$ is an i.i.d. sample of points in X with common distribution μ_X, and this concise condition for equivalence makes metric measure spaces considerably easier to study than metric spaces *per se*. A probability measure Q on the cone $\mathcal{R} := \{(r_{ij})_{(i,j)\in\mathbb{N}\times\mathbb{N}}\}$ of distance matrices is the distribution of a distance matrix corresponding to a metric measure space if and only if it is invariant and ergodic with respect to action of the infinite symmetric group and for all $\varepsilon > 0$ there exists integer N such that

$$Q\Big\{(r_{ij}) \in \mathcal{R} : \lim_{n\to\infty} \frac{\#\{j : 1 \leqslant j \leqslant n, \min_{1\leqslant i\leqslant N} r_{ij} < \varepsilon\}}{n} > 1-\varepsilon\Big\} > 1-\varepsilon, \tag{1.1}$$

see [Ver03].

Definition 1.1. Define a binary, associative, commutative operation $\boxplus$ on $\mathbb{M}$ as follows. Given two elements $\mathcal{Y} = (Y, r_Y, \mu_Y)$ and $\mathcal{Z} = (Z, r_Z, \mu_Z)$ of $\mathbb{M}$, let $\mathcal{Y} \boxplus \mathcal{Z}$ be $\mathcal{X} = (X, r_X, \mu_X) \in \mathbb{M}$, where

- $X := Y \times Z$,
- $r_X := r_Y \oplus r_Z$, where $(r_Y \oplus r_Z)((y', z'), (y'', z'')) = r_Y(y', y'') + r_Z(z', z'')$ for $(y', z'), (y'', z'') \in Y \times Z$,
- $\mu_X := \mu_Y \otimes \mu_Z$.

The distribution of the random matrix of distances for $\mathcal{Y} \boxplus \mathcal{Z}$ is the convolution of the distributions of the random matrices of distances for $\mathcal{Y}$ and $\mathcal{Z}$. The equivalence class $\mathcal{E}$ of metric measure spaces that each consist of a single point with the only possible metric and probability measure on them is the neutral element for this operation, and so $(\mathbb{M}, \boxplus)$ is a commutative semigroup with an identity. A semigroup with an identity is sometimes called a *monoid*.

Remark 1.2. We could have chosen other ways to combine the metrics r_Y and r_Z to give a metric on $Y \times Z$ that induces the product topology and results in a counterpart of $\boxplus$ that is commutative and associative. For example, by analogy with Ulam's construction we could have used one of the metrics $((y', z'), (y'', z'')) \mapsto (r_Y(y', y'')^p + r_Z(z', z'')^p)^{\frac{1}{p}}$ for $p > 1$ or the metric $((y', z'), (y'', z'')) \mapsto r_Y(y', y'') \vee r_Z(z', z'')$. We do not investigate these possibilities here.

We finish this introduction with an overview of the remainder of the paper.

We show in Section 2 that if we equip $\mathbb{M}$ with the Gromov–Prohorov metric d_{GPr} introduced in [GPW09] (see Section 12 for the definition of d_{GPr}), then the binary operation $\boxplus : \mathbb{M} \times \mathbb{M} \to \mathbb{M}$ is continuous and the metric d_{GPr} is translation invariant for the operation $\boxplus$. We

recall from [GPW09] that $(\mathbb{M}, d_{\mathrm{GPr}})$ is a complete, separable metric space. Moreover, the Gromov–Prohorov metric has the property that a sequence of elements of $\mathbb{M}$ converges to an element of $\mathbb{M}$ if and only if the corresponding sequence of associated random distance matrices described above converges in distribution to the random distance matrix associated with the limit. In Section 2 we also introduce a partial order $\leqslant$ on $\mathbb{M}$ by declaring that $\mathcal{Y} \leqslant \mathcal{Z}$ if $\mathcal{Z} = \mathcal{Y} \boxplus \mathcal{X}$ for some $\mathcal{X} \in \mathbb{M}$ and show for any $\mathcal{Z} \in \mathbb{M}$ that the set $\{\mathcal{Y} \in \mathbb{M} : \mathcal{Y} \leqslant \mathcal{Z}\}$ is compact.

A semicharacter is a map $\chi : \mathbb{M} \to [0,1]$ such that $\chi(\mathcal{Y} \boxplus \mathcal{Z}) = \chi(\mathcal{Y})\chi(\mathcal{Z})$ for all $\mathcal{Y}, \mathcal{Z} \in \mathbb{M}$. We introduce a natural family of semicharacters in Section 3. This family has the property that $\lim_{n\to\infty} \mathcal{X}_n = \mathcal{X}$ for some sequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$ and element $\mathcal{X}$ in $\mathbb{M}$ if and only if $\lim_{n\to\infty} \chi(\mathcal{X}_n) = \chi(\mathcal{X})$ for all semicharacters χ in the family. Using the semicharacters, we characterize the existence of the limit $\lim_{n\to\infty} \boxplus_{k=0}^{n} \mathcal{X}_k$ for some sequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$, and show that if the limit exists, then $\boxplus_{k=0}^{n} \mathcal{X}'_k$ converges to the same limit for any rearrangement $(\mathcal{X}'_n)_{n\in\mathbb{N}}$ of the sequence. We also use the semicharacters to prove that $(\mathbb{M}, \boxplus)$ is cancellative.

We show in Section 4 that the irreducible elements (that is, those which cannot be decomposed as a nontrivial $\boxplus$ combination of elements of $\mathbb{M}$) form a dense, G_δ subset $\mathbb{I} \subset \mathbb{M}$. We give several examples of irreducible elements; for instance, all totally geodesic metric measure spaces are irreducible. Furthermore, there are no nontrivial infinitely divisible metric measure spaces (an element $\mathcal{X} \in \mathbb{M}$ is infinitely divisible if for every $n \geqslant 2$ it can be decomposed as the $\boxplus$-sum of n identical summands).

We establish in Section 5 that $(\mathbb{M}, \boxplus)$ is a Delphic semigroup as studied in [Ken68, Dav69]. By appealing to general results for Delphic semigroups, we confirm that each metric measure space is either irreducible or has an irreducible factor and then that any element of $\mathbb{M}\backslash\{\mathcal{E}\}$ has a representation as either a finite or countable $\boxplus$ combination of irreducible elements. We further show that this representation is unique up to the order of the "factors". The uniqueness does not follow from the Delphic theory and is based on a result showing that irreducible elements are prime (an element $\mathcal{X} \in \mathbb{M}\backslash\{\mathcal{E}\}$ is prime if $\mathcal{X} \leqslant \mathcal{Y} \boxplus \mathcal{Z}$ implies that $\mathcal{X} \leqslant \mathcal{Y}$ or $\mathcal{X} \leqslant \mathcal{Z}$). The latter result is analogous to a key fact in elementary number theory that can be proved using Euclid's algorithm and leads to the uniqueness of prime factorizations for the positive integers: if an integer $p \geqslant 2$ is such that the only divisors of p are 1 and p (the usual definition of p being prime and the analogue of the irreducibility property in our setting), then given any positive integers a and b such that p divides ab it must be the case that either

p divides a or p divides b (the analogue of primality in our setting) – see, for example, [And94, Corollary 2-3, Theorem 2-5].

In Section 6 we investigate the counting measure on the family $\mathbb{I}$ of irreducible elements that is obtained by taking an element of $\mathbb{M}$ and assigning a unit mass to each irreducible element (counted according to multiplicity) in its factorization. We show that this mapping from elements of $\mathbb{M}$ to counting measures on $\mathbb{M}$ concentrated on $\mathbb{I}$ is measurable in a natural sense.

Given $\mathcal{X} \in \mathbb{M}$ and $a > 0$, we define the rescaled metric measure space $a\mathcal{X} := (X, ar_X, \mu_X) \in \mathbb{M}$. We show in Section 7 that if $(a\mathcal{X}) \boxplus (b\mathcal{X}) = c\mathcal{X}$ for some $\mathcal{X} \in \mathbb{M}$ and $a, b, c > 0$, then $\mathcal{X} = \mathcal{E}$, so the second distributivity law certainly does not hold for this scaling operation.

We begin the study of random elements of $\mathbb{M}$ in Section 8 by defining a counterpart of the usual Laplace transform in which exponential functions are replaced by semicharacters. Two random elements of $\mathbb{M}$ have the same distribution if and only if their Laplace transforms are equal. A random element in $\mathbb{M}$ can be viewed, via its decomposition into irreducibles, as a point process on the set $\mathbb{I}$ of irreducible elements of $\mathbb{M}$.

We introduce the appropriate notion of infinitely divisible random elements of $\mathbb{M}$ in Section 9 and obtain an analogue of the classical Lévy–Hiňcin–Itô description of infinitely divisible real-valued random variables. Our approach to this result is probabilistic and involves constructing for any infinitely divisible random element a Lévy process that at time 1 has the same distribution as the given random element. Our setting resembles that of nonnegative infinitely divisible random variables and so there is no counterpart of a Gaussian component in this description. Also, there is no deterministic component because the only constant that is infinitely divisible is the trivial space $\mathcal{E}$.

Using the scaling operation on $\mathbb{M}$ we define stable random elements of $\mathbb{M}$ in Section 10. We determine how the Lévy–Hiňcin–Itô description specializes to such random elements and also verify that there is a counterpart of the LePage series that represents a stable bounded metric measure space as an "infinite weighted sum" of independent identically distributed random elements in $\mathbb{M}$ with a suitable independent sequence of coefficients.

The representation of random elements of $\mathbb{M}$ as point processes on the set $\mathbb{I}$ of irreducible spaces makes it possible in Section 11 to introduce a thinning operation that takes an element of $\mathbb{M}$ and produces another by randomly discarding some of the irreducible factors. Lévy

processes on $\mathbb{M}$ necessarily have nondecreasing sample paths with respect to the partial order $\leqslant$, but by combining thinning with the addition of independent random increments one can produce Markov processes with sample paths that are not monotone. Also, thinning can be used to define a notion of discrete stable random elements in $\mathbb{M}$.

For ease of reference we summarize some facts about the Gromov–Prohorov metric in Section 12. Many of our arguments can be carried through using alternative metrics on $\mathbb{M}$ or its subfamilies such as the $\mathbb{D}$-metric studied in [Stu06]. Lastly, in Section 13 we obtain a bound on the Laplace transform of nonnegative random variables that was useful in Section 3.

2. Topological and order properties

Lemma 2.1. *The operation $\boxplus : \mathbb{M} \times \mathbb{M} \to \mathbb{M}$ is continuous. More specifically, if $\mathcal{X}_i, \mathcal{Y}_i$, $i = 1, 2$, are elements of $\mathbb{M}$, then*

$$d_{\mathrm{GPr}}(\mathcal{X}_1 \boxplus \mathcal{X}_2, \mathcal{Y}_1 \boxplus \mathcal{Y}_2) \leqslant d_{\mathrm{GPr}}(\mathcal{X}_1, \mathcal{Y}_1) + d_{\mathrm{GPr}}(\mathcal{X}_2, \mathcal{Y}_2)\,.$$

Proof. Let ϕ_{X_i} and ϕ_{Y_i} be isometries from X_i and Y_i to a common metric measure space $\mathcal{Z}_i$, $i = 1, 2$. The combined function (ϕ_{X_1}, ϕ_{X_2}) (resp. (ϕ_{Y_1}, ϕ_{Y_2})) maps $X_1 \times X_2$ (resp. $Y_1 \times Y_2$) isometrically into $Z_1 \times Z_2$. The result now follows from Lemma 12.1. □

A proof similar to that of Lemma 2.1 using Lemma 12.2 establishes the following result.

Lemma 2.2. *The metric d_{GPr} is translation invariant for the operation $\boxplus$. That is, if $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}$ are elements of $\mathbb{M}$, then*

$$d_{\mathrm{GPr}}(\mathcal{X}_1 \boxplus \mathcal{Y}, \mathcal{X}_2 \boxplus \mathcal{Y}) = d_{\mathrm{GPr}}(\mathcal{X}_1, \mathcal{X}_2)\,.$$

In particular,

$$d_{\mathrm{GPr}}(\mathcal{X}_1 \boxplus \mathcal{X}_2, \mathcal{X}_1) = d_{\mathrm{GPr}}(\mathcal{X}_2, \mathcal{E})\,.$$

Definition 2.3. Given $\mathcal{X} = (X, r_X, \mu_X) \in \mathbb{M}$, write $\mathrm{diam}(\mathcal{X})$ for the (possibly infinite) diameter of the metric space X; that is,

$$\mathrm{diam}(\mathcal{X}) := \sup\{r_X(x', x'') : x', x'' \in X\}.$$

The next result is obvious.

Lemma 2.4. *The diameter is an additive functional on $(\mathbb{M}, \boxplus)$; that is,*

$$\mathrm{diam}(\mathcal{X} \boxplus \mathcal{Y}) = \mathrm{diam}(\mathcal{X}) + \mathrm{diam}(\mathcal{Y})$$

for all $\mathcal{X}, \mathcal{Y} \in \mathbb{M}$.

Remark 2.5. The function diam is not continuous even on the family $\mathbb{K}$ of *compact* metric measure spaces. For example, let $\mathcal{X}_n = (\{0,1\}, r, \mu_n)$, where $r(0,1) = 1$, $\mu_n\{0\} = 1 - \frac{1}{n}$ and $\mu_n\{1\} = \frac{1}{n}$. Then, $\mathcal{X}_n$ converges to the trivial space $\mathcal{E}$, whereas $\operatorname{diam}(\mathcal{X}_n) = 1 \not\to 0 = \operatorname{diam}(\mathcal{E})$.

Lemma 2.6. *The function* diam *is lower semicontinuous on* $\mathbb{M}$. *That is, if the sequence* $\mathcal{X}_n \to \mathcal{X}$ *in* $\mathbb{M}$ *as* $n \to \infty$, *then* $\operatorname{diam}(\mathcal{X}) \leqslant \liminf_{n\to\infty} \operatorname{diam}(\mathcal{X}_n)$.

Proof. Suppose that the sequence $\mathcal{X}_n$ converges to $\mathcal{X}$, $(\xi_k^{(n)})_{k\in\mathbb{N}}$ are i.i.d. in X_n with the common distribution μ_{X_n}, and $(\xi_k)_{k\in\mathbb{N}}$ are i.i.d. in X with the common distribution μ_X. Observe for any k that $\max_{1\leqslant i<j\leqslant k}(r_{X_n}(\xi_i^{(n)}, \xi_j^{(n)})$ converges in distribution to $\max_{1\leqslant i<j\leqslant k}(r_X(\xi_i, \xi_j))$. It suffices to note that $\max_{1\leqslant i<j\leqslant k}(r_{X_n}(\xi_i^{(n)}, \xi_j^{(n)}))$ is increasing in k and converges almost surely to $\operatorname{diam}(\mathcal{X}_n)$ as $k \to \infty$ and that $\max_{1\leqslant i<j\leqslant k}(r_X(\xi_i, \xi_j))$ is increasing in k and converges almost surely to $\operatorname{diam}(\mathcal{X})$ as $k \to \infty$. □

Definition 2.7. Define a partial order $\leqslant$ on $\mathbb{M}$ by setting $\mathcal{Y} \leqslant \mathcal{Z}$ if $\mathcal{Z} = \mathcal{Y} \boxplus \mathcal{X}$ for some $\mathcal{X} \in \mathbb{M}$.

The symmetry and transitivity of $\leqslant$ is obvious. The antisymmetry is apparent from Lemma 2.8 below. This partial order is the dual of the *Green* or *divisibility* order (see [Gri01, Section I.4.1]). The identity $\mathcal{E}$ is the unique minimal element.

Lemma 2.8. *If* $\mathcal{X} \leqslant \mathcal{Y} \leqslant \mathcal{Z}$, *then* $d_{\mathrm{GPr}}(\mathcal{X}, \mathcal{Y}) \leqslant d_{\mathrm{GPr}}(\mathcal{X}, \mathcal{Z})$.

Proof. In view of Proposition 3.6(a) and Lemma 2.2, it suffices to assume that $\mathcal{X} = \mathcal{E}$. If $\mathcal{Z} = \mathcal{Y} \boxplus \mathcal{V}$, then (12.1) yields that

$$\begin{aligned} d_{\mathrm{GPr}}(\mathcal{Z}, \mathcal{E}) &= \inf_{y\in\mathcal{Y}, v\in\mathcal{V}} \inf\{\varepsilon > 0 : \mu_Y \otimes \mu_V\{(y', v') : \\ &\qquad\qquad r_Y(y, y') + r_V(v, v') \geqslant \varepsilon\} \leqslant \varepsilon\} \\ &\geqslant \inf_{y\in\mathcal{Y}, v\in\mathcal{V}} \inf\{\varepsilon > 0 : \mu_Y \otimes \mu_V\{(y', v') : r_X(y, y') \geqslant \varepsilon\} \leqslant \varepsilon\} \\ &= d_{\mathrm{GPr}}(\mathcal{Y}, \mathcal{E})\,. \end{aligned}$$

□

An element of a semigroup with an identity is a *unit* if it has an inverse and a semigroup with an identity is said to be *reduced* if the only unit is the identity (see [Cli38, Section 1].

Corollary 2.9. *The semigroup* $(\mathbb{M}, \boxplus)$ *is reduced.*

Proof. Suppose that $\mathcal{E} = \mathcal{X} \boxplus \mathcal{Y}$, then $\mathcal{E} \leqslant \mathcal{X} \leqslant \mathcal{E}$ and the antisymmetry of the partial order $\leqslant$ gives that $\mathcal{E} = \mathcal{X} = \mathcal{Y}$. □

Lemma 2.10. a) *For any compact set $\mathbb{S} \subset \mathbb{M}$, the set $\bigcup_{\mathcal{Z}\in\mathbb{S}}\{\mathcal{Y} \in \mathbb{M} : \mathcal{Y} \leqslant \mathcal{Z}\}$ is compact.*

b) *For any compact set $\mathbb{S} \subset \mathbb{M}$, the set $\{(\mathcal{Y}, \mathcal{Z}) \in \mathbb{M}^2 : \mathcal{Z} \in \mathbb{S}, \mathcal{Y} \leqslant \mathcal{Z}\}$ is compact.*

c) *The map K from $\mathbb{M}$ to the compact subsets of $\mathbb{M}$ defined by $K(\mathcal{X}) := \{\mathcal{Y} \in \mathbb{M} : \mathcal{Y} \leqslant \mathcal{X}\}$ is upper semicontinuous. That is, if $F \subseteq \mathbb{M}$ is closed, then $\{\mathcal{X} \in \mathbb{M} : K(\mathcal{X}) \cap F \neq \varnothing\}$ is closed. Equivalently, if $\mathcal{X}_n \to \mathcal{X}$, and $\mathcal{Y}_n \in K(\mathcal{X}_n)$ converges to $\mathcal{Y}$, then $\mathcal{Y} \in K(\mathcal{X})$.*

Proof. (a) We first show that $\bigcup_{\mathcal{Z}\in\mathbb{S}}\{\mathcal{Y} \in \mathbb{M} : \mathcal{Y} \leqslant \mathcal{Z}\}$ is pre-compact. Given $\varepsilon > 0$, we know from [GPW09, Theorem 2] that there exist $K > 0$ and $\delta > 0$ such that for all $\mathcal{Z} \in \mathbb{S}$

$$\mu_Z \otimes \mu_Z\{(z', z'') \in Z \times Z : r_Z(z', z'') > K\} \leqslant \varepsilon$$

and

$$\mu_Z\{z' \in Z : \mu_Z\{z'' \in Z : r_Z(z', z'') < \varepsilon\} \leqslant \delta\} \leqslant \varepsilon.$$

If $\mathcal{Y} \leqslant \mathcal{Z}$ for some $\mathcal{Z} \in \mathbb{S}$, then, by definition, there is a $\mathcal{W} \in \mathbb{M}$ such that $\mathcal{Z} = \mathcal{Y} \boxplus \mathcal{W}$, and so

$$\begin{aligned}
&\mu_Y \otimes \mu_Y\{(y', y'') \in Y \times Y : r_Y(y', y'') > K\} \\
&\quad \leqslant (\mu_Y \otimes \mu_Y) \otimes (\mu_W \otimes \mu_W)\{((y', y''), (w', w'')) \in (Y \times Y) \times (W \times W) : \\
&\qquad r_Y(y', y'') + r_W(w', w'') > K\} \\
&\quad = \mu_Z \otimes \mu_Z\{(z', z'') \in Z \times Z : r_Z(z', z'') > K\} \\
&\quad \leqslant \varepsilon.
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\mu_Y\{y' \in Y : \mu_Y\{y'' \in Y : r_Y(y', y'') < \varepsilon\} \leqslant \delta\} \\
&\quad = \mu_Y \otimes \mu_W\{(y', w') \in Y \times W : \\
&\qquad \mu_y \otimes \mu_W\{(y'', w'') \in Y \times W : r_Y(y', y'') < \varepsilon\} \leqslant \delta\} \\
&\quad \leqslant \mu_Y \otimes \mu_W\{(y', w') \in Y \times W : \mu_Y \otimes \mu_W\{(y'', w'') \in Y \times W : \\
&\qquad r_Y(y', y'') + r_W(w', w'') < \varepsilon\} \leqslant \delta\} \\
&\quad = \mu_Z\{z' \in Z : \mu_Z\{z'' \in Z : r_Z(z', z'') < \varepsilon\} \leqslant \delta\} \\
&\quad \leqslant \varepsilon\,.
\end{aligned}$$

It follows from [GPW09, Theorem 2] that $\bigcup_{\mathcal{Z}\in\mathbb{S}}\{\mathcal{Y} \in \mathbb{M} : \mathcal{Y} \leqslant \mathcal{Z}\}$ is pre-compact.

We now show that $\bigcup_{\mathcal{Z}\in\mathbb{S}}\{\mathcal{Y} \in \mathbb{M} : \mathcal{Y} \leqslant \mathcal{Z}\}$ is closed, and hence compact. Suppose now that $(\mathcal{Y}_n)_{n\in\mathbb{N}}$ is a sequence in $\bigcup_{\mathcal{Z}\in\mathbb{S}}\{\mathcal{Y} \in \mathbb{M} : \mathcal{Y} \leqslant \mathcal{Z}\}$ that converges to a limit $\mathcal{Y}_\infty$. For each $n \in \mathbb{N}$ we can find $\mathcal{Z}_n \in \mathbb{S}$ and $\mathcal{W}_n \in \bigcup_{\mathcal{Z}\in\mathbb{S}}\{\mathcal{Y} \in \mathbb{M} : \mathcal{Y} \leqslant \mathcal{Z}\}$ such that $\mathcal{Z}_n = \mathcal{Y}_n \boxplus \mathcal{W}_n$. From the

above we can find a subsequence $(n(k))_{k\in\mathbb{N}}$, $\mathcal{Z}_\infty \in \mathbb{S}$ and $\mathcal{W}_\infty \in \mathbb{M}$ such that $\lim_{k\to\infty} \mathcal{Z}_{n(k)} = \mathcal{Z}_\infty$ and $\lim_{k\to\infty} \mathcal{W}_{n(k)} = \mathcal{W}_\infty$. By the continuity of the semigroup operation established in Lemma 2.1,

$$\mathcal{Y}_\infty \boxplus \mathcal{W}_\infty = \lim_{k\to\infty} (\mathcal{Y}_{n(k)} \boxplus \mathcal{W}_{n(k)}) = \lim_{k\to\infty} \mathcal{Z}_{n(k)} = \mathcal{Z}_\infty,$$

which implies that $\mathcal{Y}_\infty \leqslant \mathcal{Z}_\infty \in \mathbb{S}$ (and also $\mathcal{W}_\infty \leqslant \mathcal{Z}_\infty \in \mathbb{S}$). Therefore, $\bigcup_{\mathcal{Z}\in\mathbb{S}} \{\mathcal{Y} \in \mathbb{M} : \mathcal{Y} \leqslant \mathcal{Z}\}$ is closed and hence compact.
(b) Because $\{(\mathcal{Y}, \mathcal{Z}) \in \mathbb{M}^2 : \mathcal{Z} \in \mathbb{S},\, \mathcal{Y} \leqslant \mathcal{Z}\}$ is a subset of the compact set $(\bigcup_{\mathcal{Z}\in\mathbb{S}} \{\mathcal{Y} \in \mathbb{M} : \mathcal{Y} \leqslant \mathcal{Z}\}) \times \mathbb{S}$, it suffices to show that the former set is closed, but this follows from an argument similar to that which completed the proof of part (a).
(c) This is immediate from (b). □

3. Semicharacters

Following the standard terminology in semigroup theory, a *semicharacter* is a map $\chi : \mathbb{M} \to [0, 1]$ such that $\chi(\mathcal{Y} \boxplus \mathcal{Z}) = \chi(\mathcal{Y})\chi(\mathcal{Z})$ for all $\mathcal{Y}, \mathcal{Z} \in \mathbb{M}$.

Definition 3.1. Denote by $\mathbb{A}$ the set consisting of the empty set and the arrays $A = (a_{ij})_{1\leqslant i<j\leqslant n} \in \mathbb{R}_+^{\binom{n}{2}}$ for $n \geqslant 2$. For each $A \in \mathbb{A}$ define a semicharacter χ_A by setting $\chi_\varnothing \equiv 1$ and

$$\chi_A((X, r_X, \mu_X)) := \int_{X^n} \exp\left(- \sum_{1\leqslant i<j\leqslant n} a_{ij} r_X(x_i, x_j) \right) \mu_X^{\otimes n}(dx) \tag{3.1}$$

if $A \neq \varnothing$. Note that $\chi_A(\mathcal{X}) > 0$ for all $A \in \mathbb{A}$ and $\mathcal{X} \in \mathbb{M}$. We often need the particular semicharacter

$$\chi_1(\mathcal{X}) := \int_{X^2} \exp(-r_X(x_1, x_2))\, \mu_X^{\otimes 2}(dx) \tag{3.2}$$

defined by taking as $A \in \mathbb{A}$ an array with the single element 1.

As we recalled in the Introduction, a metric measure space (X, r_X, μ_X) is uniquely determined by the distribution of the infinite random matrix of distances $(r_X(\xi_i, \xi_j))_{(i,j)\in\mathbb{N}\times\mathbb{N}}$, where $(\xi_k)_{k\in\mathbb{N}}$ is an i.i.d. sample of points in X with common distribution μ_X. The next lemma follows immediately from this observation and the unicity of Laplace transforms.

Lemma 3.2. a) *Two elements $\mathcal{X}, \mathcal{Y} \in \mathbb{M}$ are equal if and only if $\chi_A(\mathcal{X}) = \chi_A(\mathcal{Y})$ for all $A \in \mathbb{A}$.*
b) *If $\mathcal{Y} \leqslant \mathcal{X}$, then $\chi_A(\mathcal{X}) \geqslant \chi_A(\mathcal{Y})$ for all $A \in \mathbb{A}$.*

Remark 3.3. Note that if $A' \in \mathbb{R}_+^{\binom{n'}{2}}$ and $A'' \in \mathbb{R}_+^{\binom{n''}{2}}$, then $\chi_{A'}\chi_{A''} = \chi_A$, where $A \in \mathbb{R}_+^{\binom{n'+n''}{2}}$ is given by

$$a_{ij} = \begin{cases} a'_{ij}, & 1 \leqslant i < j \leqslant n' \\ a''_{i-n',j-n'}, & n'+1 \leqslant i < j < n'+n''. \end{cases}$$

It follows that $\{\chi_A : A \in \mathbb{A}\}$ is a semigroup with identity $\chi_\varnothing \equiv 1$.

Remark 3.4. Not all semicharacters of $\mathbb{M}$ are of the form χ_A for some $A \in \mathbb{A}$. For example, if $A \in \mathbb{A}$ and $\beta > 0$, then $\mathcal{X} \mapsto \chi_A(\mathcal{X})^\beta$ is a (continuous) semicharacter. If X has two points, say 0 and 1, that are distance r apart and $\mu_X(\{0\}) = (1-p)$ and $\mu_X(\{1\}) = p$ for some $0 < p < 1$, then taking A to be the array with the single element a we have $\chi_A(\mathcal{X}) = (1-p)^2 + p^2 + 2p(1-p)\exp(-ar)$ and it is not hard to see from considering just $\mathcal{X}$ of this special type that for $\beta \neq 1$ the semicharacter χ_A^β is not of the form $\chi_{A'}$ for some other $A \in \mathbb{A}$.

It follows from Lemma 2.4 that $\mathcal{X} \mapsto \exp(-\operatorname{diam}(\mathcal{X}))$ is a (discontinuous) semicharacter on $\mathbb{M}$. Also, if $A \in \mathbb{A}$ and $b > 0$, then

$$\left(\int_{X^n} \exp\left(\sum_{1 \leqslant i < j \leqslant n} a_{ij} r_X(x_i, x_j) \right) \mu_X^{\otimes n}(dx) \right)^{-b}$$

is a (discontinuous) semicharacter. These last two examples are connected by the observation that

$$\exp(-\operatorname{diam}(\mathcal{X})) = \lim_{t \to \infty} \left(\int_{X^2} \exp\left(t\, r_X(x_1, x_2)\right)\, \mu_X^{\otimes 2}(dx) \right)^{-\frac{1}{t}}.$$

Lemma 3.5. *A sequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$ converges to $\mathcal{X} \in \mathbb{M}$ if and only if $\lim_{n\to\infty} \chi_A(\mathcal{X}_n) = \chi_A(\mathcal{X})$ for all $A \in \mathbb{A}$.*

Proof. For $n \in \mathbb{N}$, let $(\xi_k^{(n)})_{k\in\mathbb{N}}$ be an i.i.d. sequence of X_n-valued random variables with common distribution μ_{X_n}, and let $(\xi_k)_{k\in\mathbb{N}}$ be an i.i.d. sequence of X-valued random variables with common distribution μ_X. It follows from [GPW09, Theorem 5] that $\mathcal{X}_n$ converges to $\mathcal{X}$ if and only if the distribution of $(r_{X_n}(\xi_i^{(n)}, \xi_j^{(n)}))_{1\leqslant i<j\leqslant m}$ converges to that of $(r_X(\xi_i, \xi_j))_{1\leqslant i<j\leqslant m}$ for all $m \in \mathbb{N}$. The result is now a consequence of the equivalence between the weak convergence of probability measures on $\mathbb{R}_+^{\binom{m}{2}}$ and the convergence of their Laplace transforms. □

In the usual terminology of semigroup theory, part (a) of the following result says that the semigroup $(\mathbb{M}, \boxplus)$ is *cancellative* (see [Gri01, Section II.1.1]).

Proposition 3.6. a) *Suppose that* $\mathcal{Y}, \mathcal{Z}', \mathcal{Z}'' \in \mathbb{M}$ *satisfy* $\mathcal{Y} \boxplus \mathcal{Z}' = \mathcal{Y} \boxplus \mathcal{Z}''$*, then* $\mathcal{Z}' = \mathcal{Z}''$*. If* $\mathcal{Y} \boxplus \mathcal{Z}' \preccurlyeq \mathcal{Y} \boxplus \mathcal{Z}''$*, then* $\mathcal{Z}' \preccurlyeq \mathcal{Z}''$*.*

b) *Consider sequences* $(\mathcal{X}_n)_{n\in\mathbb{N}}$ *and* $(\mathcal{Y}_n)_{n\in\mathbb{N}}$ *in* $\mathbb{M}$*. Set* $\mathcal{Z}_n := \mathcal{X}_n \boxplus \mathcal{Y}_n$*. Suppose that* $\mathcal{X} := \lim_{n\to\infty} \mathcal{X}_n$ *and* $\mathcal{Z} := \lim_{n\to\infty} \mathcal{Z}_n$ *exist. Then,* $\mathcal{Y} := \lim_{n\to\infty} \mathcal{Y}_n$ *exists and* $\mathcal{Z} = \mathcal{X} \boxplus \mathcal{Y}$*.*

Proof. a) For each semicharacter χ_A, $A \in \mathbb{A}$, we have $\chi_A(\mathcal{Y})\chi_A(\mathcal{Z}') = \chi_A(\mathcal{X}) = \chi_A(\mathcal{Y})\chi_A(\mathcal{Z}'')$ and so $\chi_A(\mathcal{Z}') = \chi_A(\mathcal{Z}'')$, which implies that $\mathcal{Z}' = \mathcal{Z}''$. In case of the inequality, $\mathcal{Y} \boxplus \mathcal{Z}' \boxplus \mathcal{W} = \mathcal{Y} \boxplus \mathcal{Z}''$, so that $\mathcal{Z}' \boxplus \mathcal{W} = \mathcal{Z}''$ and hence $\mathcal{Z}' \preccurlyeq \mathcal{Z}''$.
b) By Lemma 2.10(a), the sequence $(\mathcal{Y}_n)_{n\in\mathbb{N}}$ is pre-compact. Any subsequential limit $\mathcal{Y}_\infty$ will satisfy $\mathcal{Z} = \mathcal{X} \boxplus \mathcal{Y}_\infty$. It follows from part (a) that $\mathcal{Y} := \lim_{n\to\infty} \mathcal{Y}_n$ exists and $\mathcal{Z} = \mathcal{X} \boxplus \mathcal{Y}$ in view of Lemma 3.2(a). □

Remark 3.7. It is a consequence of Proposition 3.6(a) and the discussion in Section 1.10 of [CP61] that the semigroup $(\mathbb{M}, \boxplus)$ can be embedded into a group $\mathbb{G}$ as follows. Equip $\mathbb{M} \times \mathbb{M}$ with the equivalence relation $\equiv$ defined by $(\mathcal{W}, \mathcal{X}) \equiv (\mathcal{Y}, \mathcal{Z})$ if $\mathcal{W} \boxplus \mathcal{Z} = \mathcal{X} \boxplus \mathcal{Y}$. It is not hard to see that $\equiv$ is indeed an equivalence relation, the only property that is not completely obvious is transitivity. However, if $(\mathcal{U}, \mathcal{V}) \equiv (\mathcal{W}, \mathcal{X})$ and $(\mathcal{W}, \mathcal{X}) \equiv (\mathcal{Y}, \mathcal{Z})$, then, by definition, $\mathcal{U} \boxplus \mathcal{X} = \mathcal{V} \boxplus \mathcal{W}$ and $\mathcal{W} \boxplus \mathcal{Z} = \mathcal{X} \boxplus \mathcal{Y}$ so that

$$\begin{aligned}(\mathcal{U} \boxplus \mathcal{Z}) \boxplus (\mathcal{X} \boxplus \mathcal{W}) &= (\mathcal{U} \boxplus \mathcal{X}) \boxplus (\mathcal{W} \boxplus \mathcal{Z}) \\ &= (\mathcal{V} \boxplus \mathcal{W}) \boxplus (\mathcal{X} \boxplus \mathcal{Y}) = (\mathcal{V} \boxplus \mathcal{Y}) \boxplus (\mathcal{X} \boxplus \mathcal{W}),\end{aligned}$$

from which we see that $\mathcal{U} \boxplus \mathcal{Z} = \mathcal{V} \boxplus \mathcal{Y}$ and hence $(\mathcal{U}, \mathcal{V}) \equiv (\mathcal{Y}, \mathcal{Z})$. The elements of the group $\mathbb{G}$ are the equivalence classes for this relation. We write $\boxplus$ for the binary operation on $\mathbb{G}$ and define it to be the operation that takes the equivalence classes of $(\mathcal{W}, \mathcal{X})$ and $(\mathcal{Y}, \mathcal{Z})$ to the equivalence class of $(\mathcal{W} \boxplus \mathcal{Y}, \mathcal{X} \boxplus \mathcal{Z})$. It is clear that this operation is well-defined, associative and commutative. The identity element is the equivalence class of $(\mathcal{E}, \mathcal{E})$ and the inverse of the equivalence class of $(\mathcal{Y}, \mathcal{Z})$ is the equivalence class of $(\mathcal{Z}, \mathcal{Y})$.

It will be convenient for us to have various ways of measuring how far a metric measure space $\mathcal{X}$ is from the trivial space $\mathcal{E}$. The most obvious such measure is simply the Gromov–Prohorov distance $d_{\mathrm{GPr}}(\mathcal{X}, \mathcal{E})$. Note from Lemma 2.1 that $d_{\mathrm{GPr}}(\mathcal{X}_1 \boxplus \mathcal{X}_2, \mathcal{E}) \leqslant d_{\mathrm{GPr}}(\mathcal{X}_1, \mathcal{E}) + d_{\mathrm{GPr}}(\mathcal{X}_2, \mathcal{E})$ for $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{M}$. It follows from Lemma 3.5 that a sequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$ is such that $d_{\mathrm{GPr}}(\mathcal{X}_n, \mathcal{E}) \to 0$ if and only if $\chi_A(\mathcal{X}_n) \to 1$ for all $A \in \mathbb{A}$ and so $D_A(\mathcal{X}) := -\log \chi_A(\mathcal{X})$ is also a measure of how far $\mathcal{X}$ is from $\mathcal{E}$. Observe that $D_A(\mathcal{X}_1 \boxplus \mathcal{X}_2) = D_A(\mathcal{X}_1) + D_A(\mathcal{X}_2)$ for $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{M}$. To

simplify notation, we set

$$D(\mathcal{X}) := -\log \chi_1(\mathcal{X}). \tag{3.3}$$

It is a consequence of Lemma 3.8 below that $d_{\mathrm{GPr}}(\mathcal{X}_n, \mathcal{E}) \to 0$ if and only if $D(\mathcal{X}_n) \to 0$.

The equivalence between convergence in the Gromov–Prohorov distance and convergence in distribution of the corresponding random distance matrices implies that if we set

$$R(\mathcal{X}) := \int_{X^2} (r_X(x_1, x_2) \wedge 1)\, \mu_X^{\otimes 2}(dx), \tag{3.4}$$

then $d_{\mathrm{GPr}}(\mathcal{X}_n, \mathcal{E}) \to 0$ if and only if $R(\mathcal{X}_n) \to 0$. It is clear that $R(\mathcal{X}_1 \boxplus \mathcal{X}_2) \leqslant R(\mathcal{X}_1) + R(\mathcal{X}_2)$ for $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{M}$. One last quantity that is useful for measuring how far bounded metric measure spaces are from $\mathcal{E}$ is the diameter. Recall from Lemma 2.4 that $\mathrm{diam}(\mathcal{X}_1 \boxplus \mathcal{X}_2) = \mathrm{diam}(\mathcal{X}_1) + \mathrm{diam}(\mathcal{X}_2)$ for $\mathcal{X}_1, \mathcal{X}_2 \in \mathbb{M}$. The following result establishes a number of relationships between these various objects.

Lemma 3.8. a) *For each $A \in \mathbb{A}$, there exist constants $a \geqslant b > 0$ such that, for all $\mathcal{X} \in \mathbb{M}$,*

$$\chi_1(\mathcal{X})^a \leqslant \chi_A(\mathcal{X}) \leqslant \chi_1(\mathcal{X})^b$$

and hence

$$bD(\mathcal{X}) \leqslant D_A(\mathcal{X}) \leqslant aD(\mathcal{X}).$$

b) *For each $\mathcal{X} \in \mathbb{M}$,*

$$\frac{1}{4} R(\mathcal{X}) \leqslant d_{\mathrm{GPr}}(\mathcal{X}, \mathcal{E}) \leqslant \sqrt{R(\mathcal{X})}.$$

c) *There exist constants $C > c > 0$ such that for each $\mathcal{X} \in \mathbb{M}$*

$$c(D(\mathcal{X}) \wedge 1) \leqslant R(\mathcal{X}) \leqslant C(D(\mathcal{X}) \wedge 1).$$

d) *For each $\mathcal{X} \in \mathbb{M}$, $d_{\mathrm{GPr}}(\mathcal{X}, \mathcal{E}) \leqslant \mathrm{diam}(\mathcal{X})$, $D(\mathcal{X}) \leqslant \mathrm{diam}(\mathcal{X})$ and $R(\mathcal{X}) \leqslant \mathrm{diam}(\mathcal{X})$.*

Proof. Consider the the first inequality in part (a) for $A \in \mathbb{A} \cap \mathbb{R}^{\binom{n}{2}}$. The triangle inequality yields that

$$\sum_{1 \leqslant i < j \leqslant n} a_{ij} r_X(x_i, x_j) \leqslant c \sum_{i=2}^{n} r_X(x_1, x_i)$$

for a certain constant c. Therefore,

$$\chi_A(\mathcal{X}) \geqslant \int_X \left(\int_X \exp(-c r_X(x, y)) \mu_X(dy) \right)^{n-1} \mu_X(dx)$$

$$\geqslant \left(\int_{X^2} \exp(-c r_X(x,y))\, \mu_X(dx)\, \mu_X(dy) \right)^{n-1}$$
$$\geqslant (\chi_1(\mathcal{X}))^{(c \vee 1)(n-1)},$$

where the two last inequalities follow from Jensen's inequality.

Regarding the second inequality in part (a), there exist $1 \leqslant i' < j' \leqslant n$ such that $0 < a_{i'j'} =: \alpha$. Because $\sum_{1 \leqslant i < j \leqslant n} a_{ij} r_X(x_i, x_j) \geqslant \alpha r_X(x_{i'}, x_{j'})$, we have $\chi_A(\mathcal{X}) \leqslant \chi_\alpha(\mathcal{X})$. If $\alpha \geqslant 1$, then $\chi_\alpha(\mathcal{X}) \leqslant \chi_1(\mathcal{X})$, whereas if $\alpha < 1$, then $\chi_\alpha(\mathcal{X}) \leqslant \chi_1(\mathcal{X})^\alpha$ by Jensen's inequality. Therefore, $\chi_A(\mathcal{X}) \leqslant \chi_\alpha(\mathcal{X}) \leqslant \chi_1(\mathcal{X})^{\alpha \wedge 1}$.

For the first inequality in (b), we begin by recalling (12.1) which says that

$$d_{\mathrm{GPr}}(\mathcal{X}, \mathcal{E}) = \inf_{x \in X} \inf\{\varepsilon > 0 : \ \mu_X\{y \in X : \ r_X(x,y) \geqslant \varepsilon\} \leqslant \varepsilon\}.$$

Suppose that $d_{\mathrm{GPr}}(\mathcal{X}, \mathcal{E}) < \gamma$ where $0 < \gamma \leqslant 1$. There is then an $x \in X$ such that $\mu_X\{y \in X : \ r_X(x,y) \geqslant \gamma\} \leqslant \gamma$. Hence, by the triangle inequality

$$\begin{aligned}
R(\mathcal{X}) &= \int_{X^2} (r_X(y_1, y_2) \wedge 1)\, \mu_X^{\otimes 2}(dy) \\
&\leqslant \int_{X^2} ([r_X(x, y_1) + r_X(x, y_2)] \wedge 1)\, \mu_X^{\otimes 2}(dy) \\
&\leqslant 2 \int_X (r_X(x,y) \wedge 1)\, \mu_X(dy) \\
&\leqslant 2 \left[\gamma \mu_X\{y \in X : \ r_X(x,y) < \gamma\} + \mu_X\{y \in X : \ r_X(x,y) \geqslant \gamma\} \right] \\
&\leqslant 4\gamma,
\end{aligned}$$

and the inequality follows.

Turning to the second inequality in part (b), suppose that $R(\mathcal{X}) < \gamma$ where $0 < \gamma \leqslant 1$. There must then be an $x \in \mathcal{X}$ for which $\int_X (r_X(x,y) \wedge 1)\, \mu_X(dy) < \gamma$ and hence $\varepsilon \mu_X\{y \in X : \ r_X(x,y) \geqslant \varepsilon\} < \gamma$ for $0 < \varepsilon \leqslant 1$. Take $\varepsilon = \sqrt{\gamma}$ to see that $\mu_X\{y \in X : \ r_X(x,y) \geqslant \sqrt{\gamma}\} < \sqrt{\gamma}$, as required.

Part (c) is immediate from Lemma 13.1. and part (d) is obvious. □

Proposition 3.9. a) *The sequence* $(\boxplus_{k=0}^n \mathcal{X}_k)_{n \in \mathbb{N}}$ *converges in* $\mathbb{M}$ *if and only if* $\lim_{m,n \to \infty,\, m<n} \boxplus_{k=m+1}^n X_k = \mathcal{E}$.

b) *The sequence* $(\boxplus_{k=0}^n \mathcal{X}_k)_{n \in \mathbb{N}}$ *converges in* $\mathbb{M}$ *if and only if* $\sum_{k \in \mathbb{N}} D(\mathcal{X}_k) < \infty$ *or, equivalently, if and only if* $\sum_{k \in \mathbb{N}} R(\mathcal{X}_k) < \infty$.

c) *The sequence* $(\boxplus_{k=0}^n \mathcal{X}_k)_{n\in\mathbb{N}}$ *converges in* $\mathbb{M}$ *if and only if there exists* $\mathcal{Z} \in \mathbb{M}$ *such that* $\boxplus_{k=0}^n \mathcal{X}_k \leqslant \mathcal{Z}$ *for all* $n \in \mathbb{N}$*, in which case* $\lim_{n\to\infty} \boxplus_{k=0}^n \mathcal{X}_k \leqslant \mathcal{Z}$*.*

d) *Suppose that* $(\mathcal{Y}_n)_{n\in\mathbb{N}}$ *is a sequence in* $\mathbb{M}$ *such that* $\mathcal{Y}_0 \geqslant \mathcal{Y}_1 \geqslant \cdots$*. Then,* $\lim_{n\to\infty} \mathcal{Y}_n$ *exists.*

e) *Suppose that* $(\mathcal{X}_n)_{n\in\mathbb{N}}$ *is a sequence such that* $\lim_{n\to\infty} \boxplus_{k=0}^n \mathcal{X}_k = \mathcal{Y}$ *for some* $\mathcal{Y} \in \mathbb{M}$*. Suppose further that* $(\mathcal{X}'_n)_{n\in\mathbb{N}}$ *is a sequence that is obtained by re-ordering the sequence* $(\mathcal{X}_n)_{n\in\mathbb{N}}$*. Then,* $\lim_{n\to\infty} \boxplus_{k=0}^n \mathcal{X}'_k = \mathcal{Y}$ *also.*

f) *The sequence* $(\boxplus_{k=0}^n \mathcal{X}_k)_{n\in\mathbb{N}}$ *converges to a bounded metric measure space if and only if* $\sum_{n\in\mathbb{N}} \operatorname{diam}(\mathcal{X}_n) < \infty$*.*

g) *A sufficient condition for the sequence* $(\boxplus_{k=0}^n \mathcal{X}_k)_{n\in\mathbb{N}}$ *to converge in* $\mathbb{M}$ *is that* $\sum_{n\in\mathbb{N}} d_{\mathrm{GPr}}(\mathcal{X}_n, \mathcal{E}) < \infty$ *and a necessary condition is that* $\lim_{n\to\infty} d_{\mathrm{GPr}}(\mathcal{X}_n, \mathcal{E}) = 0$*.*

Proof. (a) By the completeness of $(\mathbb{M}, d_{\mathrm{GPr}})$, the convergence of $\boxplus_{k=0}^n \mathcal{X}_k$ as $n \to \infty$ is equivalent to

$$\lim_{m,n\to\infty} d_{\mathrm{GPr}}\left(\boxplus_{k=0}^m \mathcal{X}_k, \boxplus_{k=0}^n \mathcal{X}_k\right) = 0.$$

However, if $m < n$, then Lemma 2.2 gives

$$d_{\mathrm{GPr}}\left(\boxplus_{k=0}^m \mathcal{X}_k, \boxplus_{k=0}^n \mathcal{X}_k\right) = d_{\mathrm{GPr}}\left(\boxplus_{k=m+1}^n \mathcal{X}_k, \mathcal{E}\right).$$

(b) It suffices to prove the claim for D because the claim for R will then follow from Lemma 3.8(c).

Suppose that $\sum_{k\in\mathbb{N}} D(\mathcal{X}_k) < \infty$. For $m < n$,

$$\begin{aligned} d_{\mathrm{GPr}}\left(\boxplus_{k=m+1}^n \mathcal{X}_k, \mathcal{E}\right) &\leqslant \sqrt{CD\left(\boxplus_{k=m+1}^n \mathcal{X}_k\right)} \\ &= \sqrt{C\left(\sum_{k=m+1}^n D(\mathcal{X}_k)\right)} \end{aligned}$$

for some constant C by parts (b) and (c) of Lemma 3.8. It is then a consequence of part (a) that $\boxplus_{k=0}^n \mathcal{X}_k$ converges as $n \to \infty$.

Conversely, if $\lim_{n\to\infty} \boxplus_{k=0}^n \mathcal{X}_k = \mathcal{Y}$ exists, then $\sum_{k=0}^n D(\mathcal{X}_k) = D(\boxplus_{k=0}^n \mathcal{X}_k) \to D(\mathcal{Y})$ by Lemma 3.5.

(c) Suppose that $\boxplus_{k=0}^n \mathcal{X}_k \leqslant \mathcal{Z}$ for all $n \in \mathbb{N}$. It follows from Lemma 3.2(b) that $\sum_{k=0}^n D(X_k) = D(\boxplus_{k=0}^n \mathcal{X}_k) \leqslant D(\mathcal{Z})$ for all $n \in \mathbb{N}$, and so part (b) gives that $\boxplus_{k=0}^n \mathcal{X}_k$ converges as $n \to \infty$. We note that

an alternative proof of this direction can be given along the lines of the proof of part (d).

Conversely, suppose that $\lim_{n\to\infty} \boxplus_{k=0}^{n} \mathcal{X}_k =: \mathcal{Y}$ exists. We know from one direction of part (b) that $\sum_{k\in\mathbb{N}} D(\mathcal{X}_k) < \infty$ so that $\sum_{k=m+1} D(\mathcal{X}_k) < \infty$ and hence, by the other direction of part (b), $\lim_{n\to\infty} \boxplus_{k=m+1}^{n} \mathcal{X}_k =: \mathcal{Y}_m$ exists for all $m \in \mathbb{N}$. We have $\boxplus_{k=0}^{m} \mathcal{X}_k \boxplus \mathcal{Y}_m = \mathcal{Y}$ for all $m \in \mathbb{N}$ and hence $\boxplus_{k=0}^{m} \mathcal{X}_k \preccurlyeq \mathcal{Y}$ for all $m \in \mathbb{N}$. We note that Proposition 3.6(b) can be used to give an alternative proof of this direction.

(d) By Lemma 2.10(a) any subsequence of $(\mathcal{Y}_n)_{n\in\mathbb{N}}$ has a further subsequence that converges. For any $A \in \mathbb{A}$, the sequence $(\chi_A(\mathcal{Y}_n))_{n\in\mathbb{N}}$ is nondecreasing by Lemma 3.2(b) and hence convergent. By Lemma 3.5, all of the convergent subsequences produced in this manner converge to the same limit, and so the sequence $(\mathcal{Y}_n)_{n\in\mathbb{N}}$ itself converges to that limit.

(e) It follows from Lemma 3.5 that $\sum_{n\in\mathbb{N}} D_A(\mathcal{X}_n) = D_A(\mathcal{Y})$. It is well-known that all rearrangements of a convergent sequence with nonnegative terms converge to the same limit. Thus, $\sum_{n\in\mathbb{N}} D_A(\mathcal{X}'_n) = \sum_{n\in\mathbb{N}} D_A(\mathcal{X}_n) = D_A(\mathcal{Y})$, implying that $\lim_{n\to\infty} \chi_A(\boxplus_{k=0}^{n} \mathcal{X}'_k) = \chi_A(\mathcal{Y})$ and hence, by Lemma 3.5, that $\lim_{n\to\infty} \boxplus_{k=0}^{n} \mathcal{X}'_k = \mathcal{Y}$.

(f) Suppose that $\lim_{n\to\infty} \boxplus_{k=0}^{n} \mathcal{X}_k = \mathcal{Y}$, where $\mathcal{Y}$ is bounded. Since $\boxplus_{k=0}^{n} \mathcal{X}_k \preccurlyeq \mathcal{Y}$, $\sum_{k=0}^{n} \operatorname{diam}(\mathcal{X}_k) = \operatorname{diam}(\boxplus_{k=0}^{n} \mathcal{X}_k) \leqslant \operatorname{diam}(\mathcal{Y})$, and so $\sum_{n\in\mathbb{N}} \operatorname{diam}(\mathcal{X}_n) < \infty$.

Conversely, suppose that $\sum_{n\in\mathbb{N}} \operatorname{diam}(\mathcal{X}_n) < \infty$. It follows from Lemma 3.8(d) that $\sum_{n\in\mathbb{N}} D(\mathcal{X}_n) < \infty$ and hence $\boxplus_{k=0}^{n} \mathcal{X}_k$ converges to $\mathcal{Y} \in \mathbb{M}$ as $n \to \infty$.

The diameter is lower semicontinuous by Lemma 2.6 and so $\operatorname{diam}(\mathcal{Y}) \leqslant \liminf \sum_{k=0}^{n} \operatorname{diam}(\mathcal{X}_k) < \infty$.

(g) This part is immediate from part (a) and the observation that $d_{\mathrm{GPr}}(X_n, \mathcal{E}) \leqslant d_{\mathrm{GPr}}(\boxplus_{k=m+1}^{n}, \mathcal{E}) \leqslant \sum_{k=m+1}^{n} d_{\mathrm{GPr}}(X_k, \mathcal{E})$ by Lemma 2.8 and Lemma 2.1. Alternatively, the result follows from part (b) and Lemma 3.8(b). □

Remark 3.10. Proposition 3.9(e) gives that if $(\mathcal{X}_s)_{s\in S}$ is a countable collection of elements of $\mathbb{M}$, then the existence of $\lim_{n\to\infty} \boxplus_{k=0}^{n} \mathcal{X}_{s_k}$ for some listing $(s_n)_{n\in\mathbb{N}}$ implies the existence for any other listing, with the same value for the limit. We will therefore unambiguously denote the limit when it exists by the notation $\boxplus_{s\in S} \mathcal{X}_s$. Moreover, a necessary and sufficient condition for $\boxplus_{s\in S} \mathcal{X}_s$ to exist is that $\sum_{s\in S} D(\mathcal{X}_s) < \infty$.

We finish this section with a technical result that will be used to handle certain measurability issues in Section 6. We use the notation

$\mathcal{Y}^{\boxplus n}$ for $\mathcal{Y} \in \mathbb{M}$ and $n \in \mathbb{N}$ to denote $\mathcal{Y} \boxplus \cdots \boxplus \mathcal{Y}$, where there are n terms and we adopt the convention that this quantity is $\mathcal{E}$ for $n = 0$.

Corollary 3.11. a) *For all $n \in \mathbb{N}$, the set $\{(\mathcal{X}, \mathcal{Y}) \in \mathbb{M}^2 : \mathcal{Y}^{\boxplus n} \leqslant \mathcal{X}\}$ is closed.*

b) *The function $M : \mathbb{M}^2 \to \mathbb{N}$ defined by $M(\mathcal{X}, \mathcal{Y}) = \max\{n \in \mathbb{N} : \mathcal{Y}^{\boxplus n} \leqslant \mathcal{X}\}$ is upper semicontinuous and hence Borel.*

Proof. Part (a) is immediate from Proposition 3.6(b) for $n = 1$. If $n \geqslant 2$, let $\mathcal{Y}_k^{\boxplus n} \boxplus \mathcal{W}_k = \mathcal{X}_k$ for all k. If $\mathcal{X}_k \to \mathcal{X}$ and $\mathcal{Y}_k \to \mathcal{Y}$, then $\mathcal{Y}_k^{\boxplus n} \to \mathcal{Y}^{\boxplus n}$ and the statement again follows from Proposition 3.6(b).

For part (b), $\{(\mathcal{X}, \mathcal{Y}) \in \mathbb{M}^2 : M(\mathcal{X}, \mathcal{Y}) \geqslant n\} = \{(\mathcal{X}, \mathcal{Y}) \in \mathbb{M}^2 : \mathcal{Y}^{\boxplus n} \leqslant \mathcal{X}\}$ is a closed set for all $n \in \mathbb{N}$ by part (a), and this is equivalent to the upper semicontinuity of M. □

4. Irreducibility and infinite divisibility

Definition 4.1. An element $\mathcal{X} \in \mathbb{M}$ is *irreducible* if $\mathcal{X} \neq \mathcal{E}$ and $\mathcal{Y} \leqslant \mathcal{X}$ for $\mathcal{Y} \in \mathbb{M}$ implies that $\mathcal{Y}$ is either $\mathcal{E}$ or $\mathcal{X}$ (see [Cli38, Section 1]). We write $\mathbb{I}$ for the set of irreducible elements of $\mathbb{M}$.

It is not clear *a priori* that $\mathbb{I}$ is nonempty. For example, the semigroup $\mathbb{R}_+$ with the usual addition operation has no irreducible elements in the sense of the general definition in [Cli38]. The following two results show that $\mathbb{I}$ is certainly nonempty.

Proposition 4.2. *The sets $\mathbb{I}$ and $\mathbb{M}\backslash\mathbb{I}$ are dense subsets of $\mathbb{M}$. Moreover, the set $\mathbb{I}$ is a G_δ subset of $\mathbb{M}$.*

Proof. It is easy to see that $\mathbb{M}\backslash\mathbb{I}$ is a dense subset of $\mathbb{M}$: for any $\mathcal{X} \in \mathbb{M}$ and $\mathcal{Z} \in \mathbb{M}\backslash\{\mathcal{E}\}$ the elements $\mathcal{X}_n := \mathcal{X} \boxplus (\frac{1}{n}\mathcal{Z})$ belong to $\mathbb{M}\backslash\mathbb{I}$ and converge to $\mathcal{X}$ as $n \to \infty$.

We next show that $\mathbb{I}$ is dense in $\mathbb{M}$. As in the proof of [GPW09, Proposition 5.6], the subset of $\mathbb{F} \subset \mathbb{M}$ consisting of compact metric measure spaces with finitely many points is dense in $\mathbb{M}$. If we are given a finite metric measure space (W, r_W, μ_W), then convergence of a sequence of probability measures in the Prohorov metric on (W, r_W) is just pointwise convergence of the probabilities assigned to each point of W. The set of probability measures that assign positive probability to all points of W is thus just the relative interior of the $(\#W - 1)$-dimensional simplex thought of as a subset of $\mathbb{R}^{\#W}$ equipped with the usual Euclidean topology. Suppose that (W, r_W) is isometric to $(U \times V, r_U \oplus r_V)$ for some nontrivial finite compact metric spaces (U, r_U) and (V, r_V) – if this is not the case, then (W, r_W, μ_w) is already irreducible. The probability measures on $U \times V$ that are of the form $\mu_U \otimes \mu_V$

form a $(\#U-1)+(\#V-1)$-dimensional surface in the $(\#U\times\#V-1)$-dimensional simplex of probability measures on $U\times V$ and, in particular, the former set is nowhere dense. Thus, even if (W,r_W) is isometric to $(U\times V, r_U\oplus r_V)$, any probability measure on W that is the isometric image of a probability measure on $U\times V$ of the form $\mu_U\otimes\mu_V$ is arbitrarily close to probability measures on W that are not isometric images of probability measures of this form, and it follows that $\mathbb{I}$ is dense in $\mathbb{M}$.

We now show that the set $\mathbb{I}$ is a G_δ. This is equivalent to showing that $\mathbb{M}\backslash\mathbb{I}$ is an F_σ.

Let χ_1 be the semicharacter defined by (3.2). Recall that $\chi_1(\mathcal{X})=1$ if and only if $\mathcal{X}=\mathcal{E}$. For $0<\varepsilon<\frac{1}{2}$ set

$$\mathbb{L}_\varepsilon := \{\mathcal{X}\in\mathbb{M} : \exists\mathcal{Y}\leqslant\mathcal{X},\ \chi_1(\mathcal{X})^{1-\varepsilon}\leqslant\chi_1(\mathcal{Y})\leqslant\chi_1(\mathcal{X})^{\varepsilon}\}.$$

Note that $\mathbb{L}_{\varepsilon'}\supseteq\mathbb{L}_{\varepsilon''}$ for $\varepsilon'\leqslant\varepsilon''$ and $\bigcup_{0<\varepsilon<\frac{1}{2}}\mathbb{L}_\varepsilon=\mathbb{M}\backslash\mathbb{I}$, so it suffices to show that the $\mathbb{L}_\varepsilon$ are closed. Suppose that $(\mathcal{X}_n)_{n\in\mathbb{N}}$ is a sequence of elements of $\mathbb{L}_\varepsilon$ that converges to $\mathcal{X}\in\mathbb{M}$. For each $n\in\mathbb{N}$ there exist $\mathcal{Y}_n$ and $\mathcal{Z}_n$ in $\mathbb{M}$ such that $\mathcal{X}_n=\mathcal{Y}_n\boxplus\mathcal{Z}_n$ and $\chi_1(\mathcal{X}_n)^{1-\varepsilon}\leqslant\chi_1(\mathcal{Y}_n)\leqslant\chi_1(\mathcal{X}_n)^{\varepsilon}$. By Lemma 2.10(a) and Proposition 3.6(b), there is a subsequence $(n_k)_{k\in\mathbb{N}}$ such that $\lim_{k\to\infty}\mathcal{Y}_{n_k}=\mathcal{Y}$ and $\lim_{k\to\infty}\mathcal{Z}_{n_k}=\mathcal{Z}$ for $\mathcal{Y},\mathcal{Z}\in\mathbb{M}$ such that $\mathcal{X}=\mathcal{Y}\boxplus\mathcal{Z}$. Thus, $\mathcal{Y}\leqslant\mathcal{X}$ and $\chi_1(\mathcal{X})^{1-\varepsilon}\leqslant\chi_1(\mathcal{Y})\leqslant\chi_1(\mathcal{X})^{\varepsilon}$, so that $\mathcal{X}\in\mathbb{L}_\varepsilon$, as required. □

A theorem of Alexandrov, see [Kec95, Theorem 3.11], says that a subspace of a Polish space is Polish in the relative topology if and only if it is a G_δ-set; therefore, the space $\mathbb{I}$ with the relative topology inherited from $\mathbb{M}$ is Polish.

Remark 4.3. It is not difficult to construct concrete examples of irreducible elements of $\mathbb{M}$.

We first recall that a metric space (W,r_W) is *totally geodesic* if for any pair of points $w',w''\in W$ there is a unique map $\phi:[0,r_W(w',w'')]\to W$ such that $\phi(0)=w'$, $\phi(r_W(w',w''))=w''$ and $r_W(\phi(s),\phi(t))=|s-t|$ for $s,t\in[0,r_W(w',w'')]$; that is, any two points of W are joined by a unique geodesic segment.

Any nontrivial closed subset X of a totally geodesic, complete, separable metric space W is irreducible no matter what measure it is equipped with because such a space (X,r_W) cannot be isometric to a space of the form $(Y\times Z, r_Y\oplus r_Z)$ for nontrivial Y and Z. To see this, suppose that the claim is false. There will then be four distinct points a,b,c,d in X that are isometric images of points of the form (y',z'), (y'',z'), (y',z''), (y'',z'') in $Y\times Z$. Suppose that (X,r_W) is a closed subset of the totally geodesic, complete, separable metric space

(W, r_W). We have

$$\begin{aligned} r_W(a,b) &= r_W(c,d), \\ r_W(a,c) &= r_W(b,d), \\ r_W(a,d) &= r_W(a,b) + r_W(b,d), \\ r_W(a,d) &= r_W(a,c) + r_W(c,d), \\ r_W(b,c) &= r_W(a,b) + r_W(c,a), \end{aligned}$$

and

$$r_W(b,c) = r_W(b,d) + r_W(c,d).$$

It follows from the third and fourth equalities that b and c are on the geodesic segment between a and d. We may therefore suppose that (W, r_W) is a closed subinterval of $\mathbb{R}$ and, without loss of generality, that $a < b < c < d$. The fifth and sixth equalities are then impossible.

There are many totally geodesic, complete, separable metric spaces. A Banach space $(X, \|\,\|)$ is totally geodesic if and only if it is *strictly convex*; that is, $x \neq y$ and $\|x'\| = \|x''\| = 1$ imply that $\|ax' + (1-a)x''\| < 1$ for all $0 < a < 1$ [Bea85, Section 3.I.1]. Strict convexity of $(X, \|\,\|)$ is implied by *uniform convexity*; that is, for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $\|x'\| = \|x''\| = 1$ and $\|x' - x''\| \geqslant \varepsilon$ imply $\|\frac{x'+x''}{2}\| \leqslant 1 - \delta$. Any Hilbert space is uniformly convex and the Banach spaces $L^p(S, \mathcal{S}, \lambda)$, $1 < p < \infty$, where λ is a σ-finite measure, are uniformly convex [Bea85, Section 3.II.1]. Also, any real tree is, by definition, totally geodesic and any ultrametric space is isometric to a subset of a real tree.

Definition 4.4. An element of a semigroup is said to be *infinitely divisible* if, for each, $n \geqslant 2$, it can be represented as the sum of n identical summands.

Proposition 4.5. *There are no nontrivial infinitely divisible metric measure spaces.*

Proof. Suppose that $\mathcal{X} = (X, r_X, \mu_X)$ is a nontrivial infinitely divisible metric measure space. Thus, for every $n \in \mathbb{N}$ we have $\mathcal{X} = \mathcal{X}_n^{\boxplus 2^n}$ for some metric measure space $\mathcal{X}_n = (X_n, r_{X_n}, \mu_{X_n})$. We may suppose that $X_0 = X$, $r_{X_0} = r_X$ and $\mu_{X_0} = \mu_X$, and that for all $n \in \mathbb{N}$ there is an isometry $\phi_{n,n+1}$ from X_n equipped with r_{X_n} to X_{n+1}^2 equipped with $r_{X_{n+1}} \oplus r_{X_{n+1}}$ such that the push-forward of μ_{X_n} by $\phi_{n,n+1}$ is $\mu_{X_{n+1}} \otimes \mu_{X_{n+1}}$. Let ξ_i, $i \in \mathbb{N}$, be independent identically distributed random elements of X with common distribution μ_X. Define $(\xi_{ni1}, \ldots, \xi_{ni2^n})$, $n \in \mathbb{N}$, $i \in \mathbb{N}$, recursively by $\xi_{0i1} = \xi_i$ and $(\xi_{n+1,i,2k-1}, \xi_{n+1,i,2k}) = \phi_{n,n+1}(\xi_{nik})$ for $k \in \{1, \ldots, 2^n\}$. The ξ_{nik}, $i \in \mathbb{N}$, $k \in \{1, \ldots, 2^n\}$ are random elements of X_n with distribution μ_{X_n},

$r_{X_n}(\xi_{nik},\xi_{njk}) = r_{X_{n+1}}(\xi_{n+1,i,2k-1},\xi_{n+1,j,2k-1}) + r_{X_{n+1}}(\xi_{n+1,i,2k},\xi_{n+1,j,2k})$, and consequently $r_X(\xi_i,\xi_j) = \sum_{k=1}^{2^n} r_{X_n}(\xi_{nik},\xi_{njk})$.

For $i \neq j$ the nonnegative random variable $r_X(\xi_i,\xi_j)$ is clearly infinitely divisible. These random variables are not almost surely zero and they are identically distributed. Their common distribution does not have a nontrivial deterministic component because that would mean that for some $c > 0$ we would have $r_X(\xi_i,\xi_j) \geqslant c$ for all $i \neq j$, which is impossible because almost surely for all $i \in \mathbb{N}$ we must have $\inf_{j\in\mathbb{N},\, j\neq i} r_X(\xi_i,\xi_j) = 0$ if $(\xi_h)_{h\in\mathbb{N}}$ is an independent identically distributed sequence of random elements of X with common distribution μ_X. In particular, these random variables are not bounded, because a bounded infinitely divisible random variable is almost surely constant. It follows that the metric r_X is unbounded.

Let ν be the Lévy measure associated with the common infinitely divisible distribution of $r_X(\xi_i,\xi_j)$ for $i \neq j$. This is a (nontrivial) measure on $\mathbb{R}_{++} := (0,\infty)$ that satisfies $\int_{\mathbb{R}_{++}} (x \wedge 1)\, \nu(dx) < \infty$ and it is the limit as $n \to \infty$ of the measures

$$\sum_{k=1}^{2^n} \mathbb{P}\{r_{X_n}(\xi_{nik},\xi_{njk}) \in \cdot\} = 2^n \int_{X_n^2} \mathbb{1}\{r_{X_n}(y,z) \in \cdot\}\, \mu_{X_n}^{\otimes 2}(dy,dz),$$

where the limit is in the sense of vague convergence of measures on $\mathbb{R}_{++}$.

For $K > 0$, set

$$R_n^K(i,j) := \sum_{k=1}^{2^n} (r_{X_n}(\xi_{nik},\xi_{njk}) \wedge K).$$

As $n \to \infty$, $R_n^K(i,j)$ converges almost surely to an infinitely divisible random variable $R^K(i,j)$ with

$$\mathbb{E}[R^K(i,j)] = \int_{\mathbb{R}_{++}} (x \wedge K)\, \nu(dx) < \infty,$$

and $R^K(i,j) = r_X(\xi_i,\xi_j)$ for all K sufficiently large almost surely. The random matrix $(r_X(\xi_i,\xi_j))_{i,j\in\mathbb{N}}$ satisfies the necessary and sufficient condition (1.1) to be the matrix of pairwise distances for a sample from a metric measure space, and it follows easily that the same is true of the random matrix $(R^K(i,j))_{i,j\in\mathbb{N}}$. Because the random matrix $(R^K(i,j))_{i,j\in\mathbb{N}}$ is infinitely divisible, the underlying metric measure space that gives rise to this matrix of pairwise distances is also infinitely divisible. We may therefore suppose without loss of generality that the random variables $r_X(\xi_i,\xi_j)$ are integrable.

It is clear from Fubini's theorem that

$$\mathbb{E}[r_X(y,\xi_j)] = \int_X r_X(y,z)\,\mu_X(dz) < \infty, \quad \mu_X\text{-a.e. } y \in X.$$

Because r_X is unbounded, the function $y \mapsto \int_X r_X(y,z)\,\mu_X(dz)$ is also unbounded, and since μ_X has full support, each of the random variables $\mathbb{E}[r_X(\xi_i,\xi_j)\,|\,\xi_i]$, $i \neq j$, are unbounded. These random variables are equal for a fixed i as j varies and as i varies the common values are independent and identically distributed. Moreover,

$$\mathbb{E}[r_X(\xi_i,\xi_j)\,|\,\xi_i] = \sum_{k=1}^{2^n} \int_{X_n} r_{X_n}(\xi_{nik},z)\,\mu_{X_n}(dz)$$

for all $n \in \mathbb{N}$, and so $\mathbb{E}[r_X(\xi_i,\xi_j)\,|\,\xi_i]$ is infinitely divisible and, being unbounded, this random variable cannot be constant almost surely.

Given $\varepsilon > 0$, set

$$I^\varepsilon_{nik} = \mathbb{1}\left\{\int_{X_n} r_{X_n}(\xi_{nik},z)\,\mu_{X_n}(dz) > \varepsilon\right\}.$$

For ε sufficiently small, $\sum_{k=1}^{2^n} I^\varepsilon_{nik}$ converges almost surely as $n \to \infty$ to a nontrivial random variable J^ε_i that has a Poisson distribution. Moreover, for $\varepsilon', \varepsilon'' > 0$ and $i \neq j$, $\sum_{k=1}^{2^n} I^{\varepsilon'}_{nik} I^{\varepsilon''}_{njk} = 0$ for all n sufficiently large almost surely by the independence of $\{\xi_{nik} : n \in \mathbb{N},\, 1 \leqslant k \leqslant 2^n\}$ and $\{\xi_{njk} : n \in \mathbb{N},\, 1 \leqslant k \leqslant 2^n\}$.

By the triangle inequality,

$$\begin{aligned}\int_{X_n} r_{X_n}(y'',z)\,\mu_{X_n}(dz) &\geqslant \int_{X_n} \left[r_{X_n}(y',z) - r_{X_n}(y',y'')\right]\,\mu_{X_n}(dz)\\ &= \int_{X_n} r_{X_n}(y',z)\,\mu_{X_n}(dz) - r_{X_n}(y',y'')\end{aligned}$$

and hence

$$r_{X_n}(y',y'') \geqslant \int_{X_n} r_{X_n}(y',z)\,\mu_{X_n}(dz) - \int_{X_n} r_{X_n}(y'',z)\,\mu_{X_n}(dz).$$

Therefore, if $\int_{X_n} r_{X_n}(y',z)\,\mu_{X_n}(dz) > \varepsilon'$ and $\int_{X_n} r_{X_n}(y'',z)\,\mu_{X_n}(dz) \leqslant \varepsilon''$ for $\varepsilon' > \varepsilon'' > 0$, then $r_{X_n}(y',y'') > \varepsilon' - \varepsilon''$. Thus,

$$\begin{aligned} r_X(\xi_i,\xi_j) &= \sum_{k=1}^{2^n} r_{X_n}(\xi_{nik},\xi_{njk}) \\ &\geqslant \sum_{k=1}^{2^n} I_{nik}^{\varepsilon'}(1-I_{njk}^{\varepsilon''}) r_{X_n}(\xi_{nik},\xi_{njk}) \\ &\geqslant \sum_{k=1}^{2^n} I_{nik}^{\varepsilon'}(1-I_{njk}^{\varepsilon''})(\varepsilon'-\varepsilon'') \end{aligned}$$

and so on the event $\{\sum_{k=1}^{2^n} I_{nik}^{\varepsilon'} I_{njk}^{\varepsilon''} = 0\}$

$$r_X(\xi_i,\xi_j) \geqslant (\varepsilon'-\varepsilon'') \sum_{k=1}^{2^n} I_{nik}^{\varepsilon'}.$$

Consequently,

$$r_X(\xi_i,\xi_j) \geqslant \varepsilon J_i^{\varepsilon}$$

for all $i \neq j$ almost surely. This, however, is impossible because if $(\xi_h)_{h\in\mathbb{N}}$ is an independent identically distributed sequence of random elements of X with common distribution μ_X, then almost surely for all $i \in \mathbb{N}$ we must have $\inf_{j\in\mathbb{N},\, j\neq i} r_X(\xi_i,\xi_j) = 0$. □

Remark 4.6. In the case of bounded metric measure spaces, a simpler and more direct proof of Proposition 4.5 is to note that if $\mathcal{X} = \mathcal{X}_n^{\boxplus n}$ for all n, then the push-forward of the probability measure $\mu_X^{\otimes 2}$ by the map $(x',x'') \to r_X(x',x'')$ is an infinitely divisible probability measure supported on $[0, \operatorname{diam}(\mathcal{X})]$ and hence it must be a point mass at zero because any infinitely divisible probability measure with bounded support is a point mass and if that point mass was not at zero, then the distribution of $(r_X(\xi_i,\xi_j))_{i,j\in\mathbb{N}}$ for an i.i.d. sequence $(\xi_k)_{k\in\mathbb{N}}$ with common distribution μ_X would certainly not satisfy the condition (1.1).

5. Arithmetic properties

The theory of *Delphic semigroups* was developed in [Ken68, Dav69] to generalize the decomposability properties of probability distributions with respect to convolution to an abstract setting. In the following we show that $(\mathbb{M}, \boxplus)$ is a Delphic semigroup. Let us associate with each converging sequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$ in $\mathbb{M}$ its limit $L((\mathcal{X}_n))$. If $\mathcal{Y}_n \leqslant \mathcal{X}_n$ for all n, then $(\mathcal{Y}_n)_{n\in\mathbb{N}}$ is a subset of $\bigcup_{\mathcal{Z}\in\mathbb{S}}\{\mathcal{Y} \in \mathbb{M} : \mathcal{Y} \leqslant \mathcal{Z}\}$ for the compact set $\mathbb{S} := (\mathcal{X}_n)_{n\in\mathbb{N}}$ and so it is compact by Lemma 2.10(a).

Thus, $(\mathcal{Y}_n)_{n\in\mathbb{N}}$ admits a convergent subsequence, so that the condition (A') from [Dav69] holds.

For each $A \in \mathbb{A}$, the function $D_A = -\log \chi_A$ is a continuous homomorphism from $(\mathbb{M}, \boxplus)$ to $(\mathbb{R}_+, +)$. In particular, the function D from (3.3) has this property. By Lemma 3.8(a), for each $A \in \mathbb{A}$ and each $\varepsilon > 0$ there is $\delta > 0$ such that, for any $\mathcal{X} \in \mathbb{M}$ satisfying $D(\mathcal{X}) \leqslant \delta$ one has $D_A(\mathcal{X}) \leqslant \varepsilon$. If $(A_k)_{k\in\mathbb{N}}$ is a countable subset of $\mathbb{A}$, such that $(A_k)_{k\in\mathbb{N}} \cap \mathbb{R}_+^{\binom{n}{2}}$ is dense in $\mathbb{R}_+^{\binom{n}{2}}$ for all $n \geqslant 2$, then the values $D_A(\mathcal{X})$ uniquely determine $\mathcal{X}$, so that the homomorphisms D_k satisfy the condition (H) of [Dav69]. By [Dav69, Theorem 3], the semigroup $(\mathbb{M}, \boxplus)$ is sequentially Delphic; in particular, it satisfies the (CLT) condition that requires that the limit of any converging null-array is infinitely divisible. By [Ken68, Theorem II], each element $\mathcal{X} \neq \mathcal{E}$ of $(\mathbb{M}, \boxplus)$ is either irreducible or has an irreducible factor or is infinitely divisible. The last is impossible by Proposition 4.5, so the next result holds.

Proposition 5.1. *Given any $\mathcal{X} \in \mathbb{M}\backslash\{\mathcal{E}\}$, there exists $\mathcal{Y} \in \mathbb{I}$ with $\mathcal{Y} \leqslant \mathcal{X}$.*

The prime numbers are the analogue of irreducible elements for the semigroup of positive integers equipped with the usual multiplication. The key to proving the Fundamental Theorem of Arithmetic (that every positive integer other than 1 has a factorization into primes that is unique up to the order of the factors) is a lemma due to Euclid which says that if a prime number divides the product of two positive integers, then it must divide one of the factors. For general commutative semigroups, the term "prime" is usually reserved for elements that exhibit the generalization of this property (see, for example, [Cli38]). Accordingly, we say that an element $\mathcal{X} \in \mathbb{M}\backslash\{\mathcal{E}\}$ is *prime* if $\mathcal{X} \leqslant \mathcal{Y} \boxplus \mathcal{Z}$ for $\mathcal{Y}, \mathcal{Z} \in \mathbb{M}$ implies that $\mathcal{X} \leqslant \mathcal{Y}$ or $\mathcal{X} \leqslant \mathcal{Z}$. Prime elements are clearly irreducible, but the converse is not *a priori* true and there are commutative, cancellative semigroups where the analogue of the converse is false.

Before showing that the notions of irreducibility and primality coincide in our setting, we need the following elementary lemma which we prove for the sake of completeness.

Lemma 5.2. *Let $\xi_{00}, \xi_{01}, \xi_{10}, \xi_{11}$ be random elements of the respective metric spaces $X_{00}, X_{01}, X_{10}, X_{11}$. Suppose that the pairs (ξ_{00}, ξ_{01}) and (ξ_{10}, ξ_{11}) are independent and that the pairs (ξ_{00}, ξ_{10}) and (ξ_{01}, ξ_{11}) are independent. Then, $\xi_{00}, \xi_{01}, \xi_{10}, \xi_{11}$ are independent.*

Proof. Suppose that $f_{ij} : X_{ij} \to \mathbb{R}$, $i, j \in \{0, 1\}$, are bounded Borel functions. Using first the independence of (ξ_{00}, ξ_{01}) and (ξ_{10}, ξ_{11}), and

then the independence of (ξ_{00}, ξ_{10}) and (ξ_{01}, ξ_{11}), we have

$$\begin{aligned}
&\mathbb{E}[f_{00}(\xi_{00}) f_{01}(\xi_{01}) f_{10}(\xi_{10}) f_{11}(\xi_{11})] \\
&\quad = \mathbb{E}[f_{00}(\xi_{00}) f_{01}(\xi_{01})]\mathbb{E}[f_{10}(\xi_{10}) f_{11}(\xi_{11})] \\
&\quad = \mathbb{E}[f_{00}(\xi_{00})]\mathbb{E}[f_{01}(\xi_{01})]\mathbb{E}[f_{10}(\xi_{10})]\mathbb{E}[f_{11}(\xi_{11})],
\end{aligned}$$

as required. □

Proposition 5.3. *All irreducible elements of* $\mathbb{M}$ *are prime. Moreover, if* $(\mathcal{Y}_n)_{n\in\mathbb{N}}$ *is a sequence of elements of* $\mathbb{M}$ *such that* $\lim_{n\to\infty} \boxplus_{k=0}^{n} \mathcal{Y}_k = \mathcal{Y}$ *exists and* $\mathcal{X} \in \mathbb{I}$ *is such that* $\mathcal{X} \preccurlyeq \mathcal{Y}$*, then* $\mathcal{X} \preccurlyeq \mathcal{Y}_n$ *for some* $n \in \mathbb{N}$*.*

Proof. Consider the first claim. Suppose that $\mathcal{X} \in \mathbb{M}$ is irreducible and $\mathcal{X} \preccurlyeq \mathcal{Y} \boxplus \mathcal{Z}$ for some $\mathcal{Y}, \mathcal{Z} \in \mathbb{M}$.

From Proposition 3.6(a) we have $\mathcal{Y} \boxplus \mathcal{Z} = \mathcal{W} \boxplus \mathcal{X}$ for some unique $\mathcal{W} \in \mathbb{M}$. From the remarks at the end of [Tar92], we may suppose that there are metric spaces $(Y', r_{Y'})$, $(X', r_{X'})$, $(X'', r_{X''})$ and $(Z'', r_{Z''})$ such that $(Y, r_Y) = (Y' \times X', r_{Y'} \oplus r_{X'})$, $(Z, r_Z) = (X'' \times Z'', r_{X''} \oplus r_{Z''})$, $(X, r_X) = (X' \times X'', r_{X'} \oplus r_{X''})$ and $(W, r_W) = (Y' \times Z'', r_{Y'} \oplus r_{Z''})$, so that $(Y \times Z, r_Y \oplus r_Z) = (W \times X, r_W \oplus r_X) = (Y' \times X' \times X'' \times Z'', r_{Y'} \oplus r_{X'} \oplus r_{X''} \oplus r_{Z''})$ (see also [Wal87] for an analogous result concerning the existence of a common refinement of two Cartesian factorizations of a (possibly infinite) graph and [AFDF00] for the case of finite metric spaces). It follows from Lemma 5.2 that there are probability measures $\mu_{Y'}$, $\mu_{X'}$, $\mu_{X''}$ and $\mu_{Z''}$ such that $\mu_Y = \mu_{Y'} \otimes \mu_{X'}$, $\mu_Z = \mu_{X''} \otimes \mu_{Z''}$, $\mu_X = \mu_{X'} \otimes \mu_{X''}$, $\mu_W = \mu_{Y'} \otimes \mu_{Z''}$, and $\mu_Y \otimes \mu_Z = \mu_W \otimes \mu_X = \mu_{Y'} \otimes \mu_{X'} \otimes \mu_{X''} \otimes \mu_{Z''}$. Thus, $\mathcal{Y} = \mathcal{Y}' \boxplus \mathcal{X}''$, $\mathcal{Z} = \mathcal{X}' \boxplus \mathcal{Z}''$, $\mathcal{X} = \mathcal{X}' \boxplus \mathcal{X}''$, $\mathcal{W} = \mathcal{Y}' \boxplus \mathcal{Z}''$, and $\mathcal{Y} \boxplus \mathcal{Z} = \mathcal{W} \boxplus \mathcal{X} = \mathcal{Y}' \boxplus \mathcal{X}' \boxplus \mathcal{X}'' \boxplus \mathcal{Z}''$. This contradicts the irreducibility of $\mathcal{X}$ unless $\mathcal{X}' = \mathcal{E}$ or $\mathcal{X}'' = \mathcal{E}$, in which case $\mathcal{X} \preccurlyeq \mathcal{Z}$ or $\mathcal{X} \preccurlyeq \mathcal{Y}$, thus establishing the first claim of the proposition.

Turning to the second claim, let $(\mathcal{Y}_n)_{n\in\mathbb{N}}$, $\mathcal{Y} \in \mathbb{M}$ and $\mathcal{X} \in \mathbb{I}$ satisfy the hypotheses of the claim. By Proposition 3.6(b), for each $n \in \mathbb{N}$ we have $\mathcal{Y} = \boxplus_{k=0}^{n} Y_k \boxplus \mathcal{Z}_n$ for some unique $\mathcal{Z}_n \in \mathbb{M}$. If there is no $n \in \mathbb{N}$ such that $\mathcal{X} \preccurlyeq \mathcal{Y}_n$, then, by the first part of the proposition, $\mathcal{X} \preccurlyeq \mathcal{Z}_n$ for all $n \in \mathbb{N}$. By Proposition 3.6(b), this means that $\mathcal{Z}_n = \mathcal{X} \boxplus \mathcal{W}_n$ for some unique $\mathcal{W}_n \in \mathbb{M}$ and hence $\chi_A(\mathcal{Z}_n) \leqslant \chi_A(\mathcal{X})$ for all $A \in \mathbb{A}$, see Lemma 3.2(b). However, $\lim_{n\to\infty} \chi_A(\boxplus_{k=0}^{n} Y_k) = \chi_A(\mathcal{Y})$ for all $A \in \mathbb{A}$ and so $\lim_{n\to\infty} \chi_A(\mathcal{Z}_n) = 1$ for all $A \in \mathbb{A}$, implying that $\chi_A(\mathcal{X}) = 1$ for all $A \in \mathbb{A}$. This, however, is impossible, since it would imply that $\mathcal{X} = \mathcal{E} \notin \mathbb{I}$. □

The next result is standard, but we include it for the sake of completeness.

Corollary 5.4. *Suppose for $\mathcal{X} \in \mathcal{K}$ and distinct $\mathcal{Y}_0, \dots, \mathcal{Y}_n \in \mathbb{I}$ that $\mathcal{Y}_k \leqslant \mathcal{X}$ for $k = 0, \dots, n$. Then, $\boxplus_{k=0}^{n} Y_k \leqslant \mathcal{X}$.*

Proof. The proof is by induction. The statement is certainly true for $n = 0$. Suppose it is true for $n = r$ and consider the case $n = r+1$. We have $\mathcal{X} = \boxplus_{k=0}^{r} Y_k \boxplus \mathcal{W}_r$ for some $\mathcal{W}_r \in \mathbb{M}$ by the inductive assumption. Because $\mathcal{Y}_{r+1} \leqslant \mathcal{X} = \boxplus_{k=0}^{r} Y_k \boxplus \mathcal{W}_r$, it follows from Proposition 5.3 that either $\mathcal{Y}_{r+1} \leqslant \mathcal{Y}_k$ for some k with $1 \leqslant k \leqslant r$ or $\mathcal{Y}_{r+1} \leqslant \mathcal{W}_r$. The former alternative is impossible because $\mathcal{Y}_0, \dots, \mathcal{Y}_r, \mathcal{Y}_{r+1} \in \mathbb{I}$ are distinct. Thus, $\mathcal{Y}_{r+1} \leqslant \mathcal{W}_r$ and we have $\mathcal{W}_r = \mathcal{Y}_{r+1} \boxplus \mathcal{W}_{r+1}$ for some $\mathcal{W}_{r+1} \in \mathbb{M}$. This implies that $\mathcal{X} = \boxplus_{k=0}^{r} Y_k \boxplus \mathcal{Y}_{r+1} \boxplus \mathcal{W}_{r+1}$ and hence $\boxplus_{k=0}^{r+1} Y_k \leqslant \mathcal{X}$, completing the inductive step. □

Theorem 5.5. *Given any $\mathcal{X} \in \mathbb{M} \backslash \{\mathcal{E}\}$, there is either a finite sequence $(\mathcal{X}_n)_{n=0}^{N}$ or an infinite sequence $(\mathcal{X}_n)_{n=0}^{\infty}$ of irreducible elements of $\mathbb{M}$ such that $\mathcal{X} = \boxplus_{k=0}^{N} \mathcal{X}_k$ in the first case and $\mathcal{X} = \lim_{n\to\infty} \boxplus_{k=0}^{n} \mathcal{X}_k$ in the second. The sequence is unique up to the order of its terms. Each irreducible element appears a finite number of times, so the representation is specified by the irreducible elements that appear and their finite multiplicities.*

Proof. As $(\mathbb{M}, \boxplus)$ is a Delphic semigroup, [Ken68, Theorem III] yields that each $\mathcal{X} \in \mathbb{M}$ admits a representation as the sum of irreducible elements. Note that each element of the sum appears only a finite number of times, since otherwise the sum would diverge by Proposition 3.9(b).

We now turn to the uniqueness claim. This may fail because $\mathcal{X}$ has two different representations as a finite sum of irreducible elements, one representation as a finite sum and another as a limit of finite sums, or two different representations as a limit of finite sums. We deal with the last case. The other two are similar and are left to the reader. Suppose then that two sequences $(\mathcal{X}'_n)_{n\in\mathbb{N}}$ and $(\mathcal{X}''_n)_{n\in\mathbb{N}}$ represent $\mathcal{X}$. An argument similar to one above shows that any particular irreducible element appears a finite number of times in each sequence. Suppose that $\mathcal{Y} \in \mathbb{I}$ appears M' times in $(\mathcal{X}'_n)_{n\in\mathbb{N}}$ and M'' times in $(\mathcal{X}''_n)_{n\in\mathbb{N}}$ with $M' \neq M''$. Assume without loss of generality that $M' > M''$. We have $\mathcal{Y}^{\boxplus M'} \boxplus \mathcal{Z}' = \mathcal{X} = \mathcal{Y}^{\boxplus M''} \boxplus \mathcal{Z}''$, where $\mathcal{Z}', \mathcal{Z}'' \in \mathbb{M}$ are such that $\mathcal{Y} \nleqslant \mathcal{Z}'$ and $\mathcal{Y} \nleqslant \mathcal{Z}''$. Using Proposition 3.6(a), $\mathcal{Y}^{\boxplus (M'-M'')} \boxplus \mathcal{Z}' = \mathcal{Z}''$. By Proposition 5.3, $\mathcal{Y}$ is prime, so that it divides one of the factors in the representation of $\mathcal{Z}''$ meaning that so $\mathcal{Y} \leqslant \mathcal{Z}''$, contrary to the assumption. □

Remark 5.6. It is an easy consequence of Theorem 5.5 that, for the partial order $\leqslant$, every pair of elements of $\mathbb{M}$ has a *join* (that is, a least upper bound) and a *meet* (that is, a greatest lower bound), and so $\mathbb{M}$

with these operations is a *lattice*. It is not hard to check that this lattice is distributive (that is, the meet operation distributes over the join operation and vice versa). Furthermore, the Gromov–Prohorov distance between $\mathcal{X}$ and $\mathcal{Y}$ equals the maximum of the distances between the meet of $\mathcal{X}$ and $\mathcal{Y}$ and either $\mathcal{X}$ or $\mathcal{Y}$.

Remark 5.7. Given $f : \mathbb{I} \to [0,1]$, the map $\chi : \mathbb{M} \to [0,1]$ that sends $\mathcal{X}$ to $\prod_n f(\mathcal{X}_n)$, where $\mathcal{X}_0, \mathcal{X}_1, \dots$ are as in Theorem 5.5, is a semicharacter.

The following result will be a key ingredient in the characterization of the infinitely divisible random elements of $\mathbb{M}$ in Theorem 9.1.

Corollary 5.8. *If $\Phi : \mathbb{R}_+ \to \mathbb{M}$ is a continuous function such that $\Phi(s) \leqslant \Phi(t)$ for $0 \leqslant s \leqslant t < \infty$, then $\Phi \equiv \mathcal{E}$.*

Proof. Suppose that Φ is a function with the stated properties. If $\Phi \not\equiv \mathcal{E}$, then there exist $0 < u < v < \infty$ such that $\Phi(u) < \Phi(v)$. It follows from Theorem 5.5 that there exists $\mathcal{Y} \in \mathbb{I}$ such that the multiplicity of $\mathcal{Y}$ in the factorization of $\Phi(v)$ is strictly greater than the multiplicity of $\mathcal{Y}$ in the factorization of $\Phi(u)$. Define $M : \mathbb{R}_+ \to \mathbb{N}$ by setting $M(s)$, $s \geqslant 0$, to be the multiplicity of $\mathcal{Y}$ in the factorization of $\Phi(s)$. This function is nondecreasing and so there must exist $u \leqslant t \leqslant v$ such that $M(t-) < M(t+)$. Thus, $\Phi(t-\varepsilon) \boxplus \mathcal{Y} \boxplus \cdots \boxplus \mathcal{Y} \leqslant \Phi(t+\varepsilon)$ for all $\varepsilon > 0$, where there are $M(t+) - M(t-)$ summands in the sum, and this contradicts the continuity of Φ by Lemma 2.8 and Lemma 2.2. □

The following result is an immediate consequence of the absence of infinitely divisible metric measure spaces.

Corollary 5.9. *If $\Phi : \mathbb{R}_+ \to \mathbb{M}$ is a function such that $\Phi(s) \boxplus \Phi(t) = \Phi(s+t)$ for $0 \leqslant s, t < \infty$, then $\Phi \equiv \mathcal{E}$.*

Remark 5.10. Although Corollary 5.9 says there are no nontrivial *additive* functions from $\mathbb{R}_+$ to $\mathbb{M}$, there do exist nontrivial *superadditive functions*; that is, functions $\Phi : \mathbb{R}_+ \to \mathbb{M}$ such that $\Phi(0) = \mathcal{E}$ and $\Phi(s) \boxplus \Phi(t) \leqslant \Phi(s+t)$ for $0 \leqslant s, t < \infty$. For example, take $\mathcal{X} \in \mathbb{M} \backslash \{\mathcal{E}\}$ and set $\Phi(t) = \mathcal{X} \boxplus \cdots \boxplus \mathcal{X}$ for $n \leqslant t < n+1$, $n \in \mathbb{N}$, where the sum has n terms and we interpret the empty sum as $\mathcal{E}$. We have

$$\Phi(s) \boxplus \Phi(t) = \Phi(\lfloor s \rfloor) \boxplus \Phi(\lfloor t \rfloor) = \Phi(\lfloor s \rfloor + \lfloor t \rfloor) \leqslant \Phi(s+t).$$

However, by Corollary 5.8 there are no nontrivial continuous superadditive functions. Furthermore, there are no superadditive functions Φ such that $\Phi(t) \neq \mathcal{E}$ for all $t > 0$.

There are also nontrivial *subadditive functions*; that is, functions $\Phi : \mathbb{R}_+ \to \mathbb{M}$ such that $\Phi(0) = \mathcal{E}$ and $\Phi(s) \boxplus \Phi(t) \geqslant \Phi(s+t)$ for

$0 \leqslant s, t < \infty$. For example, it suffices to take some $\mathcal{X} \in \mathbb{M}\backslash\{\mathcal{E}\}$ and set $\Phi(t) = \mathcal{X}$ for $t > 0$. However, there are no continuous subadditive functions because if Φ is such a function and $\mathcal{Y} \in \mathbb{I}$ is such that $\mathcal{Y} \leqslant \Phi(t)$, then it follows from $\Phi(\frac{t}{2}) \boxplus \Phi(\frac{t}{2}) \geqslant \Phi(t)$ that $\mathcal{Y} \leqslant \Phi(\frac{t}{2})$ and hence $\mathcal{Y} \leqslant \Phi(\frac{t}{2^n})$ for all $n \in \mathbb{N}$, but this contradicts the continuity of Φ at 0.

6. Prime factorizations as measures

Theorem 5.5 guarantees that any $\mathcal{X} \in \mathbb{M}$ has a unique representation as $\mathcal{X} = \boxplus_k \mathcal{Y}_k^{\boxplus m_k}$, where the $\mathcal{Y}_k \in \mathbb{I}$ are distinct, the integers m_k are positive, and we define the empty sum to be $\mathcal{E}$. Since $\boxplus_k \mathcal{Y}_k^{\boxplus m_k}$ converges, $d_{\mathrm{GPr}}(\mathcal{Y}_k, \mathcal{E}) \to 0$ as $k \to \infty$ in case of an infinite factorization, so that the number of $\mathcal{Y}_k$ outside any neighborhood of $\mathcal{E}$ is finite. It is natural to code such a factorization as the measure $\Psi(\mathcal{X}) := \sum_k m_k \delta_{\mathcal{Y}_k}$ on $\mathbb{M}$ that is concentrated on $\mathbb{I}$ and assigns mass m_k to the point $\mathcal{Y}_k$ for each k.

Denote by $\mathfrak{N}$ the family of Borel measures N on $\mathbb{M}$ such that $N(\mathbb{M}\backslash\mathbb{I}) = 0$ and $N(B) \in \mathbb{N}$ for every Borel set B that does not intersect some neighborhood of $\mathcal{E}$. Any $N \in \mathfrak{N}$ can be represented as the positive integer linear combination of Dirac measures

$$N = \sum_k m_k \delta_{\mathcal{Y}_k}$$

for distinct $\mathcal{Y}_k \in \mathbb{I}$ and positive integers m_k, where the sum may be finite or countably infinite depending on the cardinality of the support of N. Given $N \in \mathfrak{N}$ with such a representation we define a unique element of $\mathbb{M}$ by

$$\Sigma(N) := \boxplus_k \mathcal{Y}_k^{\boxplus m_k},$$

if the sum converges (recall from Proposition 3.9(e) that the convergence of the sum is independent of the order summands). Thus, $\Sigma(\Psi(\mathcal{X})) = \mathcal{X}$ for all $\mathcal{X} \in \mathbb{M}$.

It is possible to topologize $\mathfrak{N}$ with the metrizable $w^{\#}$-topology of [DVJ03, Section A2.6]. This topology is the topology generated by integration against bounded continuous functions that are supported outside a neighborhood of $\mathcal{E}$. The resulting Borel σ-field coincides with the σ-field generated by the $\mathbb{N}$-valued maps $N \mapsto N(B)$ Borel measurable, where B is a Borel subset of $\mathbb{M}$ that is disjoint from some neighborhood of $\mathcal{E}$, see [DVJ03, Theorem A2.6.III].

Proposition 6.1. *The map $\Psi : \mathbb{M} \to \mathfrak{N}$ is Borel measurable.*

Proof. The set $\{(\mathcal{X}, \mathcal{Y}) \in \mathbb{M}^2 : \mathcal{Y} \leqslant \mathcal{X}\}$ is closed by Corollary 3.11(a) and the set $\mathbb{I}$ is G_δ by Proposition 4.2. It follows that the set $\mathbb{B} :=$

$\{(\mathcal{X},\mathcal{Y}) \in \mathbb{M}^2 : \mathcal{Y} \preccurlyeq \mathcal{X},\ \mathcal{Y} \in \mathbb{I}\}$ is a G_δ subset of $\mathbb{M}^2$ and, in particular, it is Borel.

For any $\mathcal{X} \in \mathbb{M}$, the section $\mathbb{B}_\mathcal{X} := \{\mathcal{Y} \in \mathbb{M} : (\mathcal{X},\mathcal{Y}) \in \mathbb{B}\} = \{\mathcal{Y} \in \mathbb{M} : \mathcal{Y} \preccurlyeq \mathcal{X},\ \mathcal{Y} \in \mathbb{I}\}$ is countable (indeed, it is discrete with $\mathcal{E}$ as its only possible accumulation point).

By [Kec95, Exercise 18.15], the sets $\mathbb{T}_n := \{\mathcal{X} \in \mathbb{M} : \#\mathbb{B}_\mathcal{X} = n\}$, $n = 1, 2, \ldots, \infty$, are Borel and for each n there exist Borel functions $(\theta_i^{(n)})_{0 \leqslant i < n}$ such that:

- $\theta_i^{(n)} : \mathbb{T}_n \to \mathbb{M}$,
- the sets $\{(\mathcal{X},\mathcal{Y}) : \mathcal{X} \in \mathbb{T}_n,\ \mathcal{Y} = \theta_i^{(n)}(\mathcal{X})\}$, $0 \leqslant i < n$, $n = 1, 2, \ldots, \infty$, are pairwise disjoint,
- $\mathbb{B}_\mathcal{X} = \{\theta_i^{(n)}(\mathcal{X}) : 0 \leqslant i < n\}$ for $\mathcal{X} \in \mathbb{T}_n$, $n = 1, 2, \ldots, \infty$.

Recall the Borel function M from Corollary 3.11(b). For $\mathcal{X} \in \mathbb{T}_n$, the set $\{(\theta_i^{(n)}(\mathcal{X}), M(\mathcal{X}, \theta_i^{(n)}(\mathcal{X})) : 0 \leqslant i < n\}$ is a listing of the elements of the set $\{\mathcal{Y} \in \mathbb{I} : \mathcal{Y} \preccurlyeq \mathcal{X}\}$ along with their multiplicities in the prime factorization of $\mathcal{X}$. The functions $\mathcal{X} \mapsto (\theta_i^{(n)}(\mathcal{X}), M(\mathcal{X}, \theta_i^{(n)}(\mathcal{X}))$, $\mathcal{X} \in \mathbb{T}_n$, $0 \leqslant i < n$, $n = 1, 2, \ldots, \infty$, are measurable and so

$$\mathcal{X} \mapsto \Psi(\mathcal{X}) = \sum_{i=0}^{n} M(\mathcal{X}, \theta_i^{(n)}(\mathcal{X}))\delta_{\theta_i^{(n)}(\mathcal{X})}$$

for $\mathcal{X} \in \mathbb{T}_n$, provides a measurable map from $\mathbb{M}$ to $\mathfrak{N}$, see [DVJ08, Proposition 9.1.X]. □

Remark 6.2. The map Ψ is not continuous for the $w^\#$-topology. In fact, any $\mathcal{X} \in (\mathbb{M}\backslash\mathbb{I})\backslash\{\mathcal{E}\}$ is a discontinuity point, as the following argument demonstrates. Because $\mathbb{I}$ is dense in $\mathcal{K}$, it is possible to find a sequence $\mathcal{X}_n \in \mathbb{I}$ that converges to $\mathcal{X}$. Therefore, $\Psi(\mathcal{X}_n) = \delta_{\mathcal{X}_n}$, whereas $\Psi(\mathcal{X})$ has total mass at least two and the distance between any atom of $\Psi(\mathcal{X})$ and the point $\mathcal{X}_n$ is bounded away from zero uniformly in n.

We omit the straightforward proof of the next result.

Lemma 6.3. *The set $\{N \in \mathfrak{N} : \Sigma(N)$ is defined$\}$ is measurable and the restriction of the map Σ to this set is measurable.*

7. Scaling

Given $\mathcal{X} \in \mathbb{M}$ and $a > 0$, set $a\mathcal{X} := (X, ar_X, \mu_X) \in \mathbb{M}$. This *scaling* operation $(a, \mathcal{X}) \mapsto a\mathcal{X}$ is jointly continuous by Lemma 3.5 and it satisfies the first distributivity law

$$a(\mathcal{X} \boxplus \mathcal{Y}) = (a\mathcal{X}) \boxplus (a\mathcal{Y}) \quad \text{for } \mathcal{X}, \mathcal{Y} \in \mathbb{M} \text{ and } a > 0. \tag{7.1}$$

The semigroup $(\mathbb{M}, \boxplus)$ equipped with this scaling operation is a convex cone. The neutral element $\mathcal{E}$ is the origin in this cone; that is, $\lim_{a\downarrow 0} a\mathcal{X} = \mathcal{E}$ for all $\mathcal{X} \in \mathbb{M}$, which follows from Lemma 3.5. Note that $\operatorname{diam}(a\mathcal{X}) = a \operatorname{diam}(\mathcal{X})$ for $\mathcal{X} \in \mathbb{M}$ and $a > 0$.

It is immediate from (7.1) that $\mathcal{Y} \in \mathbb{I}$ if and only if $a\mathcal{Y} \in \mathbb{I}$ for all $a > 0$

Remark 7.1. The Gromov–Prohorov metric is not homogeneous for this scaling operation; that is, $d_{\mathrm{GPr}}(a\mathcal{X}, a\mathcal{Y})$ is not generally equal to $a d_{\mathrm{GPr}}(\mathcal{X}, \mathcal{Y})$ for $a > 0$ and $\mathcal{X}, \mathcal{Y} \in \mathbb{M}$. Moreover, it is not possible to equip $\mathbb{M}$ with a homogeneous metric that induces the same topology as d_{GPr}. To see that this is so, first note that for each $n \geqslant 2$ there exists $\mathcal{X}_n \in \mathbb{M} \in \{\mathcal{E}\}$ such that $d_{\mathrm{GPr}}(c\mathcal{X}_n, \mathcal{E}) \leqslant n^{-1}$ for all $c > 0$; for example, take $\mathcal{X}_n$ to be a two-point space with unit distance between the points and respective masses n^{-1} and $1 - n^{-1}$. For any sequence $(c_n)_{n\in\mathbb{N}}$, we have $d_{\mathrm{GPr}}(c_n\mathcal{X}_n, \mathcal{E}) \to 0$, while if δ is a homogeneous metric, then $\delta(c_n\mathcal{X}_n, \mathcal{E}) = \delta(c_n\mathcal{X}_n, c_n\mathcal{E}) = c_n\delta(\mathcal{X}_n, \mathcal{E})$ does not converge to zero if $c_n \to \infty$ sufficiently rapidly.

.

We have seen that $(\mathbb{M}, \leqslant)$ is a distributive lattice. There is a large literature on lattices that are equipped with an action of the additive group of the real numbers (see, for example [Kap48, Pie59, Hol69]). Using exponential and logarithms to go back and forth from one setting to the other, this work can be recast as being about lattices with an action of the group consisting of $\mathbb{R}_{++} = (0, \infty)$ equipped with the usual multiplication of real numbers. Unfortunately, one of the hypotheses usually assumed in this area translates to our setting as an assumption that $\mathcal{X} < a\mathcal{X}$ for $a > 1$. The following result shows that this is far from being the case and also that scaling operation certainly does not satisfy the second distributivity law.

Proposition 7.2. *Let $\mathcal{X}$ be a metric measure space.*

a) *If $\mathcal{X} \leqslant a\mathcal{X}$ for some $a \neq 1$, then $a > 1$ and $\mathcal{X} = \bigboxplus_{k=1}^{\infty} a^{-k}\mathcal{Z}$, where $\mathcal{Z}$ is defined by the requirement that $a\mathcal{X} = \mathcal{X} \boxplus \mathcal{Z}$.*
b) *If $(a\mathcal{X}) \boxplus (b\mathcal{X}) = c\mathcal{X}$, for some $a, b, c > 0$, then $\mathcal{X} = \mathcal{E}$.*

Proof. (a) Suppose that $\mathcal{X} \neq \mathcal{E}$ is such that $\mathcal{X} \leqslant a\mathcal{X}$ for $a \neq 1$. Recall the function $R(\mathcal{X})$ from (3.4). Because $R(\mathcal{X}) \leqslant R(a\mathcal{X})$ and $R(a\mathcal{X})$ is monotone as function of $a \in \mathbb{R}_+$, it must be the case that $a > 1$. We have $\mathcal{X} = a^{-1}\mathcal{Z} \boxplus a^{-1}\mathcal{X}$. Iterating, we have $\mathcal{X} = \bigboxplus_{k=1}^{n} a^{-k}\mathcal{Z} \boxplus a^{-n}\mathcal{X}$. Since $\chi_1(a^{-n}\mathcal{X}) \to 1$, we have $a^{-n}\mathcal{X} \to \mathcal{E}$ by Lemma 3.5. By Proposition 3.6(b) or Proposition 3.9(c), $\lim_{n\to\infty} \bigboxplus_{k=1}^{n} a^{-k}\mathcal{Z}$ exists.

(b) Suppose that $(a\mathcal{X}) \boxplus (b\mathcal{X}) = c\mathcal{X}$ for some $a, b, c > 0$. Since $R(c\mathcal{X}) = R((a\mathcal{X}) \boxplus (b\mathcal{X})) \geqslant (R(a\mathcal{X}) \vee R(b\mathcal{X}))$, we have $a \vee b \leqslant c$. An irreducible element $\mathcal{Y} \in \mathbb{I}$ appears in the factorization of $\mathcal{X}$ guaranteed by Theorem 5.5 if and only if $c\mathcal{Y} \in \mathbb{I}$ appears in the factorization of $c\mathcal{X}$, and similar remarks hold for the factorizations of $a\mathcal{X}$ and $b\mathcal{X}$. Then $c\mathcal{Y} \leqslant a\mathcal{X}$ or $c\mathcal{Y} \leqslant b\mathcal{X}$. Assume the first, so that $\frac{c}{a}\mathcal{Y} \leqslant \mathcal{X}$, so that $\frac{c}{a}\mathcal{Y}$ appears in the factorization of $\mathcal{X}$. Iteration yields $(\frac{c}{a})^n\mathcal{Y} \leqslant \mathcal{X}$ for all $n \geqslant 1$, so that the spaces $((\frac{c}{a})^n\mathcal{Y})_{n\in\mathbb{N}}$ all belong to the prime decomposition of $\mathcal{X}$ which then diverges by Proposition 3.9(g). □

Remark 7.3. While it is possible to introduce a notion of convexity for subsets of $\mathbb{M}$ using the addition and scaling in an obvious way, the absence of the second distributivity law makes the situation entirely different from the vector space case. For instance, a single point $\{\mathcal{X}\}$ is not convex for $\mathcal{X} \neq \mathcal{E}$ and its convex hull is the set of spaces of the form $a_1\mathcal{X} \boxplus \cdots \boxplus a_n\mathcal{X}$ for $a_1, \ldots, a_n \geqslant 0$ such that $a_1 + \cdots + a_n = 1$. It is a consequence of Remark 8.3 for $a_1 = \cdots = a_n = n^{-1}$ that this latter set is not even pre-compact.

Remark 7.4. The map that sends $a \in \mathbb{R}_{++}$ to the automorphism $\mathcal{X} \mapsto a\mathcal{X}$ of $(\mathbb{M}, \boxplus)$ is a homomorphism from $(\mathbb{R}_{++}, \times)$ to the group of automorphisms of $(\mathbb{M}, \boxplus)$. We can therefore define the *semidirect product* $\mathbb{M} \rtimes \mathbb{R}_{++}$ to be the semigroup consisting of the set $\mathbb{M} \times \mathbb{R}_{++}$ equipped with the operation $\circledast$ defined by

$$(\mathcal{X}, a) \circledast (\mathcal{Y}, b) := (\mathcal{X} \boxplus (a\mathcal{Y}), ab).$$

This semigroup has the identity element $(\mathcal{E}, 1)$ and is noncommutative. The semidirect product of the group $(\mathbb{G}, \boxplus)$ considered in Remark 3.7 and the group $(\mathbb{R}_{++}, \times)$ can be defined similarly. It would be interesting to extend the investigation of infinite divisibility in Section 9 to this semigroup and group, but we leave this topic for future study.

8. The Laplace transform

A random element in $\mathbb{M}$ is defined with respect to the Borel σ-algebra on $\mathbb{M}$ generated by the Gromov–Prohorov metric.

Lemma 8.1. *Two $\mathbb{M}$-valued random elements $\mathbf{X}$ and $\mathbf{Y}$ have the same distribution if and only if $\mathbb{E}[\chi_A(\mathbf{X})] = \mathbb{E}[\chi_A(\mathbf{Y})]$ for all $A \in \mathbb{A}$.*

Proof. By Lemma 3.5, the set of functions $\{\chi_A : A \in \mathbb{A}\}$ generates the Borel σ-algebra on $\mathbb{M}$. From Remark 3.3, this set is a semigroup under the usual multiplication of functions and, in particular, it is closed under multiplication. The result now follows from a standard monotone class argument. □

Remark 8.2. Recall from Section 6 the set $\mathfrak{N}$ of $\mathbb{N}$-valued measures that are concentrated on $\mathbb{I}$ and the associated measurable structure. Following the usual terminology, we define a *point process* to be a random element of $\mathfrak{N}$. By Proposition 6.1, any $\mathbb{M}$-valued random element $\mathbf{X}$ can, in the notation of Section 6, be viewed as a point process $\mathbf{N} := \Psi(\mathbf{X})$ such that $\Sigma(\mathbf{N}) = \mathbf{X}$. If we write $\mathbf{N} = \sum m_k \delta_{\mathbf{Y}_k}$ on $\mathbb{I}$, then

$$\mathbb{E}[\chi_A(\mathbf{X})] = \mathbb{E}[\chi_A(\Sigma(\Psi(\mathbf{X}))] = \mathbb{E}\left[\prod \chi_A(\mathbf{Y}_k)^{m_k}\right].$$

The right-hand side is the expected value of the product of the function χ_A applied to each of the atoms of $\mathbf{N}$ taking into account their multiplicities and hence it is an instance of the *probability generating functional* of the point process $\mathbf{N}$, see [DVJ08, Equation (9.4.13)].

Remark 8.3. A fairly immediate consequence of Lemma 8.1 is that there is no analogue of a law of large numbers for random elements of $\mathbb{M}$ in the sense that if $(\mathbf{X}_k)_{k\in\mathbb{N}}$ is an i.i.d. sequence of random elements of $\mathbb{M}$ that are not identically equal $\mathcal{E}$, then $\frac{1}{n}\coprod_{k=0}^{n-1}\mathbf{X}_k$ does not even have a subsequence that converges in distribution. Indeed, for $A \in \mathbb{A}$ with $A \in \mathbb{R}_+^{\binom{m}{2}}$ we have

$$\begin{aligned}
\lim_{n\to\infty}\mathbb{E}\left[\chi_A\left(\frac{1}{n}\coprod_{k=0}^{n-1}\mathbf{X}_k\right)\right] &= \lim_{n\to\infty}\left(\mathbb{E}\left[\chi_A\left(\frac{1}{n}\mathbf{X}_1\right)\right]\right)^n \\
&= \lim_{n\to\infty}\left(\int_{\mathbb{M}}\int_{X^m}\exp\left(-\frac{1}{n}\sum_{1\leqslant i<j\leqslant m}a_{ij}r_X(x_i,x_j)\right)\right. \\
&\quad \left.\times\ \mu_X^{\otimes m}(dx)\,\mathbb{P}\{\mathbf{X}_1\in d\mathcal{X}\}\right)^n \\
&= \exp\left(-\lim_{n\to\infty} n\left(1-\int_{\mathbb{M}}\int_{X^m}\exp\left(-\frac{1}{n}\sum_{1\leqslant i<j\leqslant m}a_{ij}r_X(x_i,x_j)\right)\right.\right. \\
&\quad \left.\left.\times\,\mu_X^{\otimes m}(dx)\,\mathbb{P}\{\mathbf{X}_1\in d\mathcal{X}\}\right)\right) \\
&= \exp\left(-\int_{\mathbb{M}}\int_{X^m}\sum_{1\leqslant i<j\leqslant m}a_{ij}r_X(x_i,x_j)\,\mu_X^{\otimes m}(dx)\,\mathbb{P}\{\mathbf{X}_1\in d\mathcal{X}\}\right) \\
&= \exp\left(-\sum_{1\leqslant i<j\leqslant m}a_{ij}\int_{\mathbb{M}}\int_{X^2}r_X(x_1,x_2)\,\mu_X^{\otimes 2}(dx)\,\mathbb{P}\{\mathbf{X}_1\in d\mathcal{X}\}\right).
\end{aligned}$$

If some subsequence of $\frac{1}{n}\boxplus_{k=0}^{n-1}\mathbf{X}_k$ converged in distribution to a limit $\mathbf{Y}$, then we would have

$$\int_{\mathbb{M}}\int_{Y^m}\exp\left(-\sum_{1\leqslant i<j\leqslant m}a_{ij}r_Y(y_i,y_j)\right)\mu_Y^{\otimes m}(dy)\,\mathbb{P}\{\mathbf{Y}\in d\mathcal{Y}\}$$
$$=\exp\left(-\sum_{1\leqslant i<j\leqslant m}a_{ij}\int_{\mathbb{M}}\int_{X^2}r_X(x_1,x_2)\,\mu_X^{\otimes 2}(dx)\,\mathbb{P}\{\mathbf{X}_1\in d\mathcal{X}\}\right).$$

The right-hand side is the exponential of a linear combination of a_{ij} and so corresponds to the Laplace transform of a deterministic random vector. By the unicity of Laplace transforms for nonnegative random vectors, this implies that

$$\int_{\mathbb{M}}\mu_Y^{\otimes 2}\Bigg\{(y_1,y_2)\in Y^2 : r_Y(y_1,y_2)$$
$$\neq\int_{X^2}r_X(x_1,x_2)\,\mu_X^{\otimes 2}(dx)\,\mathbb{P}\{\mathbf{X}_1\in d\mathcal{X}\}\Bigg\}\mathbb{P}\{\mathbf{Y}\in d\mathcal{Y}\}$$
$$=0,$$

and hence there is a constant $c>0$ such that for $\mathbb{P}\{\mathbf{Y}\in\cdot\}$-almost all $Y\in\mathbb{M}$ we have $r_Y(y_1,y_2)=c$ for $\mu_Y^{\otimes 2}$-almost all $(y_1,y_2)\in Y^2$, but this is impossible for a nontrivial metric space (Y,r_Y) and probability measure μ_Y with full support.

9. Infinitely divisible random elements

A random element $\mathbf{Y}$ of $\mathbb{M}$ is *infinitely divisible* if for each positive integer n there are i.i.d. random elements $\mathbf{Y}_{n1},\ldots,\mathbf{Y}_{nn}$ such that $\mathbf{Y}$ has the same distribution as $\boxplus_{k=1}^{n}\mathbf{Y}_{nk}$.

An $\mathbb{M}$-valued *Lévy process* is a $\mathbb{M}$-valued stochastic process $(\mathbf{X}_t)_{t\geqslant 0}$ such that:

- $\mathbf{X}_0=\mathcal{E}$;
- $t\mapsto\mathbf{X}_t$ is càdlàg (that is, right-continuous with left-limits);
- given $0=t_0<t_1<\ldots<t_n$, there are independent $\mathbb{M}$-valued random variables $\mathbf{Z}_{t_0t_1},\mathbf{Z}_{t_1t_2},\ldots,\mathbf{Z}_{t_{n-1}t_n}$ such that the distribution of $\mathbf{Z}_{t_mt_{m+1}}$ only depends on $t_{m+1}-t_m$ for $0\leqslant m\leqslant n-1$ and $\mathbf{X}_{t_\ell}=\mathbf{X}_{t_k}\boxplus\mathbf{Z}_{t_kt_{k+1}}\boxplus\cdots\boxplus\mathbf{Z}_{t_{\ell-1}t_\ell}$ for $0\leqslant k<\ell\leqslant n$.

An account of the general theory of infinitely divisible distributions on commutative semigroups may be found in [BCR84]. The following result is the analogue in our setting of the classical Lévy–Hinčin–Itô description of an infinitely divisible, real-valued random variable.

Theorem 9.1. a) *A random element* $\mathbf{Y}$ *of* $\mathbb{M}$ *is infinitely divisible if and only if it has the same distribution as* $\mathbf{X}_1$*, where* $(\mathbf{X}_t)_{t\geqslant 0}$ *is a Lévy process with distribution uniquely specified by that of* $\mathbf{Y}$*.*

b) *For each* $t > 0$ *there is a unique random element* $\Delta\mathbf{X}_t$ *such that* $\mathbf{X}_t = \mathbf{X}_{t-} \boxplus \Delta\mathbf{X}_t$*.*

c) *For each* $t > 0$*,* $\mathbf{X}_t = \boxplus_{0<s\leqslant t} \Delta\mathbf{X}_s$*, where the sum is a well-defined limit that does not depend on the order of the summands.*

d) *The set of points* $\{(t, \Delta\mathbf{X}_t) : \Delta\mathbf{X}_t \neq \mathcal{E}\}$ *form a Poisson point process on* $\mathbb{R}_+ \times (\mathbb{M}\backslash\{\mathcal{E}\})$ *with intensity measure* $\lambda\otimes\nu$*, where* λ *is Lebesgue measure and* ν *is a* σ*-finite measure on* $\mathbb{M}\backslash\{\mathcal{E}\}$ *such that*

$$\int (D(\mathcal{X}) \wedge 1)\, \nu(d\mathcal{X}) < \infty\,. \tag{9.1}$$

e) *Conversely, if* ν *is a* σ*-finite measure on* $\mathbb{M}\backslash\{\mathcal{E}\}$ *satisfying* (9.1)*, then there is an infinitely divisible random element* $\mathbf{Y}$ *and a Lévy process* $(\mathbf{X}_t)_{t\geqslant 0}$ *such that (a)-(d) hold, and the distributions of this random element and Lévy process are unique.*

Proof. Write $\mathbb{D}$ for the set of nonnegative dyadic rational numbers. It follows from the infinite divisibility of $\mathbf{Y}$ and the Kolmogorov extension theorem that we can build a family of random variables $(\mathbf{X}_q)_{q\in\mathbb{D}}$ such that:

- $\mathbf{X}_0 = \mathcal{E}$,
- $\mathbf{X}_1$ has the same distribution as $\mathbf{Y}$,
- Given $q_0, \ldots, q_n \in \mathbb{D}$ with $0 = q_0 < q_1 < \ldots < q_n$, there are independent $\mathbb{M}$-valued random variables $\mathbf{Z}_{q_0q_1}, \mathbf{Z}_{q_1q_2}, \ldots, \mathbf{Z}_{q_{n-1}q_n}$ such that the distribution of $\mathbf{Z}_{q_mq_{m+1}}$ only depends on $q_{m+1} - q_m$ for $0 \leqslant m \leqslant n-1$ and $\mathbf{X}_{q_\ell} = \mathbf{X}_{q_k} \boxplus \mathbf{Z}_{q_kq_{k+1}} \boxplus \cdots \boxplus \mathbf{Z}_{q_{\ell-1}q_\ell}$ for $0 \leqslant k < \ell \leqslant n$. In particular, $\mathbf{X}_p \leqslant \mathbf{X}_q$ for $p, q \in \mathbb{D}$ with $p \leqslant q$.

We claim that if $p \in \mathbb{D}$, then

$$\lim_{q\downarrow p,\, q\in\mathbb{D}} \mathbf{X}_q = \mathbf{X}_p, \quad \text{a.s.} \tag{9.2}$$

To see that this is the case, note that if $p, q \in \mathbb{D}$ with $p < q$, then $\mathbf{X}_q = \mathbf{X}_p \boxplus \mathbf{Z}_{pq}$ and it suffices to show that $\lim_{q\downarrow p,\, q\in\mathbb{D}} d_{\mathrm{GPr}}(\mathbf{Z}_{pq}, \mathcal{E}) = 0$ almost surely.

By Lemma 3.5, it will certainly suffice to show that $\lim_{q\downarrow p,\, q\in\mathbb{D}} D(\mathbf{Z}_{pq}) = 0$ a.s. However, note that if we set $T_0 = 0$ and $T_r = D(\mathbf{Z}_{p,p+r})$ for $r \in \mathbb{D}\backslash\{0\}$, then the $\mathbb{R}_+$-valued process $(T_r)_{r\in\mathbb{D}}$ has stationary independent increments. It is well-known that such a process has a càdlàg extension to the index set $\mathbb{R}_+$ and hence, in particular, $\lim_{r\downarrow 0,\, r\in\mathbb{D}} T_r = 0$.

Lemma 9.4 applied to $(\mathbf{X}_p)_{p\in\mathbb{D}}$ gives that it is possible to extend $(\mathbf{X}_p)_{p\in\mathbb{D}}$ to a Lévy process $(\mathbf{X}_t)_{t\geqslant 0}$. This establishes (a). Moreover, for each $t > 0$ there is a unique $\mathbb{M}$-valued random variable $\Delta\mathbf{X}_t$ such that $\mathbf{X}_t = \mathbf{X}_{t-} \boxplus \Delta\mathbf{X}_t$, and $\mathbf{X}_t = \boxplus_{0<s\leqslant t} \Delta\mathbf{X}_s$, where the sum is well-defined by Proposition 3.9(e). This establishes (b) and (c).

A standard argument (see, for example, [Kal02, Theorem 12.10]) shows that the set of points $\{(t, \Delta\mathbf{X}_t) : \Delta\mathbf{X}_t \neq \mathcal{E}\}$ form a Poisson point process on $\mathbb{R}_+ \times (\mathbb{M}\backslash\{\mathcal{E}\})$. The stationarity of the "increments" of $(\mathbf{X}_t)_{t\geqslant 0}$ forces the intensity measure of this Poisson point process to be of the form $\lambda \otimes \nu$, and the fact that $\sum_{0<s\leqslant t} D(\Delta\mathbf{X}_s)$ is finite for all $t \geqslant 0$ implies (9.1), see, for example, [Kal02, Corollary 12.11]. This establishes (d).

We omit the straightforward proof of (e). □

Following the usual terminology, we refer to the σ-finite measure ν in Theorem 9.1 as the Lévy measure of the infinitely divisible random element $\mathbf{Y}$ or the Lévy process $(\mathbf{X}_t)_{t\geqslant 0}$. The next result is immediate from Theorem 9.1, the multiplicative property of the semicharacters χ_A, and the usual formula for the Laplace functional of a Poisson process.

Corollary 9.2. *If $\mathbf{Y}$ is an infinitely divisible random element of $\mathbb{M}$ with Lévy measure ν, then the Laplace transform of $\mathbf{Y}$ is given by*

$$\mathbb{E}[\chi_A(\mathbf{Y})] = \exp\left(-\int (1-\chi_A(\mathcal{Y}))\,\nu(d\mathcal{Y})\right), \quad A \in \mathbb{A}. \tag{9.3}$$

Remark 9.3. In the notation of Theorem 9.1, the random measure

$$\sum_{0<t\leqslant 1} \delta_{\Delta\mathbf{X}_t}$$

is a Poisson random measure on $\mathbb{M}$ with intensity measure ν and we have $\mathbf{Y} = \mathbf{X}_1 = \boxplus_{0<t\leqslant 1} \Delta\mathbf{X}_t$. The push-forward of this random measure by the map Ψ of Proposition 6.1 is a Poisson random measure on the space $\mathfrak{N}$ of $\mathbb{N}$-valued measures that are concentrated on $\mathbb{I}$. The intensity measure of this latter Poisson random measure is the push-forward Q of the Lévy measure ν by Ψ. The "points" of the latter Poisson random measure are usually called *clusters* in the point processes literature, while Q itself is called the *KLM measure*, see [DVJ08, Definition 10.2.IV]. Let $\mathbf{N}$ be the point process on $\mathbb{I}$ obtained as the superposition of clusters; that is, $\mathbf{N} = \sum_{0<t\leqslant 1} \Psi(\Delta\mathbf{X}_t)$ is the sum of the $\mathbb{N}$-valued measures given by each individual cluster. This point process on $\mathbb{I}$ is called the *Poisson cluster process* in the Poisson point process literature. The infinite divisibility of $\mathbf{Y}$ implies the infinite divisibility of the point process $\Psi(\mathbf{Y})$ and the equality $\Psi(\mathbf{Y}) = \mathbf{N}$ is an

instance of the well-known fact that infinitely divisible point processes are Poisson cluster processes. Furthermore, (9.3) corresponds to the classical representation of the probability generating functional of an infinitely divisible point process specialized to the space $\mathbb{I}$, see [DVJ08, Theorem 10.2.V]. On the other hand, if $\mathbf{M}$ is a Poisson cluster process on $\mathbb{I}$ such that $\Sigma(\mathbf{M})$ is almost surely well-defined, then $\Sigma(\mathbf{M})$ is an infinitely divisible random element of $\mathbb{M}$, and our observations above show that all infinitely divisible random elements of $\mathbb{M}$ appear this way.

We end this section with a deterministic path-regularization result that was used in the proof of Theorem 9.1.

Lemma 9.4. *Suppose that $\Xi : \mathbb{D} \to \mathbb{M}$ is such that $\Xi(0) = \mathcal{E}$, $\Xi(p) \leqslant \Xi(q)$ for $p, q \in \mathbb{D}$ with $0 \leqslant p \leqslant q$, and $\lim_{q \downarrow p,\, q \in \mathbb{D}} \Xi(q) = \Xi(p)$ for all $p \in \mathbb{D}$. Then, $\bar{\Xi}(t) := \lim_{q \downarrow t,\, q \in \mathbb{D}} \Xi(q)$ exists for all $t \in \mathbb{R}_+$. Moreover, the function $\bar{\Xi} : \mathbb{R}_+ \to \mathbb{M}$ has the following properties:*

- $\bar{\Xi}(p) = \Xi(p)$ *for* $p \in \mathbb{D}$,
- $\bar{\Xi}(s) \leqslant \bar{\Xi}(t)$ *for* $s, t \in \mathbb{R}_+$ *with* $s \leqslant t$,
- $t \mapsto \bar{\Xi}(t)$ *is càdlàg*,
- *for* $p, q \in \mathbb{D}$ *with* $0 \leqslant p < q$, *there is a unique* $\Theta(p, q) \in \mathbb{M}$ *such that* $\Xi(q) = \Xi(p) \boxplus \Theta(p, q)$,
- *for* $0 \leqslant s < t$, *there is a unique* $\bar{\Theta}(s, t) \in \mathbb{M}$ *such that* $\bar{\Xi}(t) = \bar{\Xi}(s) \boxplus \bar{\Theta}(s, t)$ *and* $\bar{\Theta}(s, t) = \lim_{p \downarrow s,\, q \downarrow t,\, p, q \in \mathbb{D}} \Theta(p, q)$,
- *for each* $t > 0$ *there is a unique* $\Delta\bar{\Xi}(t) \in \mathbb{M}$ *such that* $\bar{\Xi}(t) = \lim_{s \uparrow t} \bar{\Xi}(s) \boxplus \Delta\bar{\Xi}(t)$,
- $\sum_{u < t \leqslant v} D(\Delta\bar{\Xi}(t)) \leqslant D(\bar{\Theta}(u, v))$ *for all* $0 \leqslant u < v$,
- *the sum* $\boxplus_{0 < s \leqslant t} \Delta\bar{\Xi}(s)$ *is well-defined for all* $t \geqslant 0$,
- $\bar{\Xi}(t) = \boxplus_{0 < s \leqslant t} \Delta\bar{\Xi}(s)$ *for all* $t \geqslant 0$.

Proof. It follows from Proposition 3.9(d) that $\lim_{q \downarrow t,\, q \in \mathbb{D}} \Xi(q) =: \bar{\Xi}(t)$ exists for all $t \geqslant 0$.

It is clear that $\bar{\Xi}(p) = \Xi(p)$ for $p \in \mathbb{D}$ and that $\bar{\Xi}(s) \leqslant \bar{\Xi}(t)$ for $s, t \in \mathbb{R}_+$ with $s \leqslant t$. It is also clear that $t \mapsto \bar{\Xi}(t)$ is right-continuous. By Proposition 3.9(c), $\bar{\Xi}(t-) := \lim_{s \uparrow t} \bar{\Xi}(s)$ exists for all $t > 0$ and $\bar{\Xi}(t-) \leqslant \bar{\Xi}(t)$ for all $t > 0$.

The existence and uniqueness of $\bar{\Theta}(s, t)$ such that $\bar{\Xi}(t) = \bar{\Xi}(s) \boxplus \bar{\Theta}(s, t)$ and the fact that $\bar{\Theta}(s, t) = \lim_{p \downarrow s,\, q \downarrow t,\, p, q \in \mathbb{D}} \Theta(p, q)$ follow from Proposition 3.6(b).

It is a consequence of Proposition 3.6(b) that $\Delta\bar{\Xi}(t)$ exists and is well-defined.

For any $0 \leqslant u < v$ and $u < t_1 < \cdots < t_n \leqslant v$ we have $\Delta\bar{\Xi}(t_1) \boxplus \cdots \boxplus \Delta\bar{\Xi}(t_n) \leqslant \bar{\Theta}(u, v)$. Hence, by Proposition 3.6(a), $\boxplus_{0 < s \leqslant t} \Delta\bar{\Xi}(s)$ is well-defined.

It is clear that $\boxplus_{0<s\leqslant t}\Delta\bar{\Xi}(s) \leqslant \bar{\Xi}(t)$ for all $t \geqslant 0$ and so we can use Proposition 3.6 to define a unique function $\Phi : \mathbb{R}_+ \to \mathbb{M}$ such that $\bar{\Xi}(t) = \Phi(t)\boxplus\boxplus_{0<s\leqslant t}\Delta\bar{\Xi}(s)$ for all $t \geqslant 0$. The function Φ is continuous and $\Phi(s) \leqslant \Phi(t)$ for $0 \leqslant s < t$. Also, $\Phi(0) = \mathbb{E}$. Corollary 5.8 gives that $\Phi \equiv \mathcal{E}$, completing the proof of the lemma. □

10. Stable random elements

A $\mathbb{M}$-valued random element $\mathbf{Y}$ is *stable* with index $\alpha > 0$ if for any $a, b > 0$ the random element $(a+b)^{\frac{1}{\alpha}}\mathbf{Y}$ has the same distribution as $a^{\frac{1}{\alpha}}\mathbf{Y}' \boxplus b^{\frac{1}{\alpha}}\mathbf{Y}''$, where $\mathbf{Y}'$ and $\mathbf{Y}''$ are independent copies of $\mathbf{Y}$. Note that a stable random element is necessarily infinitely divisible. If $\mathbf{Y}$ is stable and almost surely takes values in the space of bounded metric measure spaces, then its diameter is a nonnegative strictly stable random variable.

There is a general investigation of stable random elements of convex cones in [DMZ08]. In general, not all such objects have Laplace transforms that are of the type analogous to those described in Corollary 9.2. For example, there can be Gaussian-like distributions. However, no such complexities arise in our setting.

Theorem 10.1. *Suppose that $\mathbf{Y}$ is a nontrivial α-stable random element of $\mathbb{M}$. Then, $0 < \alpha < 1$ and the Lévy measure ν of $\mathbf{Y}$ obeys the scaling condition*

$$\nu(aB) = a^{-\alpha}\nu(B), \quad a > 0, \tag{10.1}$$

for all Borel sets $B \subseteq \mathbb{M}$. Conversely, if ν is a σ-finite measure on $\mathbb{M}\backslash\{\mathcal{E}\}$ that obeys the scaling condition for $0 < \alpha < 1$ and satisfies (9.1), then ν is the Lévy measure of an α-stable random element.

Proof. If $(\mathbf{X}_t)_{t\geqslant 0}$ is the Lévy process corresponding to $\mathbf{Y}$, then it is not difficult to check that the process $(a^{-\frac{1}{\alpha}}\mathbf{X}_{at})_{t\geqslant 0}$ has the same distribution as $(\mathbf{X}_t)_{t\geqslant 0}$, and the scaling condition for ν follows easily. Since $r_{\mathbf{Y}}(\xi_1, \xi_2)$ is a nonnegative stable random variable of index α, we necessarily have $\alpha \in (0, 1)$. The remainder of the proof is straightforward and we omit it. □

Remark 10.2. One of the conclusions of Theorem 10.1 is that there are no nontrivial α-stable random elements for $\alpha \geqslant 1$. This is also a consequence of the following argument. If $\mathbf{Y}$ was a nontrivial α-stable random element and $(\mathbf{Y}_k)_{k\in\mathbb{N}}$ was a sequence of independent copies of $\mathbf{Y}$, then $n^{-\frac{1}{\alpha}}\boxplus_{k=0}^{n-1}\mathbf{Y}_k$ would have the same distribution as $\mathbf{Y}$ and hence $\frac{1}{n}\boxplus_{k=0}^{n-1}\mathbf{Y}_k$ would certainly converge in distribution as $n \to \infty$,

but this contradicts Remark 8.3, where we observed that there is no analogue of a law of large numbers in our setting.

We finish this section with an analogue of the classical LePage representation of stable real random variables.

Theorem 10.3. *The following are equivalent for a random element* $\mathbf{Y}$ *of* $\mathbb{M}$.

a) *The random element* $\mathbf{Y}$ *is* α*-stable.*

b) *The random element* $\mathbf{Y}$ *is infinitely divisible with a Lévy measure* ν *that is of the form* $\nu(B) = \alpha \int_0^\infty \pi(t^{-1}B)\, t^{-(\alpha+1)}\, dt$ *for all Borel sets* $B \subseteq \mathbb{M}\backslash\{\mathcal{E}\}$, *where* π *is a probability measure on* $\mathbb{M}\backslash\{\mathcal{E}\}$ *such that*

$$\int_{\mathbb{M}\backslash\{\mathcal{E}\}} \int_{Z^2} r_Z^\alpha(x,y)\, \mu_Z^{\otimes 2}(dz_1, dz_2)\, \pi(d\mathcal{Z}) < \infty, \tag{10.2}$$

In particular, π *assigns all of its mass to metric measure spaces* $\mathcal{Z}$ *for which*

$$\int_{Z^2} r_Z^\alpha(x,y)\, \mu_Z^{\otimes 2}(dz_1, dz_2) < \infty. \tag{10.3}$$

c) *The random element* $\mathbf{Y}$ *has the same distribution as*

$$\biguplus_{n\in\mathbb{N}} \Gamma_n^{-\frac{1}{\alpha}} \mathbf{Z}_n, \tag{10.4}$$

where $(\Gamma_n)_{n\in\mathbb{N}}$ *is the sequence of successive arrivals of a homogeneous, unit intensity Poisson point process on* $\mathbb{R}_+$ *and* $(\mathbf{Z}_n)_{n\in\mathbb{N}}$ *is a sequence of i.i.d. random elements in* $\mathbb{M}\backslash\{\mathcal{E}\}$ *with common distribution* π *such that* (10.2) *holds.*

Proof. Suppose that $\mathbf{Y}$ is α-stable with Lévy measure ν. We know from Theorem 10.1 that ν satisfies the scaling condition (10.1) and the integrability condition (9.1).

For any $\mathcal{X} \in \mathbb{M}\backslash\{\mathcal{E}\}$ the function $t \mapsto D(t\mathcal{X})$ is strictly increasing and $\sup_{t>0} D(t\mathcal{X}) = -\log(r_X^{\otimes 2}\{(x_1,x_2) \in X^2 : x_1 = x_2\})$ with the convention $-\log(0) = \infty$. Set

$$\mathbb{V}_0 := \{\mathcal{X} \in \mathbb{M}\backslash\{\mathcal{E}\} : \sup_{t>0} D(t\mathcal{X}) > 1\}$$

and

$$\mathbb{V}_k := \{\mathcal{X} \in \mathbb{M}\backslash\{\mathcal{E}\} : 2^{-k} < \sup_{t>0} D(t\mathcal{X}) \leqslant 2^{-(k-1)}\}$$

for $k \geqslant 1$. The sets $\mathbb{V}_k$, $k \in \mathbb{N}$, are disjoint, their union is $\mathbb{M}\backslash\{\mathcal{E}\}$, and $\mathcal{X} \in \mathbb{V}_k$ for some $k \in \mathbb{N}$ if and only if $t\mathcal{X} \in \mathbb{V}_k$ for all $t \in \mathbb{R}_{++}$.

Define $\tau : \mathbb{M}\backslash\{\mathcal{E}\} \to \mathbb{R}_{++}$ as follows. For $\mathcal{X} \in \mathbb{V}_k$, set $\tau(\mathcal{X}) := \inf\{t > 0 : D(t^{-1}\mathcal{X}) \leqslant 2^{-k}\}$. Observe that $\tau(s\mathcal{X}) = s\tau(\mathcal{X})$ for all $\mathcal{X} \in \mathbb{M}\backslash\{\mathcal{E}\}$ and $s \in \mathbb{R}_{++}$ and that $D(\tau(\mathcal{X})^{-1}\mathcal{X}) = 2^{-k}$ for $\mathcal{X} \in \mathbb{V}_k$. Note that if $\tau(\mathcal{X})^{-1}\mathcal{X} = \tau(\mathcal{Y})^{-1}\mathcal{Y}$, then $\{t\mathcal{X} : t > 0\} = \{t\mathcal{Y} : t > 0\}$, whereas if $\tau(\mathcal{X})^{-1}\mathcal{X} \neq \tau(\mathcal{Y})^{-1}\mathcal{Y}$, then $\{t\mathcal{X} : t > 0\} \cap \{t\mathcal{Y} : t > 0\} = \varnothing$. In other words, the set $\{\tau(\mathcal{X})^{-1}\mathcal{X} : \mathcal{X} \in \mathbb{M}\backslash\{\mathcal{E}\}\}$ is a cross-section of orbit representatives for the action of the group $\mathbb{R}_{++}$ on $\mathbb{M}\backslash\{\mathcal{E}\}$.

For each $k \in \mathbb{N}$, the maps $\mathcal{X} \mapsto (\tau(\mathcal{X})^{-1}\mathcal{X}, \tau(\mathcal{X}))$ and $(\mathcal{Y}, t) \mapsto t\mathcal{Y}$ are mutually inverse Borel bijections between the Borel sets $\mathbb{V}_k$ and $\mathbb{S}_k \times \mathbb{R}_{++}$, where $\mathbb{S}_k := \{\mathcal{X} \in \mathbb{V}_k : \tau(\mathcal{X}) = 1\}$. Let $\tilde{\nu}$ be the push-forward of ν by the map $\mathcal{X} \mapsto (\tau(\mathcal{X})^{-1}\mathcal{X}, \tau(\mathcal{X}))$ and define a measure ρ_k on $\mathbb{S}_k$ by $\rho_k(A) = \tilde{\nu}(A \times [1, \infty))$. Since

$$\rho_k(\mathbb{S}_k) = \tilde{\nu}(\mathbb{S}_k \times [1, \infty)) \leqslant \nu\{\mathcal{X} \in \mathbb{M}\backslash\{\mathcal{E}\} : D(\mathcal{X}) \geqslant 2^{-k}\},$$

it follows from (9.1) that the total mass of ρ_k is finite.

The scaling property (10.1) of ν is equivalent to the scaling property $\tilde{\nu}(A \times sB) = s^{-\alpha}\tilde{\nu}(A \times B)$ for $s \in \mathbb{R}_{++}$ and Borel sets $A \subseteq \mathbb{S}_k$ and $B \subseteq \mathbb{R}_{++}$. Thus, if we let θ be the measure on $\mathbb{R}_{++}$ given by $\theta(dt) = \alpha t^{-(\alpha+1)}\, dt$, then

$$\begin{aligned}\tilde{\nu}(A \times [b, \infty)) &= \tilde{\nu}(A \times b[1, \infty)) \\ &= b^{-\alpha}\tilde{\nu}(A \times [1, \infty)) \\ &= \rho_k(A) \times \theta([b, \infty))\end{aligned}$$

for $A \subseteq \mathbb{S}_k$. Therefore the restriction of $\tilde{\nu}$ to $\mathbb{S}_k \times \mathbb{R}_{++}$ is $\rho_k \otimes \theta$ and hence the restriction of ν to $\mathbb{V}_k$ is the push-forward of $\rho_k \otimes \theta$ by the map $(\mathcal{Y}, t) \mapsto t\mathcal{Y}$.

We can think of ρ_k as a measure on all of $\mathbb{V}_k$. For $c_k \in \mathbb{R}_{++}$, let η_k be the measure on $\mathbb{V}_k$ that assigns all of its mass to the set $c_k\mathbb{S}_k$ and is given by $\eta_k(A) = c_k^{\alpha}\rho_k(c_k^{-1}A)$. We have

$$\begin{aligned}\eta_k \otimes \theta\{(\mathcal{Y}, t) : t\mathcal{Y} \in B\} &= \int \eta_k(t^{-1}B)\alpha t^{-(\alpha+1)}\, dt \\ &= \int c_k^{\alpha}\rho_k(c_k^{-1}t^{-1}B)\alpha t^{-(\alpha+1)}\, dt \\ &= \int \rho_k(s^{-1}B)\alpha s^{-(\alpha+1)}\, ds \\ &= \rho_k \otimes \theta\{(\mathcal{Y}, t) : t\mathcal{Y} \in B\} \\ &= \nu(B),\end{aligned}$$

and so η_k is a finite measure with total mass $c_k^{\alpha}\rho_k(\mathbb{S}_k)$ that has the property that the push-forward of $\eta_k \otimes \theta$ by the map $(\mathcal{Y}, t) \mapsto t\mathcal{Y}$ is the restriction of ν to $\mathbb{V}_k$.

We can regard η_k as being a finite measure on all of $\mathbb{M}\backslash\{\mathcal{E}\}$ and, by choosing the constants c_k, $k \in \mathbb{N}$, appropriately we can arrange for $\pi := \sum_{k\in\mathbb{N}} \eta_k$ to be a probability measure. We have

$$\nu(B) = \alpha \int_0^\infty \pi(t^{-1}B)\, t^{-(\alpha+1)}\, dt$$

for all Borel sets $B \subseteq \mathbb{M}\backslash\{\mathcal{E}\}$.

It follows from (9.1) that

$$\int_0^\infty \int_{\mathbb{M}\backslash\{\mathcal{E}\}} (D(t\mathcal{Z}) \wedge 1) t^{-(\alpha+1)}\, \pi(d\mathcal{Z})\, dt < \infty.$$

By Lemma 3.8, for any σ-finite measure λ on $\mathbb{M}$ the integral

$$\int_0^\infty \int_{\mathbb{M}\backslash\{\mathcal{E}\}} (D(t\mathcal{Z}) \wedge 1) t^{-(\alpha+1)}\, \lambda(d\mathcal{Z})\, dt$$

is bounded above and below by constant multiples of

$$\int_{\mathbb{M}\backslash\{\mathcal{E}\}} \int_{Z^2} ((tr_Z(z_1, z_2)) \wedge 1) t^{-(\alpha+1)}\, \mu_Z^{\otimes 2}(dz_1, dz_2)\, \lambda(d\mathcal{Z})\, dt.$$

The latter integral is a constant multiple of

$$\int_{\mathbb{M}\backslash\{\mathcal{E}\}} \int_{Z^2} r_Z^\alpha(x, y)\, \mu_Z^{\otimes 2}(dz_1, dz_2)\, \lambda(d\mathcal{Z})$$

because

$$\int_0^\infty ((tr) \wedge 1) t^{-(\alpha+1)}\, dt = \frac{1}{\alpha(1-\alpha)} r^\alpha.$$

for any $r \geqslant 0$.

This completes the proof that (a) implies (b). The proof that (b) implies (a) simply involves checking that the measure ν satisfies the scaling property (10.1) and the integrability property (9.1). The former is obvious and the latter follows from the argument immediately above.

The proof that (a) and (b) are equivalent to (c) requires showing that ν is a measure satisfying the conditions of (b) if and only if the points of a Poisson random measure on $\mathbb{M}\backslash\{\mathcal{E}\}$ with intensity ν have the same distribution as the random set $(\Gamma_n^{-\frac{1}{\alpha}} \mathbf{Z}_n)_{n\in\mathbb{N}}$. However, if $(\mathbf{Z}_n, \Gamma_n)_{n\in\mathbb{N}}$ are as in (c), then they are the points of a Poisson random measure on $(\mathbb{M}\backslash\{\mathcal{E}\}) \times \mathbb{R}_{++}$ with intensity $\pi \otimes \lambda$, where λ is Lebesgue measure, and so $(\mathbf{Z}_n, \Gamma_n^{-\frac{1}{\alpha}})_{n\in\mathbb{N}}$ are the points of a Poisson random measure with intensity $\pi \otimes \theta$, where the measure θ is as above. □

Remark 10.4. The probability measure π in Theorem 10.3 is not unique. However, in the proof that (a) implied (b) the π that was constructed was concentrated on a set $\mathbb{T}$ with the property that for all $\mathcal{X} \in \mathbb{M}\backslash\{\mathcal{E}\}$

there is a unique $t \in \mathbb{R}_{++}$ such that $t\mathcal{X} \in \mathbb{T}$. If part (b) holds with a π that is supported on a set $\mathbb{U}$ with this property, then π is the unique probability measure concentrated on $\mathbb{U}$ that leads to a representation of ν in the manner described in the theorem.

Remark 10.5. It follows readily from Theorem 10.3 that a bounded metric measure space is α-stable if and only if it admits a representation of the form (10.4) where $(\mathbf{Z}_n)_{n\in\mathbb{N}}$ is any sequence of i.i.d. random elements in $\mathbb{M}$ such that $\operatorname{diam}(\mathbf{Z}_n) = \gamma$ almost surely for a suitable constant γ. An alternative proof of this fact can be carried out using [DMZ08, Theorems 3.6 and 7.14].

Example 10.6. We can construct an α-stable random element $\mathbf{Y}$ by considering the LePage series in which the $\mathbf{Z}_n$ are copies of some common nonrandom bounded metric measure space. In this case on the set of full probability where $\sum_{n\in\mathbb{N}} \Gamma_n^{-\frac{1}{\alpha}} < \infty$, $\mathbf{Y}$ is the infinite Cartesian product $Y := Z^\infty$ equipped with the random metric

$$r_Y((z'_n),(z''_n)) := \sum_{n\in\mathbb{N}} \Gamma_n^{-\frac{1}{\alpha}} r_Z(z'_n, z''_n)$$

and the probability measure $\mu_Y := \mu_Z^{\otimes\infty}$.

11. Thinning

Recall the map Ψ that associates with each $\mathcal{X} \in \mathbb{M}$ an $\mathbb{N}$-valued measure on $\mathbb{I}$. For $p \in [0,1]$, the independent p-thinning of an $\mathbb{N}$-valued measure $N := \sum_k m_k \delta_{y_k}$ is defined in the usual way as $N^{(p)} := \sum_k \xi_k \delta_{y_k}$, where ξ_k, $k \in \mathbb{N}$, are independent binomial random variables with parameters m_k and p. In other words, each atom of N is retained with probability p and otherwise eliminated independently of all other atoms and taking into account the multiplicities.

Applying an independent p-thinning procedure to the point process $\mathbf{N} := \Psi(\mathbf{X})$ generated by random element $\mathbf{X}$ in $\mathbb{M}$ yields an $\mathbb{M}$-valued random element $\mathbf{X}^{(p)} := \Sigma(\mathbf{N}^{(p)})$ that we call the *p-thinning* of $\mathbf{X}$. Note that the $\mathbf{X}^{(p)} \leqslant \mathbf{X}$, $\mathbf{X}^{(0)} = \mathcal{E}$, $\mathbf{X}^{(1)} = \mathbf{X}$, and for $0 \leqslant p, q \leqslant 1$ the random element $(\mathbf{X}^{(p)})^{(q)}$ has the same distribution as the random element $\mathbf{X}^{(pq)}$. It is possible to build an $\mathbb{M}$-valued strong Markov process $(\mathbf{X}_t)_{t\geqslant 0}$ so that the conditional distribution of $\mathbf{X}_{s+t}$ given $\{\mathbf{X}_s = \mathcal{X}\}$ is the e^{-t}-thinning of $\mathcal{X}$; each irreducible factor of $\mathbf{X}$ is equipped with an independent exponential random clock that has expected value 1 and the factor appears in the decomposition of $\mathbf{X}_t$ into irreducibles provided its clock has not rung by time t.

Also, if $\mathbf{X}$ and $\mathbf{Y}$ are independent random elements and $\mathbf{X}^{(p)}$ and $\mathbf{Y}^{(p)}$ are constructed to be independent, then $\mathbf{X}^{(p)} \boxplus \mathbf{Y}^{(p)}$ has the same distribution as $(\mathbf{X} \boxplus \mathbf{Y})^{(p)}$. It follows from this last property that, for fixed $A \in \mathbb{A}$ and $0 \leqslant p \leqslant 1$, the map

$$\mathcal{X} \mapsto \mathbb{E}[\chi_A(\mathcal{X}^{(p)})] = \prod (1 - p + p\chi_A(\mathcal{Y}_n))$$

is a semicharacter, where the product ranges over the factors that appear in the factorization of $\mathcal{X}$ into a sum of irreducible elements of $\mathbb{M}$ (repeated, of course, according to their multiplicities). This is a particular case of the construction in Remark 5.7.

The thinning operation can be used to construct $\mathbb{M}$-valued stochastic processes that are not necessarily increasing or decreasing in the $\leqslant$ partial order by combining the $\boxplus$ addition of independent random increments with thinning; that is, the semigroup of the process is the Trotter product of the semigroup of a Lévy process and the semigroup of the Markov process introduced above that evolves in such a way that the value of the process at time t is the e^{-t}-thinning of its value at time 0.

Furthermore, the thinning procedure is the key ingredient for defining a notion of discrete stability analogous to that in [DMZ11]. A random metric measure space $\mathbf{X}$ is said to be *discrete stable* of index α if $\mathbf{X}$ coincides in distribution with $\mathbf{X}_1^{(t^{1/\alpha})} \boxplus \mathbf{X}_2^{((1-t)^{1/\alpha})}$ for $0 \leqslant t \leqslant 1$, where $\mathbf{X}_1$ and $\mathbf{X}_2$ are independent copies of $\mathbf{X}$. By an application of general results from [DMZ11] it is possible to conclude that such an $\mathbf{X}$ corresponds to a doubly stochastic (Cox) Poisson process on $\mathbb{I}$ whose random intensity measure is stable. The simplest example is $\mathbf{X} := \mathcal{Y}^{\boxplus N}$, where $\mathcal{Y} \in \mathbb{I}$ and N is an $\mathbb{N}$-valued discrete α-stable random variable; any such random variable N has a probability generating function of the form $\mathbb{E}[s^N] = \exp(-c(1-s)^\alpha)$, $s \in [0,1]$, where $c \in \mathbb{R}_{++}$.

12. The Gromov-Prohorov metric

We follow the definition of the *Gromov-Prohorov* metric in [GPW09].

Recall that the distance in the *Prohorov metric* between two probability measures μ_1 and μ_2 on a common metric space (Z, r_Z) is defined by

$$d_{\mathrm{Pr}}^{(Z,r_Z)}(\mu_1, \mu_2) := \inf\{\varepsilon > 0 : \mu_1(F) \leqslant \mu_2(F^\varepsilon) + \varepsilon,\ \forall F \text{ closed}\},$$

where

$$F^\varepsilon := \{z \in Z : r_Z(z, z') < \varepsilon, \text{ for some } z' \in F\}.$$

An alternative characterization of the Prohorov metric due to Strassen (see, for example, [EK86, Theorem 3.1.2] or [Dud02, Corollary 11.6.4])

is that

$$d_{\mathrm{Pr}}^{(Z,r_Z)}(\mu_1,\mu_2) = \inf_{\pi} \inf\{\varepsilon > 0 : \pi\{(z,z') \in Z \times Z : r_Z(z,z') \geqslant \varepsilon\} \leqslant \varepsilon\},$$

where the infimum is over all probability measures π such that $\pi(\cdot \times Z) = \mu_1$ and $\pi(Z \times \cdot) = \mu_2$.

The following result is no doubt well-known, but we include it for completeness. Recall that if (X, r_X) and (Y, r_Y) are two metric spaces, then $r_X \oplus r_Y$ is the metric on the Cartesian product $X \times Y$ given by $r_X \oplus r_Y((x',y'),(x'',y'')) = r_X(x',x'') + r_Y(y',y'')$.

Lemma 12.1. *Suppose that μ_1 and μ_2 (resp. ν_1 and ν_2) are probability measures on a metric space (X, r_X) (resp. (Y, r_Y)). Then,*

$$d_{\mathrm{Pr}}^{(X\times Y, r_X\oplus r_Y)}(\mu_1 \otimes \nu_1, \mu_2 \otimes \nu_2) \leqslant d_{\mathrm{Pr}}^{(X,r_X)}(\mu_1,\mu_2) + d_{\mathrm{Pr}}^{(Y,r_Y)}(\nu_1,\nu_2).$$

Proof. This is immediate from the observation that if α and β are probability measures on $X \times X$ and $Y \times Y$, respectively, such that

$$\alpha\{(x',x'') \in X \times X : r_X(x',x'') \geqslant \gamma\} \leqslant \gamma$$

and

$$\beta\{(y',y'') \in Y \times Y : r_Y(y',y'') \geqslant \delta\} \leqslant \delta$$

for $\gamma, \delta > 0$, then

$$\begin{aligned} \alpha \otimes \beta\{((x',y'),(x'',y'')) \in (X \times Y) \times (X \times Y) \\ : r_X(x',x'') + r_Y(y',y'') \geqslant \gamma + \delta\} \\ \leqslant \gamma + \delta, \end{aligned}$$

where, with a slight abuse of notation, we identify the measure $\alpha \otimes \beta$ on $(X \times X) \times (Y \times Y)$ with its push-forward on $(X \times Y) \times (X \times Y)$ by the map $((x',x''),(y',y'')) \mapsto ((x',y'),(x'',y''))$. □

The next lemma is also probably well-known.

Lemma 12.2. *Suppose that μ_1 and μ_2 are two probability measures on a metric space (X, r_X) and ν is a probability measure on another metric space (Y, r_Y). Then,*

$$d_{\mathrm{Pr}}^{(X\times Y, r_X\oplus r_Y)}(\mu_1 \otimes \nu, \mu_2 \otimes \nu) = d_{\mathrm{Pr}}^{(X,r_X)}(\mu_1,\mu_2).$$

Proof. It follows from Lemma 12.1 that

$$\begin{aligned} d_{\mathrm{Pr}}^{(X\times Y, r_X\oplus r_Y)}(\mu_1 \otimes \nu, \mu_2 \otimes \nu) &\leqslant d_{\mathrm{Pr}}^{(X,r_X)}(\mu_1,\mu_2) + d_{\mathrm{Pr}}^{(Y,r_Y)}(\nu,\nu) \\ &= d_{\mathrm{Pr}}^{(X,r_X)}(\mu_1,\mu_2). \end{aligned}$$

On the other hand, suppose that π is a probability measure on $(X \times Y) \times (X \times Y)$ such that $\pi(\cdot \times (X \times Y)) = \mu_1 \otimes \nu$, $\pi((X \times Y) \times \cdot) = \mu_2 \otimes \nu$ and

$$\pi\{((x', y'), (x'', y'')) \in (X \times Y) \times (X \times Y) : r_X(x', x'') + r_Y(y', y'') \geqslant \varepsilon\} \leqslant \varepsilon$$

for some $\varepsilon > 0$. If ρ is the push-forward of π by the map $((x', y'), (x'', y'')) \mapsto (x', x'')$, then it is clear that $\rho(\cdot \times X) = \mu_1$, $\rho(X \times \cdot) = \mu_2$ and

$$\rho\{((x', x'') \in X \times X : r_X(x', x'') \geqslant \varepsilon\} \leqslant \varepsilon,$$

and hence

$$d_{\mathrm{Pr}}^{(X, r_X)}(\mu_1, \mu_2) \leqslant d_{\mathrm{Pr}}^{(X \times Y, r_X \oplus r_Y)}(\mu_1 \otimes \nu, \mu_2 \otimes \nu).$$

□

The *Gromov-Prohorov metric* is a metric on the space of equivalence classes of metric measure space (recall that two metric measure spaces are equivalent if there is an isometry mapping one to the other such that the probability measure on the first is mapped to the probability measure on the second). Given two metric measure spaces $\mathcal{X} = (X, r_X, \mu_X)$ and $\mathcal{Y} = (Y, r_Y, \mu_Y)$, the Gromov-Prohorov distance between their equivalence classes is

$$d_{\mathrm{GPr}}(\mathcal{X}, \mathcal{Y}) := \inf_{(\phi_X, \phi_Y, Z)} d_{\mathrm{Pr}}^{(Z, r_Z)}(\mu_X \circ \phi_X^{-1}, \mu_Y \circ \phi_Y^{-1}),$$

where the infimum is taken over all metric spaces (Z, r_Z) and isometric embeddings ϕ_X of X and ϕ_Y of Y into Z, and $\mu_X \circ \phi_X^{-1}$ (resp. $\mu_Y \circ \phi_Y^{-1}$) denotes the push-forward of μX by ϕ_X (resp. μ_Y by ϕ_Y). It is easy to see that

$$d_{\mathrm{GPr}}(\mathcal{X}, \mathcal{E}) = \inf_{x \in X} \inf\{\varepsilon > 0 : \ \mu_X\{y \in X : \ r_X(x, y) \geqslant \varepsilon\} \leqslant \varepsilon\}. \tag{12.1}$$

13. Inequalities for Laplace transforms

In this section we prove two inequalities about Laplace transforms of nonnegative random variables that were used in the proof of Lemma 3.8.

Lemma 13.1. *There are constants $\kappa', \kappa'' > 0$ such that for any nonnegative random variable ξ we have*

$$\kappa'((-\log(\mathbb{E}[\exp(-\xi)])) \wedge 1) \leqslant \mathbb{E}[\xi \wedge 1] \leqslant \kappa''((-\log(\mathbb{E}[\exp(-\xi)])) \wedge 1).$$

Proof. Consider the first inequality. Recall that $1-\exp(-u) \leqslant u$ for all $u \in \mathbb{R}$. Furthermore, there is a constant $\gamma > 0$ such that $1-u \geqslant -\gamma \log u$ for $e^{-1} \leqslant u \leqslant 1$. Thus,

$$\begin{aligned}\mathbb{E}[\xi \wedge 1] &\geqslant 1 - \mathbb{E}[\exp(-(\xi \wedge 1))] \\ &\geqslant \gamma(-\log(\mathbb{E}[\exp(-(\xi \wedge 1))])).\end{aligned}$$

It will therefore suffice to show that there is a constant $\delta > 0$ such that

$$\begin{aligned}-\log(\mathbb{E}[\exp(-(\xi \wedge 1))]) &\geqslant \delta((-\log(\mathbb{E}[\exp(-\xi)])) \wedge 1) \\ &= \delta(-\log(\mathbb{E}[\exp(-\xi)] \vee e^{-1}))\end{aligned}$$

or, equivalently, that

$$\mathbb{E}[\exp(-(\xi \wedge 1))] \leqslant (\mathbb{E}[\exp(-\xi)] \vee e^{-1})^{\delta} = \mathbb{E}[\exp(-\xi)]^{\delta} \vee e^{-\delta}.$$

That is, we need to show that we can choose $\delta > 0$ such that if

$$\mathbb{E}[\exp(-(\xi \wedge 1))] > e^{-\delta}, \tag{13.1}$$

then

$$\mathbb{E}[\exp(-(\xi \wedge 1))] \leqslant \mathbb{E}[\exp(-\xi)]^{\delta}.$$

Moreover, since

$$\mathbb{E}[\exp(-(\xi \wedge 1))] = \mathbb{E}[\exp(-\xi)\mathbb{1}\{\xi < 1\}] + e^{-1}\mathbb{P}\{\xi \geqslant 1\}$$

and

$$\mathbb{E}[\exp(-\xi)] \geqslant \mathbb{E}[\exp(-\xi)\mathbb{1}\{\xi < 1\}],$$

it will be enough to establish that

$$1 + \frac{e^{-1}\mathbb{P}\{\xi \geqslant 1\}}{\mathbb{E}[\exp(-\xi)\mathbb{1}\{\xi < 1\}]} \leqslant \mathbb{E}[\exp(-\xi)\mathbb{1}\{\xi < 1\}]^{(\delta-1)}. \tag{13.2}$$

Suppose that (13.1) holds. In that case

$$\begin{aligned}e^{-\delta} &< \mathbb{E}[\exp(-(\xi \wedge 1))] \\ &= \mathbb{E}[\exp(-\xi)\mathbb{1}\{\xi < 1\}] + e^{-1}\mathbb{P}\{\xi \geqslant 1\} \\ &\leqslant 1 - \mathbb{P}\{\xi \geqslant 1\} + e^{-1}\mathbb{P}\{\xi \geqslant 1\},\end{aligned}$$

so that

$$\mathbb{P}\{\xi \geqslant 1\} < \frac{1 - e^{-\delta}}{1 - e^{-1}}$$

and

$$\mathbb{E}[\exp(-\xi)\mathbb{1}\{\xi < 1\}] > \frac{e^{-\delta} - e^{-1}}{1 - e^{-1}}.$$

Therefore (13.2) will hold when

$$1 + \frac{e^{-1}(1 - e^{-\delta})}{e^{-\delta} - e^{-1}} = \frac{e^{-\delta}(1 - e^{-1})}{e^{-\delta} - e^{-1}} \leqslant \left(\frac{e^{-\delta} - e^{-1}}{1 - e^{-1}}\right)^{\delta-1}$$

or, after some rearrangement, when

$$e^{-\delta} \leqslant \left(\frac{e^{-\delta} - e^{-1}}{1 - e^{-1}} \right)^{\delta}$$

or, equivalently,

$$e^{-1} \leqslant \frac{e^{-\delta} - e^{-1}}{1 - e^{-1}}.$$

This is certainly possible by taking δ sufficiently small. Numerically, the upper bound on satisfactory values of δ given by this argument is approximately 0.51012.

Now consider the second inequality in the statement of the lemma. Recall that $-\log(1-u) \geqslant u$ for $0 \leqslant u < 1$. Moreover, there is a constant $0 < \beta < 1$ such that $\exp(-u) \leqslant 1 - \beta u$ for $0 \leqslant u \leqslant 1$. We have

$$\begin{aligned}
(-\log \mathbb{E}[\exp(-\xi)]) \wedge 1 &= -\log(\mathbb{E}[\exp(-\xi)] \vee e^{-1}) \\
&\geqslant -\log(\mathbb{E}[\exp(-\xi) \vee e^{-1}]) \\
&= -\log(\mathbb{E}[\exp(-(\xi \wedge 1)]) \\
&\geqslant 1 - \mathbb{E}[\exp(-(\xi \wedge 1)] \\
&\geqslant \beta \mathbb{E}[\xi \wedge 1],
\end{aligned}$$

where we used Jensen's inequality for the first inequality. □

Acknowledgments: This work commenced while the authors were attending a symposium the Institut Mittag-Leffler of the Royal Swedish Academy of Sciences to honor the scientific work of Olav Kallenberg. The authors are grateful to the referee for careful reading of the manuscript, for drawing the authors' attention to the results from the theory of Delphic semigroups and for the encouragement to extend the setting from compact to general metric measure spaces.

References

[AFDF00] S. Avgustinovich and D. Fon-Der-Flaass, *Cartesian products of graphs and metric spaces*, European J. Combin. **21** (2000), no. 7, 847–851. MR 1787898 (2001h:54012)

[And94] George E. Andrews, *Number theory*, Dover Publications, Inc., New York, 1994, Corrected reprint of the 1971 original [Dover, New York; MR0309838 (46 #8943)]. MR 1298627

[BCR84] Christian Berg, Jens Peter Reus Christensen, and Paul Ressel, *Harmonic analysis on semigroups*, Graduate Texts in Mathematics, vol. 100, Springer-Verlag, New York, 1984, Theory of positive definite and related functions. MR 747302 (86b:43001)

[Bea85] Bernard Beauzamy, *Introduction to Banach spaces and their geometry*, second ed., North-Holland Mathematics Studies, vol. 68, North-Holland Publishing Co., Amsterdam, 1985, Notas de Matemática [Mathematical Notes], 86. MR 889253 (88f:46021)

[BP95] Ehrhard Behrends and Jan Pelant, *The cancellation law for compact Hausdorff spaces and vector-valued Banach-Stone theorems*, Arch. Math. (Basel) **64** (1995), no. 4, 341–343. MR 1319005 (96h:46046)

[Cli38] A. H. Clifford, *Arithmetic and ideal theory of commutative semigroups*, Ann. of Math. (2) **39** (1938), no. 3, 594–610. MR 1503427

[CP61] A. H. Clifford and G. B. Preston, *The algebraic theory of semigroups. Vol. I*, Mathematical Surveys, No. 7, American Mathematical Society, Providence, R.I., 1961. MR 0132791 (24 #A2627)

[Dav69] Rollo Davidson, *More Delphic theory and practice*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **13** (1969), 191–203. MR 0263128 (41 #7733)

[DMZ08] Youri Davydov, Ilya Molchanov, and Sergei Zuyev, *Strictly stable distributions on convex cones*, Electron. J. Probab. **13** (2008), no. 11, 259–321. MR 2386734 (2009e:60036)

[DMZ11] ———, *Stability for random measures, point processes and discrete semigroups*, Bernoulli **17** (2011), no. 3, 1015–1043. MR 2817615 (2012h:60160)

[dR52] Georges de Rham, *Sur la reductibilité d'un espace de Riemann*, Comment. Math. Helv. **26** (1952), 328–344. MR 0052177 (14,584a)

[Dud02] R. M. Dudley, *Real analysis and probability*, Cambridge Studies in Advanced Mathematics, vol. 74, Cambridge University Press, Cambridge, 2002, Revised reprint of the 1989 original. MR 1932358 (2003h:60001)

[DVJ03] D. J. Daley and D. Vere-Jones, *An introduction to the theory of point processes. Vol. I*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2003, Elementary theory and methods. MR 1950431 (2004c:60001)

[DVJ08] ———, *An introduction to the theory of point processes. Vol. II*, second ed., Probability and its Applications (New York), Springer, New York, 2008, General theory and structure. MR 2371524 (2009b:60150)

[EK86] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, 1986, Characterization and convergence. MR 838085 (88a:60130)

[FL08] Thomas Foertsch and Alexander Lytchak, *The de Rham decomposition theorem for metric spaces*, Geom. Funct. Anal. **18** (2008), no. 1, 120–143. MR 2399098 (2010c:53061)

[Fou71] Gilles Fournier, *On a problem of S. Ulam*, Proc. Amer. Math. Soc. **29** (1971), 622. MR 0278262 (43 #3993)

[GPW09] Andreas Greven, Peter Pfaffelhuber, and Anita Winter, *Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees)*, Probab. Theory Related Fields **145** (2009), no. 1-2, 285–322. MR 2520129 (2011c:60008)

[Gri01] P. A. Grillet, *Commutative semigroups*, Advances in Mathematics (Dordrecht), vol. 2, Kluwer Academic Publishers, Dordrecht, 2001. MR 2017849 (2004h:20089)

[Gro99] Misha Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Progress in Mathematics, vol. 152, Birkhäuser Boston Inc., Boston, MA, 1999, Based on the 1981 French original [MR0682063 (85e:53051)], With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates. MR 1699320 (2000d:53065)

[Gru70] Peter Gruber, *Zur Charakterisierung konvexer Körper. Über einen Satz von Rogers und Shephard. II*, Math. Ann. **184** (1970), 79–105. MR 0256266 (41 #922)

[GW84] R. L. Graham and P. M. Winkler, *Isometric embeddings of graphs*, Proc. Nat. Acad. Sci. U.S.A. **81** (1984), no. 22, Phys. Sci., 7259–7260. MR 768159 (85k:05040)

[GW85] ———, *On isometric embeddings of graphs*, Trans. Amer. Math. Soc. **288** (1985), no. 2, 527–536. MR 776391 (86f:05055b)

[Her94] I. Herburt, *There is no cancellation law for metric products*, Intuitive geometry (Szeged, 1991), Colloq. Math. Soc. János Bolyai, vol. 63, North-Holland, Amsterdam, 1994, pp. 155–160. MR 1383621 (97d:51023)

[Hol69] W. Holsztyński, *Lattices with real numbers as additive operators*, Dissertationes Math. Rozprawy Mat. **62** (1969), 86. MR 0269560 (42 #4455)

[Imr71] Wilfried Imrich, *Über das schwache Kartesische Produkt von Graphen*, J. Combinatorial Theory Ser. B **11** (1971), 1–16. MR 0280401 (43 #6121)

[Kal02] Olav Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR 1876169 (2002m:60002)

[Kap48] Irving Kaplansky, *Lattices of continuous functions. II*, Amer. J. Math. **70** (1948), 626–634. MR 0026240 (10,127a)

[Kec95] Alexander S. Kechris, *Classical descriptive set theory*, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, New York, 1995. MR 1321597 (96e:03057)

[Ken68] David G. Kendall, *Delphic semi-groups, infinitely divisible regenerative phenomena, and the arithmetic of p-functions*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **9** (1968), 163–195. MR 0229746 (37 #5320)

[Mau81] R. Daniel Mauldin (ed.), *The Scottish Book*, Birkhäuser Boston, Mass., 1981, Mathematics from the Scottish Café, Including selected papers presented at the Scottish Book Conference held at North Texas State University, Denton, Tex., May 1979. MR 666400 (84m:00015)

[Mil70] Donald J. Miller, *Weak cartesian product of graphs*, Colloq. Math. **21** (1970), 55–74. MR 0274327 (43 #92)

[Mos92] Maria Moszyńska, *On the uniqueness problem for metric products*, Glas. Mat. Ser. III **27(47)** (1992), no. 1, 145–158. MR 1230118 (94h:54035)

[Pie59] R. S. Pierce, *Translation lattices*, Mem. Amer. Math. Soc. no. **32** (1959), 66 pp. (1959). MR 0107613 (21 #6338)

[Sab60] Gert Sabidussi, *Graph multiplication*, Math. Z. **72** (1959/1960), 446–457. MR 0209177 (35 #80)

[SH11] Willi-Hans Steeb and Yorick Hardy, *Matrix calculus and Kronecker product*, second ed., World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2011, A practical approach to linear and multilinear algebra. MR 2807612 (2012e:15001)

[Stu06] Karl-Theodor Sturm, *On the geometry of metric measure spaces. I*, Acta Math. **196** (2006), no. 1, 65–131. MR 2237206 (2007k:53051a)

[Tar92] Claude Tardif, *Prefibers and the Cartesian product of metric spaces*, Discrete Math. **109** (1992), no. 1-3, 283–288, Algebraic graph theory (Leibnitz, 1989). MR 1192390 (93j:54019)

[Ver98] A. M. Vershik, *The universal Uryson space, Gromov's metric triples, and random metrics on the series of natural numbers*, Uspekhi Mat. Nauk **53** (1998), no. 5(323), 57–64. MR 1691182 (2000b:53055)

[Ver03] ———, *Random and universal metric spaces*, Fundamental mathematics today (Russian), Nezavis. Mosk. Univ., Moscow, 2003, pp. 54–88. MR 2072636 (2005g:60018)

[Ver04] ———, *Random metric spaces and universality*, Uspekhi Mat. Nauk **59** (2004), no. 2(356), 65–104. MR 2086637 (2005g:60019)

[Viz63] V. G. Vizing, *The cartesian product of graphs*, Vyčisl. Sistemy No. **9** (1963), 30–43. MR 0209178 (35 #81)

[Wal87] James W. Walker, *Strict refinement for graphs and digraphs*, J. Combin. Theory Ser. B **43** (1987), no. 2, 140–150. MR 904402 (89a:05128)

[Win87] Peter Winkler, *The metric structure of graphs: theory and applications*, Surveys in combinatorics 1987 (New Cross, 1987), London Math. Soc. Lecture Note Ser., vol. 123, Cambridge Univ. Press, Cambridge, 1987, pp. 197–221. MR 905281 (88h:05090)

[Zer01] E. Santillan Zeron, *Cancellation laws in topological products*, Houston J. Math. **27** (2001), no. 1, 67–74. MR 1843913 (2002d:54003)

Department of Statistics #3860, 367 Evans Hall, University of California, Berkeley, CA 94720-3860, USA

E-mail address: evans@stat.berkeley.edu

University of Bern, Institute of Mathematical Statistics and Actuarial Science, Sidlerstrasse 5, CH-3012 Bern, SWITZERLAND

E-mail address: ilya.molchanov@stat.unibe.ch