Albert-Ludwigs Universität Freiburg im Breisgau
Fakultät für Mathematik und Physik
Institut für Mathematische Stochastik

Compact metric measure spaces and Λ-coalescents

Diploma Thesis

by

Florian Holger Biehler

under the Supervision of
Prof. Dr. P. Pfaffelhuber

May 2011

Abstract. We study topological properties of the space of compact metric measure spaces equipped with the Gromov-weak topology. In particular, we show a characterization of relative compactness and give necessary and sufficient conditions on a metric measure space to be (locally) compact.

The latter is used to study random metric measure spaces which arise from Λ-coalescents. Λ-coalescents are stochastic processes, which start with an infinite number of lines and evolve through multiple mergers in an exchangeable setting. We show, that the resulting Λ-coalescent measure tree is compact if and only if the Λ-coalescent comes down from infinite, i.e. only consists of finitely many lines at any positive time. If the Λ-coalescent stays infinite, then the Λ-coalescent measure tree is not even locally compact.

Zusammenfassung. Wir untersuchen topologische Eigenschaften im Raum der kompakten metrischen Maßräume, der mit mit der Gromov-schwachen Topologie ausgestattet ist. Insbesondere zeigen wir eine Charakterisierung der relativen Kompaktheit und geben hinreichende und notwendige Bedingungen an einen metrischen Maßraum, damit dieser (lokal) kompakt ist.

Letzteres findet Anwendung in der Untersuchung von zufälligen metrischen Maßräumen, welche im Kontext der Λ-Koaleszenten auftreten. Λ-Koaleszenten sind stochastische Prozesse, welche mit einer unendlichen Anzahl an Linien starten und sich durch mehrfache Verschmelzungen in einer austauschbaren Umgebung entwickeln. Wir zeigen, dass der daraus resultierende Λ-Koaleszent Maßraum kompakt ist, genau dann wenn der Λ-Koaleszent von unendlich herunterkommt, d.h. zu jeder positiven Zeit nur noch aus endlich vielen Linien besteht. Wenn der Λ-Koaleszent unendlich bleibt, dann ist der Λ-Koaleszent Maßraum nicht einmal mehr lokal kompakt.

Vorab

»Auch der erste Schritt gehört zum Wege.«

[Arthur Schnitzler]

Ich betrachte nachdenklich den Wegweiser, darauf steht in großen Lettern geschrieben »Zukunft«, und er zeigt gleichzeitig in verschiedenste Richtungen. Ich blicke auf den Weg, der schließlich zu dieser Arbeit geführt hat, zurück, denn wenn das eine Wegstück zu Ende geht und ein Neues beginnt, dann denkt man gerne wieder an den Anfang, und vielleicht erkennt man besser, wohin man gehen will, wenn man weiß woher man gekommen ist.

Das Straßenschild am Anfang des Weges kann ich noch deutlich lesen, obwohl es ein langer Weg war, »Eckerstraße«. Man sagt, der Mut stelle sich die Wege kürzer vor [Goethe] und sicherlich auch weniger steinig, Aber man wird von ihm dann doch immer zu deren Ende geführt. Ich sehe mich, wie ich die Grundlagen der Mathematik versuche zu verstehen, wie ich im Hochgefühl gelöster Aufgaben bade und im Frust ungelöster Probleme ertrinke, wie ich mich auf die Vordiplomsprüfungen vorbereite, sie bestehe, unterrichte, ins Ausland gehe und wie ich mich schließlich zu dieser Arbeit entschließe.

Die Faszination, die die Mathematik auf mich ausübt, wollte stehts die Hintergründe dieser begreifen und trieb mich voran. Die Beschäftigung an dieser Arbeit und das damit verbundene von Vorlesungen gelöste Denken ließ mich das bis dahin Gelernte neu und mit einem tieferen Verständnis erfahren und schührte nur noch mehr den Wunsch, Mathematiker zu sein.

Ich kann nicht behaupten, im Laufe dieses Weges der Selbe geblieben zu sein, aber das ist auch gut so, [Klaus Wowereit]. Im Lauf der vergangenen sechs Jahre des Studiums habe ich mir zwar sicherlich einiges an Mathematikwissen aneignen können, die wachsende mathematische Intuition beeinflusste mich selbst im alltäglichen Denken, aber vielmehr lernte ich über mich selbst, wer ich bin und wo mein Platz in dieser Welt ist. Nichts übt den Geist mehr, als das Bemühen, Rätselhaftes zu ergründen, [Goethe], nichts formt den Geist mehr, als das Bemühen, immer besser zu werden, nichts erhöht die Grenze zur Frustration mehr, als ein Studium der Mathematik.

Natürlich gab es viele Höhen und Tiefen auf diesem Weg, die ihre Spuren hinterließen. Unterwegs traf ich viele Menschen, von denen viele mir heute noch sehr wichtig sind, und die mich unterstützten, förderten, forderten oder einfach nur ablenkten von Frust und Arbeit. Allen voran möchte ich Louisa danken, die mir über lange Zeit eine sehr gute Freundin und mehr war. Unsere Diskussionen über diverse Übungsaufgaben stellten stehts mein bisheriges Verständnis in Frage und sie lehrte mich, mich nicht immer ganz so ernst zu nehmen. Ich danke meinen Freunden Andrea, deren Art so sehr anders ist als die meine, aber die ich dennoch oder gerade deswegen sehr schätze, Raija, die immer ein offenes Ohr für mich hat, Heiko, der mir seit nun mehr als fünfzehn Jahren ein guter Freund ist, Kerstin, Corinna, Christian und vielen mehr..
Meinem Vater danke ich für seine finanzielle Unterstützung.

Nicht vergessen möchte ich meine Kommilitonen und Lehrer, die mir nicht nur bei mathematischen Problemen gerne weiter geholfen haben; Heinz, der jeden Vortragenden mit seinen Fragen ins Schwitzen bringt, Bene und Nico, vor allem aber dem Betreuer dieser Arbeit Peter Pfaffelhuber, der sich stets Zeit nahm und mir mit Rat und Tat zur Seite stand.

Mein Blick richtet sich nun wieder voraus auf den Wegweiser. Ich versuche die Augen leicht zusammenkneifend in die Ferne, die sich hinter dem Schild weit erstreckt, zu schauen; vergebens. Dann mache ich den ersten Schritt..

Freiburg im Breisgau, Mai 2011 Florian Holger Biehler

Contents

CHAPTER 1

Introduction

Metric structures arise frequently in probability theory. Common examples are the euclidean space $\mathbb{R}^d$, the space of càdlàg paths, equipped with the Skorohod metric or the space of probability measures, equipped with the Prohorov metric. Also, prominent examples are random tress ([Ald93], [Ber09]). We discuss in $\mathbb{M}$, the space of *measure-preserving isometry classes of metric measure spaces*, a more general model.

The space $\mathbb{M}$ is build as follows. Take a metric space equipped with a probability measure, to keep the possibility of random sampling. It is necessary to choose an arbitrary but fix set and to let the spaces be subsets of it, to avoid problems which arise from the Zermelo-Fraenkel axioms. Then, we build classes by using the equivalence relation of being measure-preserving isometric.

Following the philosophy of Aldous, [Ald93], a sequence of metric measure spaces converges to a limit metric measure space if and only if all randomly sampled finite subspaces converge to the corresponding subspace. The resulting topology is referred to as the *Gromov-weak topology*. It turns out that the space $\mathbb{M}$ equipped with this topology is metrizable by a complete metric, the *Gromov-Prohorov metric*. Actually, $\mathbb{M}$ is also separable, hence Polish and suitable for probability theory, [GPW09].

Results on weak convergence and stochastic process theory require besides an underlying Polish space informations about tightness of probability measures. Therefore, a characterization of the (relatively) compact sets is necessary.

In the context of Riemannian geometry, such foundations have already been laid by Gromov, [Gro99], or Burago et al., [BBI01]. These authors study convergence of (isometry classes of) compact metric spaces. Here, we mainly refer to work of Greven, Pfaffelhuber, Winter, Depperschmidt et al. ([GPW09], [GPW10], [Win07], [DGP11a], [DGP11b]) on the space $\mathbb{M}$.

In this work, we focus on the subspace $\mathbb{M}_c$ of (isometry classes of) *compact metric measure spaces*. Firstly, we give necessary and sufficient conditions on a metric measure space to be (locally) compact. This leads to a characterization of the relative compact sets in $\mathbb{M}_c$ and to a criterion for tightness in the space of probability measures with respect to the Borel-σ-algebra.

A class of random trees is given by *coalescent processes*, where a subset of an infinite number of lines can merge at random and the distance of two points is proportional to their coalescence time. The complexity of this class of processes is properly described by the concepts of Λ-*coalescents*, where any set of lines can merge to a single line. Hence, Λ-coalescents are also known as coalescents with multiple collisions. If any set of lines can merge to several lines at the same time, we speak of simultaneous multiple collisions and of Ξ-coalescents, [Sch00a].

Around 1982 Kingman ([Kin82a], [Kin00], [Kin82b]) studied coalescent processes with binary mergers. In 1999 Pitman, [Pit99], established the notion of Λ-coalescent as a Markovian stochastic process with the state space of all partitions of $\mathbb{N}$. For a generalization of Λ-coalescents to a spatial setting we refer to the work of Limic and Sturm, [LS06]. Since the coalescent theory expanded at a quick pace over the last decades, we refer to a review from N. Berestycki, [Ber09], for a recent survey on coalescent theory.

There are several applications of coalescent theory mainly in theoretical population genetics but also, for instance, in statistical physics. The most known case is Kingman's coalescent for populations with constant size, low offspring variability and in equilibrium. The relevance of the Kingman's coalescent is founded in the relationship to the Moran model and the Wright-Fisher diffusion. But there is a natural need for more general models.

In spin glass models of statistical physics the Bolthausen-Sznitman coalescent turns out as a universal scaling limit, [BS98].

It is important to know, if a coalescent process *comes down from infinity*, i.e. only consists of finitely many lines at any positive time. Schweinsberg gives in [Sch00b] a necessary and sufficient condition on a Λ-coalescent to come down from infinity.

Looking into this matter, we focus on the resulting metric space of a Λ-coalescent, i.e. the Λ-*coalescent measure tree*. This space exists if and only if the Λ-coalescent is free of *dust*, i.e. there are almost surely no singleton blocks, [GPW09].

In this work, we show that the Λ-coalescent comes down from infinity if and only if the corresponding Λ-coalescent measure tree is compact. See also [Eva00] for a similar result for the Kingman's coalescent. Moreover, if the Λ-coalescent stays infinite, then the corresponding metric measure space is not even locally compact.

Outline. The goal of the present work is the following. We concentrate on (locally) compact metric measure spaces and give a characterization of these in Theorem 2.32 and Theorem 2.36, respectively. Moreover, we give in Theorem 2.37 a characterization of relative compactness in $\mathbb{M}_c$. We apply these general results to the Λ-coalescent measure tree in Theorem 3.36.

The rest of the thesis is organized as follows. In Chapter 2 we give a review over the theory of metric measure spaces. Section 2.1 refers to the first definitions and results on $\mathbb{M}$. We show that the space $\mathbb{M}$ equipped with the Gromov-weak topology is Polish, Theorem 2.15, and characterize relative compactness in Theorem 2.18. Then Section 2.2 focuses on $\mathbb{M}_c$.

Section 2.3 treats tightness in the space $\mathcal{M}_1(\mathbb{M})$ and $\mathcal{M}_1(\mathbb{M}_c)$ of probability measures on $\mathbb{M}$ and $\mathbb{M}_c$ with respect to the Borel-σ-algebra, respectively. We extend the characterization of tightness in $\mathcal{M}_1(\mathbb{M})$ as given in Proposition 2.45 to a criterion of tightness in $\mathcal{M}_1(\mathbb{M}_c)$, Proposition 2.47.

In Chapter 3 we study the Λ-coalescent. Firstly, in Section 3.1 we give a basic building block of coalescent theory considering exchangeable random partitions. The main part is proving Kingman's representation, Theorem 3.5. In Section 3.2, we define the Λ-coalescent process and prove Theorem

3.10 which ensures existence and uniqueness of the Λ-coalescent. As an important corollary about dust, i.e. about the singleton blocks, in a coalescent we have Corollary 3.13. In Section 3.3 we show a characterization of the Λ-coalescent coming down from infinity from Schweinsberg, Theorem 3.22. Finally, in Section 3.4 we deal with the connection between the Λ-coalescent and the corresponding metric measure space and prove the existence of the Λ-coalescent measure tree in Theorem 3.33.

To ensure self-containment of this work we recall some vocabulary and results from general metric topology, Section 4.1, from the set of (isometry classes of) compact metric spaces, Section 4.2, and from general probability and measure theory, Section 4.3.

CHAPTER 2

The space of metric measure spaces

The second chapter emphasizes on the set of metric measure spaces $\mathbb{M}$. We give a review on the theory of metric measure spaces as studied by Greven et al., mainly from [GPW09] and [DGP11a]. It turns out that this space equipped with the Gromov-weak topology is metrizable and Polish, Theorem 2.15.

We focus on topological properties. Hence, we characterize compactness and local compactness of a metric measure space in Theorem 2.32 and Theorem 2.36, respectively. Then, we turn to relative compactness in $\mathbb{M}$ with respect to the Gromov-weak topology as characterized in Theorem 2.18. This extends to Theorem 2.37, which gives a characterization of relative compactness in $\mathbb{M}_c$, the space of compact metric measure spaces.

Finally, in Section 2.3, we study tightness in the space of probability measures on the Borel-σ-algebra of $\mathbb{M}$ and $\mathbb{M}_c$, respectively.

2.1 Metric measure spaces

In this section we give a review on the theory of random metric measure spaces. We introduce the space of metric measure spaces $\mathbb{M}$ and recall some results due to Greven, Pfaffelhuber et al., [GPW09].

Firstly, we define the Gromov-weak topology on $\mathbb{M}$ and state some equivalent characterizations, since there is a need for several approaches for our proofs. We then show equivalence in Theorem 2.26. In addition, we show in Theorem 2.15 that $\mathbb{M}$ equipped with the Gromov-weak topology is a Polish space. The main part here is Theorem 2.18, which gives a characterization of relative compactness with respect to the Gromov-weak topology.

As usual, given a topological space $(X, \mathscr{O})$, we denote by $\mathscr{M}_f(X)$ the space of all finite measures on the Borel-σ-algebra $\mathscr{B}(X)$. In particular, $\mathscr{M}_1(X)$ denotes the space of probability measures on $\mathscr{B}(X)$. The *support* *supp*(μ) of $\mu \in \mathscr{M}_1(X)$ is the smallest closed set $X_\mu \subset X$ such that

$$\mu(X \setminus X_\mu) = 0.$$

The *push forward* of μ under a measurable map φ from X into another metric space (Z, r_Z) is the probability measure $\varphi_*\mu \in \mathscr{M}_1(Z)$ defined for all $A \in \mathscr{B}(Z)$ by

$$\varphi_*\mu(A) := \mu(\varphi^{-1}(A)).$$

We denote by $\Rightarrow$ weak convergence in $\mathscr{M}_1(X)$ and by $\mathbb{P}[\cdot]$ the expectation operator.

Definition 2.1 (Metric measure space)**.** Fix any set $\mathscr{R}$. A *metric measure space* is a complete and separable metric space (X, r), where $X \subset \mathscr{R}$, which is equipped with a probability measure μ. We write $\mathbb{M}$ for the space of measure-preserving isometry classes of metric measure spaces, where we say that (X, r, μ) and (X', r', μ') are *measure-preserving isometric* if there exists an isometry φ between the support of μ on (X, r) and of μ' on (X', r') such that $\mu' = \varphi_*\mu$. It is clear that the property of being measure-preserving isometric is an equivalence relation. We abbreviate $\mathcal{X} = \overline{(X, r, \mu)} = (X, r, \mu)$ for a whole measure-preserving isometry class whenever no confusion seems to be possible.

Remark 2.2 (On Zermelo-Fraenkel axioms). One may wonder why we fix in the definition of metric measure spaces a arbitrary set $\mathscr{R}$. We have to be careful to deal with sets in the sense of Zermelo-Fraenkel axioms. Without the restriction $X \subset \mathscr{R}$, $\mathbb{M}$ is not a set, not even a class (in the sense, that elements of a class are sets). It just does not exist. The set $\mathscr{R}$ may be arbitrary but to keep things easy one can think of $\mathscr{R} = \mathbb{R}$.

Following [GPW09], we equip $\mathbb{M}$ with the following topology. A sequence of metric measure spaces converges if and only if all finite subspaces sampled by the measure sitting on the corresponding metric space converge. Firstly, we define a map which is invariant under measure-preserving isometries. For a metric space (X, r) let

$$R^{(X,r)} : \begin{cases} X^{\mathbb{N}} & \to \mathbb{R}_+^{\binom{\mathbb{N}}{2}} \\ ((x_i)_{i\in\mathbb{N}}) & \mapsto (r(x_i, x_j))_{1\leq i<j} \end{cases}$$

be the map which sends a sequence of points in X to its distance matrix . Moreover, for a metric measure space (X, r, μ), we define the distance matrix distribution by

$$\nu^{(X,r,\mu)} := (R^{(X,r)})_* \mu^{\otimes\mathbb{N}} \in \mathscr{M}_1(\mathbb{R}_+^{\binom{\mathbb{N}}{2}}),$$

where $\mu^{\otimes\mathbb{N}}$ is the infinite product measure of μ with respect to the product σ-field on $\mathbb{R}_+^{\binom{\mathbb{N}}{2}}$.

Definition 2.3 (Distance matrix distribution)**.** The *distance matrix distribution* $\nu^{\mathcal{X}}$ of $\mathcal{X} \in \mathbb{M}$ is the distance matrix distribution $\nu^{(X,r,\mu)}$ of an arbitrary representative $(X, r, \mu) \in \mathcal{X}$. Note that $\nu^{(X,r,\mu)}$ depends on (X, r, μ) only through its measure-preserving isometry class. Therefore $\nu^{\mathcal{X}}$ is well-defined.

Remark 2.4. Note that, if we define

$$R_n^{(X,r)} : \begin{cases} X^{n} & \to \mathbb{R}_+^{\binom{n}{2}} \\ ((x_i)_{1\leq i\leq n}) & \mapsto (r(x_i, x_j))_{1\leq i<j\leq n} \end{cases}$$

and $\nu_n^{(X,r,\mu)} := (R_n^{(X,r)})_* \mu^{\otimes n} \in \mathscr{M}_1(\mathbb{R}_+^{\binom{n}{2}})$, then we get the *finite distance matrix distribution* $\nu_n^{\mathcal{X}}$. In addition, let

$$\pi_n^{n+1} : \mathbb{R}_+^{\binom{n+1}{2}} \to \mathbb{R}_+^{\binom{n}{2}}$$

be the restriction operator, forgetting the last column and row. Then we have that

$$\nu_n^{\mathcal{X}} = (\pi_n^{n+1})_* \nu_{n+1}^{\mathcal{X}}.$$

Moreover, since $\mathbb{R}_+^{\binom{\mathbb{N}}{2}}$ is Polish, Proposition 4.2, $\nu^{\mathcal{X}}$ appears as the projective limit of the measures $\nu_n^{\mathcal{X}}$, i.e.

$$\nu^{\mathcal{X}} = \varprojlim_{n\to\infty} \nu_n^{\mathcal{X}}.$$

See also Theorem 4.31 and Definition 4.32.

Remark 2.5 (Gromov's reconstruction theorem). By Gromov's reconstruction theorem [Gro99, $3\frac{1}{2}$.5. or $3\frac{1}{2}$.7.] metric measure spaces are uniquely determined by their distance matrix distribution. If $\nu^{\mathcal{X}} = \nu^{\mathcal{X}'}$ for $\mathcal{X} = (X, r, \mu)$ and $\mathcal{X}' = (X', r', \mu')$ in $\mathbb{M}$, then there exists equidistributed sequences $(x_i)_{i\in\mathbb{N}}$ and $(x'_i)_{i\in\mathbb{N}}$ in X and X', respectively, such that for all $i, j \in \mathbb{N}$, $r(x_i, x_j) = r'(x'_i, x'_j)$. The map $x_i \mapsto x'_i$ is isometric and can be continuously extended to an isometry $\psi : X \to X'$. Since these sequences are equidistributed, this isometry sends μ to μ'.

We base our notion of convergence in $\mathbb{M}$ on the convergence of distance matrix distributions.

Definition 2.6 (Gromov-weak topology, version 1)**.** A sequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$ in $\mathbb{M}$ is said to *converge Gromov-weakly* to $\mathcal{X}$ in $\mathbb{M}$ if and only if

$$\nu^{\mathcal{X}_n} \xRightarrow{n\to\infty} \nu^{\mathcal{X}}$$

in the weak topology on $\mathscr{M}_1(\mathbb{R}_+^{\binom{\mathbb{N}}{2}})$. We call the corresponding topology $\mathscr{O}_{\mathbb{M}}$ on $\mathbb{M}$ the *Gromov-weak topology* .

There is a characterization of the Gromov-weak convergence, Lemma 2.10. Firstly, we introduce polynomials on $\mathbb{M}$.

Definition 2.7 (Polynomials)**.** A function $\Phi = \Phi^{n,\phi} : \mathbb{M} \to \mathbb{R}$ is called a *polynomial* (of degree n with respect to the test function ϕ) on $\mathbb{M}$ if and only if there exists a bounded continuous function $\phi : [0,\infty)^{\binom{n}{2}} \to \mathbb{R}$ such that

$$\Phi((X, r, \mu)) = \int \mu^{\otimes n}(d(x_1, ..., x_n))\phi\Big((r(x_i, x_j))_{1\le i<j\le n}\Big),$$

where $\mu^{\otimes n}$ is the n-fold product measure of μ. We denote by $\mathscr{A}$ the algebra of all polynomials on $\mathbb{M}$..

Remark 2.8. Note that for $\mathcal{X} \in \mathbb{M}$ and a polynomial Φ we have that

$$\Phi(\mathcal{X}) = \int \nu^{\mathcal{X}}(d(r_{i,j})_{1\le i<j})\phi\Big((r_{i,j})_{1\le i<j\le n}\Big). \tag{2.1}$$

Lemma 2.9 (Polynomials separate points). *The algebra $\mathscr{A}$ of all polynomials is a rich enough class to determine a metric measure space, i.e. $\mathscr{A}$ separates points in $\mathbb{M}$ (Definition 4.17).*

Proof. Let $\mathcal{X} = (X, r, \mu)$ and $\mathcal{X}' = (X', r', \mu')$ be some metric measure spaces with $\Phi(\mathcal{X}) = \Phi(\mathcal{X}')$ for all $\Phi \in \mathscr{A}$. Since by Proposition 4.20 the algebra $\{\phi \in \mathcal{C}_b(\mathbb{R}^{\binom{n}{2}}), n \in \mathbb{N}\}$ is separating in $\mathscr{M}_1(\mathbb{R}^{\binom{n}{2}})$, we have that $\nu^{\mathcal{X}} = \nu^{\mathcal{X}'}$ by equation (2.1). Therefore, by Remark 2.5, we have $\mathcal{X} = \mathcal{X}'$. □

Lemma 2.10 (Gromov-weak topology, version 2). *A sequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$ in $\mathbb{M}$ converges Gromov-weakly to $\mathcal{X}$ in $\mathbb{M}$ if and only if $\mathcal{X}_n$ converges weakly to $\mathcal{X}$ with respect to the algebra of polynomials $\mathscr{A}$, i.e. $\Phi(\mathcal{X}_n)$ converges to $\Phi(\mathcal{X})$ in $\mathbb{R}$, for all polynomials $\Phi \in \mathscr{A}$.*

In [GPW09, Theorem 5], it is shown that this is equivalent to our definition of Gromov-weak convergence. Moreover, the Gromov-weak topology is metrizable by the Gromov-Prohorov metric d_{GPr}. We give a short proof at the end of Section 2.1 in Theorem 2.26. Firstly, we recall the definition of the Gromov-Prohorov distance but we leave out the details that it is indeed a metric. For that see for example [GPW09, Lemma 5.4].

Definition 2.11 (Gromov-Prohorov metric). The *Gromov-Prohorov distance* between two metric measure spaces $\mathcal{X} = (X, r_X, \mu_X)$ and $\mathcal{Y} = (Y, r_Y, \mu_Y)$ in $\mathbb{M}$ is defined by

$$d_{GPr}(\mathcal{X}, \mathcal{Y}) := \inf_{(\varphi_X, \varphi_Y, Z)} d_{Pr}^{(Z, r_Z)}((\varphi_X)_* \mu_X, (\varphi_Y)_* \mu_Y),$$

where the infimum is taken over all isometric embeddings φ_X and φ_Y from X and Y, respectively, into some common metric space (Z, r_Z).

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

$$d_{Pr}^{(Z, r_Z)}(\mu_1, \mu_2) := \inf\{\epsilon > 0 : \mu_1(F) \leq \mu_2(F^\epsilon) + \epsilon \text{ for all closed } F \subset Z\},$$

where

$$F^\epsilon := \{z \in Z : r_Z(z, F) < \epsilon\} \text{ is the } \epsilon\text{-}enlargement \text{ of } F.$$

Remark 2.12. Note that one can choose in the definition of the Gromov-Prohorov metric for a common metric space (Z, r_Z) the disjoint union $X \sqcup Y$. Therefore, we can also write

$$d_{GPr}(\mathcal{X}, \mathcal{Y}) := \inf_{r_{X\sqcup Y}} d_{Pr}^{(X\sqcup Y, r_{X\sqcup Y})}((\varphi_X)_*\mu_X, (\varphi_Y)_*\mu_Y).$$

We need the following two lemma.

Lemma 2.13. *Fix a sequence* $(\epsilon_n)_{n\in\mathbb{N}}$ *in* $(0,1)$*. Let* $\mathcal{X}_1 = (X_1, r_1, \mu_1)$, $\mathcal{X}_2 = (X_2, r_2, \mu_2), \ldots \in \mathbb{M}$*. Then,*

$$d_{GPr}(\mathcal{X}_n, \mathcal{X}_{n+1}) < \epsilon_n \tag{2.2}$$

iff there exists a complete and separable metric space (Z, r_Z) *and isometric embeddings* $\varphi, \varphi_1, \varphi_2, \ldots$ *from* $\mathcal{X}, \mathcal{X}_1, \mathcal{X}_2, \ldots$ *into* (Z, r_Z)*, respectively, such that*

$$d_{Pr}^{(Z,r_Z)}((\varphi_n)_*\mu_n, (\varphi_{n+1})_*\mu_{n+1}) < \epsilon_n.$$

Proof. We give a sketch of the proof following [GPW09, Lemma 5.7].
Assume first that (2.2) holds. Let $Y_n = X_n \sqcup X_{n+1}$ be the disjoint union. Then by Remark 2.12, there is a metric r_{Y_n} such that

$$d_{Pr}^{(Y_n,r_{Y_n})}((\varphi_n)_*\mu_n, (\varphi_{n+1})_*\mu_{n+1}) < \epsilon_n,$$

where φ_n and φ_{n+1} are the canonical embeddings. We define a correspondence R_n, Definition 4.3, by

$$R_n := \{(x_n, x_{n+1}) \in X_n \times X_{n+1} : r_{Y_n}(\varphi_n(x_n), \varphi_{n+1}(x_{n+1})) < \epsilon_n\}.$$

Then, R_n is not empty. Hence, by Remark 4.4, we have metrics $r_{Y_n}^{R_n}$, such that

$$d_{Pr}^{(Y_n,r_{Y_n}^{R_n})}((\varphi_n)_*\mu_n, (\varphi_{n+1})_*\mu_{n+1}) \leq \epsilon_n.$$

Now, let $Z_n := \bigsqcup_{k=1}^n X_k$ and ψ_k^n be the isometric embeddings from X_k to Z_n which arise from the canonical embeddings. Then, again by Remark 4.4, we define using the correspondence given by

$$\tilde{R}_n := \{(z_n, x_{n+1}) \in Z_n \times X_{n+1} : ((\psi_k^n)^{-1}(z_n), x_{n+1}) \in R_n\}$$

metrics $r_{Z_{n+1}}^{\tilde{R}_n}$ on Z_{n+1}. Taking the limit $n \to \infty$ and then the completion, we obtain a separable and complete metric space (Z, r_Z) and isometric embeddings ψ_n from X_n to Z. Finally, the restriction of r_Z to Y_n is isometric to $(Y_n, r_{Y_n}^{R_n})$. Hence,

$$d_{Pr}^{(Z,r_Z)}((\psi_n)_*\mu_{X_n}, (\psi_{n+1})_*\mu_{X_{n+1}}) \leq \epsilon_n$$

and we are done. The converse direction is clear by definition. □

Lemma 2.14. *Let* $\mathcal{X} = (X, r, \mu), \mathcal{X}_1 = (X_1, r_1, \mu_1), \ldots \in \mathbb{M}$. *Then*

$$d_{GPr}(\mathcal{X}_n, \mathcal{X}) \xrightarrow{n\to\infty} 0, \tag{2.3}$$

iff there exists a complete and separable metric space (Z, r_Z) *and isometric embeddings* $\varphi, \varphi_1, \varphi_2, \ldots$ *from* $\mathcal{X}, \mathcal{X}_1, \mathcal{X}_2, \ldots$ *into* (Z, r_Z)*, respectively, such that*

$$d_{Pr}^{(Z,r_Z)}((\varphi_n)_*\mu_n, (\varphi)_*\mu) \xrightarrow{n\to\infty} 0. \tag{2.4}$$

Proof. As in [GPW09, Lemma 5.8], we can follow the same line of arguments as in the proof of Lemma 2.13. Just the metric r_Z extending the metrics $r, r_1, r_2, \ldots$ is built on correspondences between X and X_n. □

The following theorem from [GPW09, Proposition 5.6] ensures that the state space $\mathbb{M}$ is suitable for probability theory.

Theorem 2.15. *The space* $(\mathbb{M}, d_{GPr})$ *is Polish.*

Proof. To get completeness, it suffices to show that a Cauchy-sequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$ has a convergent subsequence. Take therefore a subsequence $(\mathcal{Y}_n)_{n\in\mathbb{N}}$ of $(\mathcal{X}_n)_{n\in\mathbb{N}}$, $\mathcal{Y}_n = (Y_n, r_n, \mu_n)$, such that

$$d_{GPr}(\mathcal{Y}_n, \mathcal{Y}_{n+1}) \leq 2^{-n}.$$

By Lemma 2.13, there is a complete and separabel metric space (Z, r_Z) and isometric embeddings φ_n from Y_n to Z such that $((\varphi_n)_*\mu_n)_{n\in\mathbb{N}}$ is a Cauchy-sequence in $\mathscr{M}_1(Z)$. Since $(\mathscr{M}_1(Z), d_{Pr})$ is complete, the sequence $((\varphi_n)_*\mu_n)_{n\in\mathbb{N}}$ converges to some $\tilde{\mu} \in \mathscr{M}_1(Z)$. Letting $\mathcal{Z} = (Z, r_Z, \tilde{\mu})$, we find

$$d_{GPr}(\mathcal{Y}_n, \mathcal{Z}) \xrightarrow{n\to\infty} 0.$$

To get separability, we partly follow the proof of [EK86, Theorem 3.2.2] or rather [GPW09, Proposition 5.6]. Let $\mathcal{X} = (X, r, \mu) \in \mathbb{M}$ and $\epsilon > 0$. Then there is $\mathcal{X}^\epsilon = (X, r, \mu^\epsilon) \in \mathbb{M}$ such that μ^ϵ is a finitely supported atomic measure on $\mathscr{B}(X)$ and $d_{Pr}(\mu, \mu^\epsilon) < \epsilon$. Then, we find that $d_{GPr}(\mathcal{X}, \mathcal{X}^\epsilon) < \epsilon$. Since $\mathcal{X}^\epsilon$ is just a »finite metric space« it can be approximated in the Gromov-Prohorov metric by finite metric spaces with rational mutual distances and weights. The set of isometry classes of finite metric spaces with rational edge-lengths is countable. Hence, we are done. □

Now, we focus on compactness in $\mathbb{M}$. Roughly speaking, a subset of $\mathbb{M}$ is relatively compact iff the corresponding sequence of probability measures puts most of their mass on subspaces of a uniformly bounded diameter and if the contribution of points not carrying much mass in their vicinity is small, [GPW09]. These two criteria lead to the following definitions.

Definition 2.16 (Distance distribution, Modulus of mass distribution)**.** Let $\mathcal{X} = (X, r, \mu) \in \mathbb{M}$.

(a) The *distance distribution* is given by $w_{\mathcal{X}} := r_* \mu^{\otimes 2}$, i.e.

$$w_{\mathcal{X}}(\cdot) := \mu^{\otimes 2}(\{(x, x') : r(x, x') \in \cdot\}).$$

(b) For $\delta > 0$, define the *modulus of mass distribution* as

$$v_\delta(\mathcal{X}) := \inf\{\epsilon > 0 : \ \mu(\{x \in X : \ \mu(B_\epsilon(x)) \leq \delta\}) \leq \epsilon\},$$

where $B_\epsilon(x)$ is the open ball with radius ϵ and center x.

Lemma 2.17. *Let $\delta > 0$, $\epsilon > 0$ and $\mathcal{X} = (X, r, \mu) \in \mathbb{M}$. If $v_\delta(\mathcal{X}) < \epsilon$, then there exists $N_\epsilon < \lfloor \frac{1}{\delta} \rfloor$ and points $x_1, \dots, x_{N_\epsilon} \in X$ such that the following holds.*

(a) For $i = 1, \dots, N_\epsilon$ we have $\mu(B_\epsilon(x_i)) > \delta$ and $\mu(\bigcup_{i=1}^{N_\epsilon} B_{2\epsilon}(x_i)) > 1 - \epsilon$.

(b) For all $i, j = 1, \dots, N_\epsilon$ with $i \neq j$ we have $r(x_i, x_j) > \epsilon$.

Proof. We follow the proof given in [GPW09, Lemma 6.9]. By definition of $v_\delta(\cdot)$, there exists $\epsilon' < \epsilon$ for which $\mu(\{x \in X : \mu(B_{\epsilon'}(x)) \leq \delta\}) \leq \epsilon'$. Since $\{x \in X : \mu(B_\epsilon(x)) \leq \delta\} \subset \{x \in X : \mu(B_{\epsilon'}(x)) \leq \delta\}$, we find that

$$\mu(\{x \in X : \mu(B_\epsilon(x)) \leq \delta\}) \leq \mu(\{x \in X : \mu(B_{\epsilon'}(x)) \leq \delta\}) \leq \epsilon' < \epsilon.$$

Let $D := \{x \in X : \mu(B_\epsilon(x)) > \delta\}$, then we have $\mu(D) > 1 - \epsilon$. Take a maximal 2ϵ separated net $\{x_i \in D : i \in I\}$, i.e. $D \subset \bigcup_{i \in I} B_{2\epsilon}(x_i)$ and for all $i \neq j$, $r(x_i, x_j) > 2\epsilon$. By [BBI01, p. 278], such a net exists in every metric space. Since

$$1 \geq \mu(\bigcup_{i \in I} B_\epsilon(x_i)) = \sum_{i \in I} \mu(B_\epsilon(x_i)) \geq | I | \delta,$$

we find that $| I | \leq \lfloor \frac{1}{\delta} \rfloor$. □

The next result from [GPW09] characterizes relative compactness in the topology induced by the Gromov-Prohorov metric on $\mathbb{M}$. We denote by $(\mathbb{X}_c, d_{GH})$ the space of (isometry classes of) compact metric spaces equipped with the Gromov-Hausdorff metric. See Section 4.2 for more.

Theorem 2.18 (Characterization of relative compactness)**.** *Let* $\Gamma \subseteq \mathbb{M}$. *The following conditions are equivalent.*

(a) The family Γ *is relatively compact in the Gromov-Prohorov topology.*

(b) The family $\{w_\mathcal{X} : \mathcal{X} \in \Gamma\}$ *is tight and* $\sup_{\mathcal{X} \in \Gamma} v_\delta(\mathcal{X}) \xrightarrow{\delta \to 0} 0$.

(c) For all $\epsilon > 0$ *there exists* $N_\epsilon \in \mathbb{N}$ *such that for all* $\mathcal{X} \in \Gamma$ *there is a subset* $X_{\epsilon,\mathcal{X}} \subseteq X$ *with*

- c_1) $\mu(X_{\epsilon,\mathcal{X}}) \geq 1 - \epsilon$,
- c_2) $X_{\epsilon,\mathcal{X}}$ *can be covered by at most* N_ϵ *balls of radius* ϵ,
- c_3) $X_{\epsilon,\mathcal{X}}$ *has diameter at most* N_ϵ.

(d) For all $\epsilon > 0$ *and* $\mathcal{X} \in \Gamma$ *there exists a compact subset* $K_{\epsilon,\mathcal{X}} \subset X$ *with*

- d_1) $\mu(K_{\epsilon,\mathcal{X}}) \geq 1 - \epsilon$,
- d_2) *the family* $K_\epsilon := \{K_{\epsilon,\mathcal{X}}, \mathcal{X} \in \Gamma\}$ *is relatively compact in* $(\mathbb{X}_c, d_{GH})$.

We give later a proof following [GPW09, Proposition 7.1]. After Example 2.19, we first introduce the random distance distribution of a given metric measure space, Definition 2.20, and prepare the proof of the theorem with Proposition 2.23.

Example 2.19. The following examples illustrate condition (b) in Theorem 2.18 for relative compactness of a family in $\mathbb{M}$ by counter-examples. The examples (a) and (b) are from [GPW09, Example 2.12].

(a) Consider the sequence of metric measure spaces defined by

$$\mathcal{X}_n := (\{1,2\}, r_n(1,2) = n, \mu_n(\{1\}) = \mu_n(\{2\}) = \frac{1}{2}).$$

A potential Gromov-weak limit object would be a metric measure space with masses $\frac{1}{2}$ within distance infinity. This clearly does not exist. Indeed, the family $\{w_{\mathcal{X}_n} = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_n : n \in \mathbb{N}\}$ is not tight. Hence, $\{\mathcal{X}_n : n \in \mathbb{N}\}$ is not relatively compact.

(b) Consider the sequence of metric measure spaces defined by

$$\mathcal{X}_n := (\{1,\dots,2^n\}, r_n(x,y) := \mathbb{1}_{\{x\neq y\}}, \mu_n := 2^{-n}\sum_{i=1}^{2^n}\delta_i).$$

A potential Gromov-weak limit object would consist of infinitely many points of mutual distance one with a uniform measure. Such a space does not exist. Indeed, we have for $\delta < 2^{-n}$ that $v_\delta(\mathcal{X}_n) = 0$ and for $\delta \geq 2^{-n}$ that $v_\delta(\mathcal{X}_n) = 1$. It follows that for all $\delta > 0$,

$$\sup_{n\in\mathbb{N}} v_\delta(\mathcal{X}_n) = 1.$$

(c) Let r_{eucl} be the euclidean metric and $N(0,n)$ the normal distribution with expectation 0 and variance n. Then, consider the sequence of metric measure spaces defined by

$$\mathcal{X}_n := (\mathbb{R}, r_{eucl}, N(0,n)).$$

A potential Gromov-weak limit object would consist of $(\mathbb{R}, r_{eucl})$ and a normal distribution with infinite variance. Since the set of variances of the family of distributions $\{N(0,n) : n \in \mathbb{N}\}$ is not bounded, the family is not tight. Hence, $\{w_{\mathcal{X}_n} : n \in \mathbb{N}\}$ is not tight.

Definition 2.20 (Random distance distribution)**.** Let $\mathcal{X} = (X, r, \mu) \in \mathbb{M}$. We define the map $r_x : X \to \mathbb{R}_+$ by $r_x(x') := r(x,x')$ and for $x \in X$, let the *distribution of distance* be $\mu_x := (r_x)_*\mu \in \mathscr{M}_1(\mathbb{R}_+)$. Moreover, we define the map $\hat{r} : X \to \mathscr{M}_1(\mathbb{R}_+)$ by $\hat{r}(x) := \mu_x$. Now let

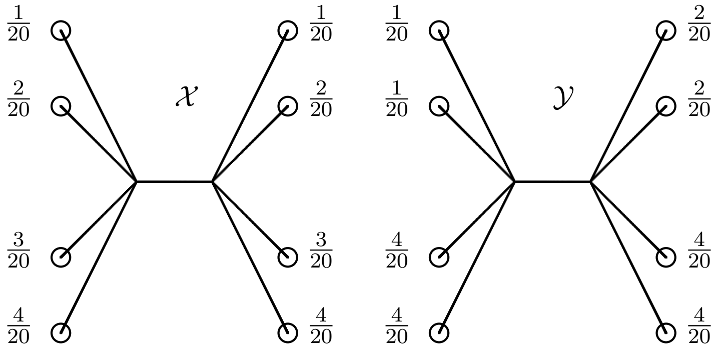
$$\hat{\mu}_{\mathcal{X}} := \hat{r}_*\mu \in \mathscr{M}_1(\mathscr{M}_1(\mathbb{R}_+))$$

be the *random distance distribution* of $\mathcal{X}$.

Remark 2.21. For $\mathcal{X} = (X, r, \mu)$ we have by definition, that

$$\hat{\mu}_{\mathcal{X}}(\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \leq \delta\}) = \mu(\{x \in X : \mu(B_\epsilon(x)) \leq \delta\}).$$

Remark 2.22. The random distance distribution does not characterize the metric measure spaces uniquely. Consider, for example, the two metric measure spaces $\mathcal{X}$ and $\mathcal{Y}$. Both consist of eight points. The distance between two points equals the minimal number of edges one has to cross to come from one point to the other.



The random distance distributions agree, since

$$\hat{\mu}_{\mathcal{X}} = \hat{\mu}_{\mathcal{Y}} = \frac{1}{10}\delta_{\frac{1}{20}\delta_0 + \frac{9}{20}\delta_2 + \frac{1}{2}\delta_3} + \frac{1}{5}\delta_{\frac{1}{10}\delta_0 + \frac{2}{5}\delta_2 + \frac{1}{2}\delta_3} + \frac{3}{10}\delta_{\frac{3}{20}\delta_0 + \frac{7}{20}\delta_2 + \frac{1}{2}\delta_3} + \frac{2}{5}\delta_{\frac{1}{5}\delta_0 + \frac{3}{10}\delta_2 + \frac{1}{2}\delta_3}.$$

But $\mathcal{X}$ and $\mathcal{Y}$ are not measure-preserving isometric, [GPW09].

Proposition 2.23 (Continuity properties)**.** *Let* $\mathcal{X}, \mathcal{X}_1, \mathcal{X}_2, \ldots$ *in* $\mathbb{M}$.

(a) If for all polynomials $\Phi \in \mathscr{A}$, $\Phi(\mathcal{X}_n) \xrightarrow{n\to\infty} \Phi(\mathcal{X})$, *then* $\hat{\mu}_{\mathcal{X}_n} \xRightarrow{n\to\infty} \hat{\mu}_{\mathcal{X}}$.

(b) If for all polynomials $\Phi \in \mathscr{A}, \Phi(\mathcal{X}_n) \xrightarrow{n\to\infty} \Phi(\mathcal{X})$, *then* $w_{\mathcal{X}_n} \xRightarrow{n\to\infty} w_{\mathcal{X}}$.

(c) Let $\delta > 0$. *If* $\hat{\mu}_{\mathcal{X}_n} \xRightarrow{n\to\infty} \hat{\mu}_{\mathcal{X}}$, *then* $\limsup_{n\to\infty} v_\delta(\mathcal{X}_n) \leq v_\delta(\mathcal{X})$.

In order to prove this result, we need the following definition.

Definition 2.24 (k^{th} moment measure)**.** for $\mathcal{X} = (X, r_X, \mu_X)$ and $k \in \mathbb{N}$ we define the k^{th} *moment measure* $\hat{\mu}^k_{\mathcal{X}} \in \mathscr{M}_1(\mathbb{R}^k_+)$ of $\hat{\mu}_{\mathcal{X}}$ by

$$\hat{\mu}^k_{\mathcal{X}}(d(r_1,\ldots,r_k)) := \int \hat{\mu}_{\mathcal{X}}(d\nu)\ \nu^{\otimes k}(d(r_1,\ldots,r_k)).$$

Remark 2.25. By [Kal02, Theorem 16.16], weak convergence of random measures is equivalent to weak convergence of all moment measures.

Proof of Proposition 2.23. We follow the proof given in [GPW09, Proposition 6.6].

(a) By the above remark, it suffices to show that for arbitrary $k \in \mathbb{N}$, $\hat{\mu}^k_{\mathcal{X}_n} \xRightarrow{n\to\infty} \hat{\mu}^k_{\mathcal{X}}$. For $k \in \mathbb{N}$, consider all $\phi \in \mathcal{C}_b(\mathbb{R}_+^{\binom{k+1}{2}})$ which depend on $(r_{ij})_{0\le i<j\le k}$ only through $(r_{0,j})_{1\le j\le k}$, i.e. there exists $\tilde{\phi} \in \mathcal{C}_b(\mathbb{R}^k_+)$ with

$$\phi((r_{ij})_{0\le i<j\le k}) = \tilde{\phi}((r_{0,j})_{1\le j\le k}).$$

Since for any $\mathcal{Y} = (Y, r_Y, \mu_Y)$,

$$\begin{aligned}\int \hat{\mu}^k_{\mathcal{Y}}(d(r_1,\dots,r_k))\tilde{\phi}(r_1,\dots,r_k) &= \int \mu_Y^{\otimes k+1}(d(u_0,u_1,\dots,u_k))\tilde{\phi}(r_Y(u_0,u_1),\dots,r_Y(u_0,u_k)) \\ &= \int \mu_Y^{\otimes k+1}(d(u_0,u_1,\dots,u_k))\phi((r_Y(u_i,u_j))_{0\le i<j\le k}),\end{aligned}$$

it follows by assumption, $\hat{\mu}^k_{\mathcal{X}_n} \xRightarrow{n\to\infty} \hat{\mu}^k_{\mathcal{X}}$. Hence, we are done.

(b) By definition, the distance distribution $w_{\mathcal{X}}$ satisfies

$$w_{\mathcal{X}} = \int_{\mathcal{M}_1(\mathbb{R}_+)} \hat{\mu}_{\mathcal{X}}(d\nu)\ \nu,$$

i.e. $w_{\mathcal{X}}$ equals the first moment measure of $\hat{\mu}_{\mathcal{X}}$. Hence by (a), we are done.

(c) Assume that $\epsilon > 0$ is such that $v_\delta(\mathcal{X}) < \epsilon$. Note that

$$v_\delta(\mathcal{X}) = \inf\{\epsilon > 0 : \hat{\mu}_{\mathcal{X}}\{\nu \in \mathcal{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \le \delta\} \le \epsilon\}.$$

Then we have

$$\hat{\mu}_{\mathcal{X}}\{\nu \in \mathcal{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \le \delta\} < \epsilon.$$

Since the set $\{\nu \in \mathcal{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \le \delta\}$ is closed in $\mathcal{M}_1(\mathbb{R}_+)$, Lemma 4.13, we got by the Portmanteau-Theorem, Theorem 4.12, that

$$\begin{aligned}\limsup_{n\to\infty} \hat{\mu}_{\mathcal{X}_n}\{\nu \in \mathcal{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \le \delta\} \\ \le \hat{\mu}_{\mathcal{X}}\{\nu \in \mathcal{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \le \delta\} < \epsilon.\end{aligned}$$

That is, we have for all but finitely many $n \in \mathbb{N}$, $v_\delta(\mathcal{X}_n) < \epsilon$. Therefore, we find $\limsup_{n\to\infty} v_\delta(\mathcal{X}_n) < \epsilon$. □

Proof of Theorem 2.18. In order to prove this theorem, we anticipate parts of Theorem 2.26, in particular $(a) \Rightarrow (c)$. But of course, this implications do not need the following proof.

Proof. We follow [GPW09, Proposition 7.1].
$(a) \Rightarrow (b)$. We consider a sequence $(\mathcal{X}_n)_{n\in\mathbb{N}} \subset \Gamma$. Since Γ is relatively compact by assumption, there is a convergent subsequence $(\mathcal{X}_{n_k})_{k\in\mathbb{N}}$. By Theorem 2.26 and Proposition 2.23, we have $w_{\mathcal{X}_{n_k}}(\cdot) \xRightarrow{k\to\infty} w_{\mathcal{X}}(\cdot)$. As the sequence was arbitrarily chosen it follows with Proposition 4.16 that the family $\{w_{\mathcal{X}} : \mathcal{X} \in \Gamma\}$ is tight.
The second part is by contradiction. Assume that $v_\delta(\mathcal{X})$ does not converge to 0 uniformly in $\mathcal{X} \in \Gamma$. Therefore, we find $\epsilon > 0$ such that for all $n \in \mathbb{N}$ there exist a sequence $(\delta_n)_{n\in\mathbb{N}}$ converging to 0 and a sequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$ in Γ with

$$v_{\delta_n}(\mathcal{X}_n) \geq \epsilon. \quad (*)$$

By assumption, there is a subsequence $(\mathcal{X}_{n_k})_{k\in\mathbb{N}}$ converging in the Gromov-Prohorov topology to a metric measure space $\mathcal{X} \in \Gamma$. By Theorem 2.26 and Proposition 2.23, $\hat{\mu}_{\mathcal{X}_{n_k}}$ converges weakly to $\hat{\mu}_{\mathcal{X}}$ as $n \to \infty$. Finally, again by Proposition 2.23, we have that $\limsup_{k\to\infty} v_{\delta_k}(\mathcal{X}_{n_k}) = 0$ which contradicts $(*)$.

$(b) \Rightarrow (c)$. By assumption, the family $\{w_{\mathcal{X}} : \mathcal{X} \in \Gamma\}$ is tight, i.e. for all $\epsilon > 0$ there exists $C(\epsilon) \in \mathbb{R}_+$ and $\delta = \delta(\epsilon)$ such that

$$\sup_{\mathcal{X}\in\Gamma} w_{\mathcal{X}}([C(\epsilon), \infty)) < \epsilon,$$

and

$$\sup_{\mathcal{X}\in\Gamma} v_\delta(\mathcal{X}) < \epsilon.$$

Now set

$$X'_{\epsilon,\mathcal{X}} := \{x \in X : \mu_X(B_{C(\frac{\epsilon^2}{4})}(x)) > 1 - \frac{\epsilon}{2}\}.$$

Then it follows, that $\mu_X(X'_{\epsilon,\mathcal{X}}) > 1 - \frac{\epsilon}{2}$. By Lemma 2.17, we can choose for all $\mathcal{X} \in \Gamma$ points $x_1, \ldots, x_{N_\epsilon^{\mathcal{X}}} \in X$ with

$$\begin{aligned} &N_\epsilon^{\mathcal{X}} \le N_\epsilon := \lfloor \frac{1}{\delta(\epsilon/2)} \rfloor, \\ &r_X(x_i, x_j) > \frac{\epsilon}{2} \text{ for } 1 \le i < j \le N_\epsilon^{\mathcal{X}}, \\ &\mu_X(\bigcup_{i=1}^{N_\epsilon^{\mathcal{X}}} B_\epsilon(x_i)) > 1 - \frac{\epsilon}{2}. \end{aligned}$$

Now set

$$X_{\epsilon,\mathcal{X}} := X'_{\epsilon,\mathcal{X}} \cap \bigcup_{i=1}^{N_\epsilon^{\mathcal{X}}} B_\epsilon(x_i).$$

Then $\mu_X(X_{\epsilon,\mathcal{X}}) > 1 - \epsilon$ and $X_{\epsilon,\mathcal{X}}$ can be covered by (at most) N_ϵ balls of radius ϵ. Since the diameter of $X'_{\epsilon,\mathcal{X}}$ is bounded by $4C(\frac{\epsilon^2}{4})$, the same is true for $X_{\epsilon,\mathcal{X}}$.

$(c) \Rightarrow (d)$. Fix $\epsilon > 0$ and set for all $n \in \mathbb{N}$, $\epsilon_n := \epsilon 2^{-(n+1)}$. Then we choose for each $n \in \mathbb{N}$, $N_{\epsilon_n} \in \mathbb{N}$ such that for all $\mathcal{X} \in \Gamma$ there exists a subset $X_{\epsilon_n,\mathcal{X}} \subset X$ such that $c_1)$, $c_2)$ and $c_3)$ holds. Without loss of generality we may assume that all $X_{\epsilon_n,\mathcal{X}}$ are closed. Otherwise we just take their closure. Now, we take for every $\mathcal{X} \in \Gamma$ a compact set $K_{\epsilon_n,\mathcal{X}} \subset X$ with $\mu_X(K_{\epsilon_n,\mathcal{X}}) > 1 - \epsilon_n$. Set

$$K_{\epsilon,\mathcal{X}} := \bigcap_{n=1}^{\infty} (X_{\epsilon_n,\mathcal{X}} \cap K_{\epsilon_n,\mathcal{X}}).$$

As intersection of a compact set with closed sets, $K_{\epsilon,\mathcal{X}}$ is also compact, and $\mu_X(K_{\epsilon,\mathcal{X}}) > 1 - \epsilon$. Let

$$\mathcal{K}_\epsilon := \{K_{\epsilon,\mathcal{X}}, \mathcal{X} \in \Gamma\}.$$

By construction $\mathcal{K}_\epsilon$ is uniformly totally bounded and by Proposition 4.7, $\mathcal{K}_\epsilon$ is relatively compact in $(\mathbb{X}_c, d_{GH})$.

$(d) \Rightarrow (a)$. The proof is in two steps. First assume that all metric spaces (X, r_X) with $\mathcal{X} = (X, r_X, \mu_X) \in \Gamma$ are compact and that the family $\{(X, r_X) : (X, r_X, \mu_X) \in \Gamma\}$ is relatively compact in $(\mathbb{X}_c, d_{GH})$.

That is, we can choose for every sequence in Γ a subsequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$, $\mathcal{X}_n = (X_n, r_{X_n}, \mu_{X_n})$, and a metric space $(X, r) \in \Gamma$ such that

$$d_{GH}(X_n, X) \xrightarrow{n\to\infty} 0.$$

By Lemma 4.5, there is a compact metric space (Z, r_Z) and isometric embeddings $\varphi_X, \varphi_{X_1}, \varphi_{X_2}, \ldots$ of $X, X_1, X_2, \ldots$, respectively, into (Z, r_Z) such that $d_H^{(Z,r_Z)}(\varphi_{X_n}(X_n), \varphi_X(X)) \xrightarrow{n\to\infty} 0$. Since Z is compact the set of measures $\{(\varphi_{X_n})_*\mu_{X_n} : n \in \mathbb{N}\}$ is tight. Therefore, relatively compact with respect to the weak topology, i.e. there is a subsequence $((\varphi_{X_{n_k}})_*\mu_{X_{n_k}})_{k\in\mathbb{N}}$ such that

$$d_{Pr}((\varphi_{X_{n_k}})_*\mu_{X_{n_k}}, \varphi_X(X)) \xrightarrow{k\to\infty} 0.$$

Then by definition, $d_{GPr}(\mathcal{X}_{n_k}, \mathcal{X}) \xrightarrow{k\to\infty} 0$ and we are done.

In the second step, let $\epsilon_n := 2^{-n}$ and fix for every $\mathcal{X} \in \Gamma$ and every $n \in \mathbb{N}$, $x \in K_{\epsilon_n,\mathcal{X}}$. We define

$$\mu_{X,n}(\cdot) := \mu_X(\cdot \cap K_{\epsilon_n,\mathcal{X}}) + (1 - \mu_X(K_{\epsilon_n,\mathcal{X}}))\delta_x(\cdot)$$

and

$$\mathcal{X}^n := (X, r_X, \mu_{X,n}), \quad \Gamma^n := \{\mathcal{X}^n : \mathcal{X} \in \Gamma\}.$$

Then, for all $\mathcal{X} \in \Gamma$ we have

$$d_{GPr}(\mathcal{X}^n, \mathcal{X}) \leq \epsilon_n.$$

Moreover, Γ^n is relatively compact in $(\mathbb{X}_c, d_{GH})$. Therefore, we find a convergent subsequence in Γ^n by the first step.
By a diagonal argument we find a subsequence $(\mathcal{X}_m)_{m\in\mathbb{N}}$ such that $(\mathcal{X}_m^n)_{m\in\mathbb{N}}$ converges for every $n \in \mathbb{N}$ to some metric measure space $\mathcal{Z}^n$. We pick a subsequence such that for all $n \in \mathbb{N}$ and $m \geq n$,

$$d_{GPr}(\mathcal{X}_m^n, \mathcal{Z}^n) \leq \epsilon_m.$$

Then for all $m, m' \geq n$,

$$d_{GPr}(\mathcal{X}_m^n, \mathcal{X}_{m'}^n) \leq 2\epsilon_n.$$

Finally, it follows that $(\mathcal{X}_m)_{m\in\mathbb{N}}$ is a Cauchy-sequence in $(\mathbb{M}, d_{GPr})$ and since $(\mathbb{M}, d_{GPr})$ is complete by Theorem 2.15, this sequence converges and we are done. □

Theorem 2.26 (Versions of the Gromov-weak topology)**.** *Let* $\mathcal{X} = (X, r, \mu)$*,* $\mathcal{X}_1 = (X_1, r_1, \mu_1)$*,* $\mathcal{X}_2 = (X_2, r_2, \mu_2)$*, … be metric measure spaces. Then the following is equivalent.*

(a) The Gromov-Prohorov metric converges, i.e.

$$d_{GPr}(\mathcal{X}_n, \mathcal{X}) \xrightarrow{n\to\infty} 0.$$

(b) The distance matrix distributions converge, i.e.

$$\nu^{\mathcal{X}_n} \xRightarrow{n\to\infty} \nu^{\mathcal{X}}.$$

(c) All polynomials converge, i.e.

$$\forall \Phi \in \mathscr{A}, \Phi(\mathcal{X}_n) \xrightarrow{n\to\infty} \Phi(\mathcal{X}).$$

Proof. We follow the proof given in [GPW09, Theorem 5].

$(a) \Rightarrow (b)$. By Lemma 2.14, there are a complete and separable metric space (Z, r_Z) and isometric embeddings $\varphi, \varphi_1, \varphi_2, \ldots$ from $(X, r), (X_1, r_1), \ldots$ into (Z, r_Z), respectively, such that $((\varphi_n)_* \mu_n)_{n\in\mathbb{N}}$ converges weakly to $\varphi_* \mu$. By the definition of the distance matrix distribution, Definition 2.3, we have that

$$\nu^{\mathcal{X}_n} = (R^{(X_n, r_n)})_* \mu_n^{\otimes\mathbb{N}} = (R^{(Z, r_Z)})_* \left(((\varphi_n)_* \mu_n)^{\otimes\mathbb{N}} \right)$$

converges weakly to

$$(R^{(Z, r_Z)})_* \left((\varphi_* \mu)^{\otimes\mathbb{N}} \right) = (R^{(X,r)})_* \mu^{\otimes\mathbb{N}} = \nu^{\mathcal{X}}$$

as $n \to \infty$.

$(b) \Rightarrow (c)$. This follows easily with the two different representations of the polynomials, Remark 2.8.

$(c) \Rightarrow (a)$. It suffices to show that the sequence $(\mathcal{X}_n)_{n\in\mathbb{N}}$ is relatively compact with respect to the Gromov-weak topology. Since the algebra of polynomials $\mathscr{A}$ separates points in $\mathbb{M}$, Lemma 2.9, this would imply that all limit points coincide and are equal $\mathcal{X}$. Therefore, we have to check the two conditions for relative compactness given in Theorem 2.18.

By Proposition 2.23, the family $\{w_{\mathcal{X}_n} : n \in \mathbb{N}\}$ is tight and we find that $\limsup_{n\to\infty} v_\delta(\mathcal{X}_n) \leq v_\delta(\mathcal{X}) \xrightarrow{\delta\to 0} 0$. □

Remark 2.27. Note that by Theorem 2.26, the above Proposition 2.23 means, that the maps $\mathcal{X} \mapsto \hat{\mu}_{\mathcal{X}}$ and $\mathcal{X} \mapsto w_{\mathcal{X}}$ are continuous with respect to the Gromov-weak topology on $\mathbb{M}$ and the weak topology on $\mathscr{M}_1(\mathscr{M}_1(\mathbb{R}_+))$ or $\mathscr{M}_1(\mathbb{R}_+)$, respectively.

2.2 Compact and locally compact metric measure spaces

We consider the space of measure-preserving isometry classes of *(locally) compact metric measure spaces*, denoted by $\mathbb{M}_c$ and $\mathbb{M}_{lc}$, respectively. A metric measure space $\mathcal{X} = (X, r, \mu)$ is *(locally) compact* if and only if there is $(X, r, \mu) \in \mathcal{X}$ such that (X, r) is (locally) compact. We equip $\mathbb{M}_c$ and $\mathbb{M}_{cl}$ with the Gromov-weak topology induced by $\mathbb{M}$.

We give a characterization of (local) compactness of a metric measure space in Theorem 2.32 and Theorem 2.36, respectively. Then, we extend the characterization of relative compactness in $\mathbb{M}$, Theorem 2.18, to $\mathbb{M}_c$. This attempt results in Proposition 2.37. Furthermore, we reformulate a criterion for relative compactness from [GPW10] in Corollary 2.41.

Remark 2.28 ($\mathbb{M}_c$ is not closed). If $\mathcal{X} = (X, r, \mu)$ is a finite metric measure space then $\mathcal{X} \in \mathbb{M}_c$. Moreover, since elements of $\mathbb{M}$ can be approximated by a sequence of finite metric measure spaces, the subspace $\mathbb{M}_c \subset \mathbb{M}$ is not closed. See also the proof of separability in Theorem 2.15.

In order to formulate the characterization of (locally) compact metric measure spaces we need the following notion of the size of ϵ-separated sets.

Definition 2.29 (Size of ϵ-separated sets)**.** Let $R := (r_{i,j})_{1 \le i < j} \in \mathbb{R}_+^{\binom{\mathbb{N}}{2}}$. For $\epsilon > 0$, define the *minimal size of an ϵ-separated set* by

$$\xi_\epsilon(R) := \max\{N \in \mathbb{N} : \exists k_1 < \cdots < k_N \ (r_{k_i,k_j})_{1 \le i < j \le N} \in (\epsilon, \infty)^{\binom{N}{2}}\}.$$

Recall from Definition 2.3 the distance matrix distribution $\nu^{\mathcal{X}}$ of a metric measure space $\mathcal{X}$.

Lemma 2.30. *Let $\mathcal{X} \in \mathbb{M}$. Then ξ_ϵ is $\nu^{\mathcal{X}}$-almost surely constant and equals*

$$\xi_\epsilon(\mathcal{X}) = \min\{N \in \mathbb{N} : (\pi_{N+1})_* \nu^{\mathcal{X}}((\epsilon, \infty)^{\binom{N+1}{2}}) = 0\},$$

where $\pi^{\mathbb{N}}_{n+1} := \pi_{N+1} : \mathbb{R}_+^{\binom{\mathbb{N}}{2}} \to \mathbb{R}_+^{\binom{N+1}{2}}$ *is the restriction operator. Moreover, for all* $\epsilon > 0$, ξ_ϵ *is lower semi-continuous and hence measurable (Definition 4.21).*

Proof. Take $\mathcal{X} = (X, r, \mu) \in \mathbb{M}$. Let $x_1, x_2, \dots \in X$ be such that

$$\xi_\epsilon((r(x_i, x_j))_{1\le i<j}) = N.$$

Then N is the maximal size of an ϵ-separated set in (X, r), $\mu^{\otimes\mathbb{N}}$-almost surely. The identity is clear by definition.

For lower semi-continuity, by Lemma 4.22, it suffices to show that for all $N \in \mathbb{N}$ the set $A_{\epsilon,N} := \{\mathcal{X} \in \mathbb{M} : \xi_\epsilon(\mathcal{X}) \le N\}$ is closed. For that, let $(\mathcal{X}_n)_{n\in\mathbb{N}}$ be a sequence in $A_{\epsilon,N}$ and $\mathcal{X} \in \mathbb{M}$ such that $\mathcal{X}_n$ converges Gromov-weakly to $\mathcal{X}$. We have to show that $\mathcal{X} \in A_{\epsilon,N}$.

Clearly, we have that for all $n \in \mathbb{N}$, $(\pi_{N+1})_*\upsilon^{\mathcal{X}_n}((\epsilon,\infty)^{\binom{N+1}{2}}) = 0$. Then, we get by the Portmanteau-Theorem that

$$(\pi_{N+1})_*\upsilon^{\mathcal{X}}(B_\epsilon^{N+1}) \le \liminf_{n\to\infty}(\pi_{N+1})_*\upsilon^{\mathcal{X}_n}((\epsilon,\infty)^{\binom{N+1}{2}}) = 0.$$

Hence, $\xi_\epsilon(\mathcal{X}) \le N$. □

Remark 2.31. For a metric space, let N_ϵ be the minimal number of ϵ-balls needed to cover (X, r). Then

$$N_\epsilon \le \xi_\epsilon \le N_{\epsilon/2}.$$

The first inequality is clear, since the points $x_1, \dots, x_{\xi_\epsilon}$ of a maximal ϵ-separated set, serve as centers of ϵ-balls covering (X, r). For the second inequality, we consider the disjoint $(\epsilon/2)$-balls around the points of the maximal ϵ-separated set. Any other set of centers of $\epsilon/2$-balls covering (X, r) must hit each $B_{\epsilon/2}(x_i)$ at least once.

Recall the random distance distribution $\hat{\mu}_{\mathcal{X}}$ of a given metric measure space $\mathcal{X}$ from Definition 2.20. The characterization of compact metric measure spaces reads as follows.

Theorem 2.32 (Characterization of compact metric measure spaces)**.** *Let* $\mathcal{X} \in \mathbb{M}$. *Then the following is equivalent.*

(a) The metric measure space $\mathcal{X}$ *is compact.*

(b) For all $\epsilon > 0$, $\xi_\epsilon(\mathcal{X}) < \infty$.

(c) For all $\epsilon > 0$ *there is* $\delta > 0$ *such that*

$$\hat{\mu}_{\mathcal{X}}(\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \leq \delta\}) = 0.$$

Proof. Let $\mathcal{X}$ be compact and fix $\epsilon > 0$. Since (X, r) is totally bounded, there is $N_{\epsilon/2} \in \mathbb{N}$, such that (X, r) can be covered by $N_{\epsilon/2}$ balls of radius $\epsilon/2$. Then by Remark 2.31, $\xi_\epsilon \leq N_{\epsilon/2} < \infty$. It follows (b).

Now, assume (b) holds. Then, (X, r) can be covered by $\xi_{\epsilon/2} < \infty$ balls of radius $\epsilon/2$. Let $x_1, \ldots, x_{\xi_{\epsilon/2}}$ be centers of such balls and define

$$\delta := \min\{\mu(B_\epsilon(x_i)) : \mu(B_\epsilon(x_i)) > 0\}.$$

Then, $\delta > 0$. Now, take any $x \in X$ and choose $i \in \{1, \ldots, \xi_{\epsilon/2}\}$ such that $x \in B_{\epsilon/2}(x_i)$. Then we have

$$\mu(B_\epsilon(x)) \geq \mu(B_{\epsilon/2}(x_i)) \geq \delta.$$

Hence, $\hat{\mu}_{\mathcal{X}}(\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \leq \delta\}) = \mu(\{x \in X : \mu(B_\epsilon(x) \leq \delta)\}) = 0$, and (c) follows.

Finally, by (c), for $\epsilon > 0$ there is $\delta > 0$ such that $\hat{\mu}_{\mathcal{X}}(\nu([0,\epsilon)) \leq \delta) = 0$. For (a), it suffices to show that (X, r) is totally bounded, or alternatively that there is a finite 2ϵ-net in X, i.e. a finite maximal 2ϵ-separated set. For this, take a maximal 2ϵ-separated set $S \subset X$. Then, by Remark 2.31,

$$1 = \mu(X) = \mu(\bigcup_{x \in S} B_{2\epsilon}(x)) \geq \mu(\bigcup_{x \in S} B_\epsilon(x)) = \sum_{x \in S} \mu(B_\epsilon(x)) \geq \mid S \mid \delta,$$

since for μ-almost all $x \in X$ we have $\mu(B_\epsilon(x)) > \delta$, by assumption. Now, $\mid S \mid \leq 1/\delta < \infty$ and $\epsilon > 0$ was arbitrary. So (X, r) is totally bounded, i.e. compact. □

As an immediate consequence of the above theorem we have a characterization of $\mathbb{M}$-valued random variables to be supported by the space of compact metric measure spaces.

Corollary 2.33 (Random compact metric measure spaces)**.** *Let* $\mathcal{X}$ *be a* $\mathbb{M}$-*valued random variable. Then the following is equivalent.*

(a) The random metric measure space $\mathcal{X}$ *is almost surely compact.*

(b) For all $\epsilon > 0$, $\mathbb{P}(\xi_\epsilon(\mathcal{X}) < \infty) = 1$.

(c) For all $\epsilon > 0$ *there is a random* $\Delta > 0$ *such that*

$$\mathbb{P}(\hat{\mu}_{\mathcal{X}}(\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \leq \Delta\}) = 0) = 1.$$

Remark 2.34. If we define a modification of the modulus of mass distribution (Definition 2.16) by

$$\tilde{v}_\delta(\mathcal{X}) := \inf\{\epsilon > 0 : \ \mu(\{x \in X : \ \mu(B_\epsilon(x)) \leq \delta\}) = 0\},$$

there follows that $\tilde{v}_\delta(\mathcal{X}) \leq \epsilon$ iff $\mu(\{x : \mu(B_\epsilon(x)) \leq \delta\}) = 0$. Therefore, condition (c) in Theorem 2.32 can be replaced by the following condition.

(c') For all $\epsilon > 0$ there is $\delta > 0$ such that $\tilde{v}_\delta(\mathcal{X}) \leq \epsilon$.

Moreover, in Corollary 2.33 the condition (c) can be replaced by

(c') For all $\epsilon > 0$ there is a random $\Delta > 0$ such that $\mathbb{P}(\tilde{v}_\Delta(\mathcal{X}) \leq \epsilon) = 1$.

Next, we come to the characterization of locally compact metric measure spaces. We need the following notion.

Definition 2.35 (δ-restriction)**.** Let $R = (r_{i,j})_{1\leq i<j} \in \mathbb{R}_+^{\binom{\mathbb{N}}{2}}$. Then for $\delta > 0$ the δ-*restriction* is defined by

$$\tau_\delta(R) := (r_{\hat{\tau}_\delta(i),\hat{\tau}_\delta(j)})_{1\leq i<j},$$

where $\hat{\tau}_\delta(0) := 1$ and recursively $\hat{\tau}_\delta(i+1) := \inf\{j > \hat{\tau}_\delta(i) : r_{1,j} \leq \delta\}$.

Theorem 2.36 (Characterization of locally compact metric measure spaces)**.** *Let* $\mathcal{X} \in \mathbb{M}$. *Then* $\mathcal{X}$ *is locally compact if and only if*

$$\nu^{\mathcal{X}}\Big(\bigcap_{0<\epsilon<\delta} \Big\{R : \xi_\epsilon(\tau_\delta(R)) < \infty\Big\}\Big) \xrightarrow{\delta\to 0} 1.$$

Proof. Let $\mathcal{X} = (X, r, \mu)$. Then $\mathcal{X}$ is locally compact iff for μ-almost all $x \in X$ there is $\delta > 0$, such that for all $0 < \epsilon < \delta$ the ball $B_\delta(x)$ can be covered by a finite number of balls with radius ϵ. Or equivalently, that the maximal ϵ-separated set in $B_\delta(x)$ is finite. Hence, it is necessary and sufficient that,

$$\begin{aligned} 1 &= \lim_{\delta\to 0} \mu^{\otimes\mathbb{N}}\Big(\bigcap_{0<\epsilon<\delta} \{(x_1,x_2,\dots) : \xi_\epsilon(\tau_\delta((r(x_i,x_j))_{1\leq i<j})) < \infty\}\Big) \\ &= \lim_{\delta\to 0} \nu^{\mathcal{X}}\Big(\bigcap_{0<\epsilon<\delta} \Big\{R : \xi_\epsilon(\tau_\delta(R)) < \infty\Big\}\Big). \end{aligned}$$

□

Turning to relative compactness in the Gromov-weak topology on $\mathbb{M}_c$, note, that a criterion for a subset $\Gamma \subset \mathbb{M}_c$ to be relatively compact is given in [GPW10], as recalled in Remark 2.43. We give in Corollary 2.41 in the present work a version of this criterion and prove it with help of the following stronger proposition which characterizes relative compactness in $\mathbb{M}_c$.

Proposition 2.37 (Characterization of relative compactness in $\mathbb{M}_c$)**.** *A set $\Gamma \subset \mathbb{M}_c$ is relatively compact in the Gromov-weak topology on $\mathbb{M}_c$ if and only if*

(a) the set Γ is relatively compact in $\mathbb{M}$,

(b) for all $\epsilon > 0$ there is $\delta > 0$ such that the set

$$\{\hat{\mu}_{\mathcal{X}}(\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \leq \delta\}) : \mathcal{X} \in \Gamma\}$$

has only accumulation point 0.

Proof. The proposition is an immediate consequence of the following Lemma 2.38 and Lemma 2.39. □

Lemma 2.38. *Let $A \subset \mathbb{R}_+$. Then 0 is the only accumulation point of A iff for all sequences $(a_n)_{n\in\mathbb{N}} \subset A$ we have $\limsup_{n\to\infty} a_n = 0$.*

Proof. The proof is by contradiction. If there is a sequence $(a_n)_{n\in\mathbb{N}}$ such that $\limsup_{n\to\infty} a_n = a > 0$, then a is another accumulation point of A. On the other hand, if we have for all sequences $(a_n)_{n\in\mathbb{N}}$ that $\limsup_{n\to\infty} a_n = 0$, then zero is an accumulation point. And if $a > 0$ is another accumulation point there is a sequence $(a_n)_{n\in\mathbb{N}}$ such that $\limsup_{n\to\infty} a_n = a$. □

Lemma 2.39. *Let $\mathcal{X} \in \mathbb{M}$ and $(\mathcal{X}_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{M}_c$ such that $\mathcal{X}_n \xrightarrow{n\to\infty} \mathcal{X}$ in the Gromov-weak topology. Then $\mathcal{X}$ is compact iff for all $\epsilon > 0$ there is $\delta > 0$ such that*

$$\limsup_{n\to\infty} \hat{\mu}_{\mathcal{X}_n}(\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \leq \delta\}) = 0.$$

Proof. By Theorem 2.32, if (X, r, μ) is compact, for all $\epsilon > 0$ there is $\delta > 0$ such that $\hat{\mu}_{\mathcal{X}}(\nu([0,\epsilon)) \leq \delta) = 0$. Note that by Lemma 4.13, the set

$\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \le \delta\}$ is closed in $\mathscr{M}_1(\mathbb{R}_+)$. We conclude by the Portmanteau-Theorem, Theorem 4.12,

$$\limsup_{n\to\infty} \hat{\mu}_{\mathcal{X}_n}(\nu([0,\epsilon)) \le \delta) \le \hat{\mu}_{\mathcal{X}}(\nu([0,\epsilon)) \le \delta) = 0.$$

On the other hand for $\epsilon > 0$ let $\delta > 0$ be such that

$$\limsup_{n\to\infty} \hat{\mu}_{\mathcal{X}_n}(\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \le \delta\}) = 0.$$

As $\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,2\epsilon]) < \frac{\delta}{2}\}$ is open in $\mathscr{M}_1(\mathbb{R}_+)$, Lemma 4.13, we again conclude by the Portmanteau-Theorem,

$$\begin{aligned}
\hat{\mu}_{\mathcal{X}}(\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,4\epsilon)) \le \frac{\delta}{4}\}) \\
&\le \hat{\mu}_{\mathcal{X}}(\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,2\epsilon]) < \frac{\delta}{2}\}) \\
&\le \limsup_{n\to\infty} \hat{\mu}_{\mathcal{X}_n}(\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,2\epsilon]) < \frac{\delta}{2}\}) \\
&\le \limsup_{n\to\infty} \hat{\mu}_{\mathcal{X}_n}(\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \le \delta\}) = 0. \quad \square
\end{aligned}$$

Remark 2.40. Note that by Theorem 2.18 the condition (a) in Proposition 2.37 can be replaced by $\{w_{\mathcal{X}} : \mathcal{X} \in \Gamma\}$ is tight.

Corollary 2.41. *A set $\Gamma \subset \mathbb{M}_c$ is relatively compact in the Gromov-weak topology on $\mathbb{M}_c$ if*

(a) the set Γ is relatively compact in $\mathbb{M}$,

(b) for all $\epsilon > 0$, $\sup_{\mathcal{X}\in\Gamma} \xi_\epsilon(\mathcal{X}) < \infty$.

Proof. By Proposition 2.37, it suffices to show that for all $\epsilon > 0$ there is $\delta > 0$ such that the set

$$\{\hat{\mu}_{\mathcal{X}}(\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \le \delta\}) : \mathcal{X} \in \Gamma\}$$

has only accumulation point 0. The proof is by contradiction. Assume there is $\epsilon > 0$ such that for all $\delta > 0$ there are $\mathcal{X}_n = (X_n, r_n, \mu_n) \in \Gamma$, such that

$$\limsup_{n\to\infty} \mu_n(\{x \in X_n : \mu_n(B_\epsilon(x)) \le \delta\}) > 0. \quad (*)$$

By assumption, for all $n \in \mathbb{N}$, there is $N \in \mathbb{N}$ such that $\xi_{\epsilon/2}(\mathcal{X}_n) \le N < \infty$. Hence, there are points $x_1^n, \dots, x_N^n \in X_n$ such that

(a) $0 < \mu_n(B_{\epsilon/2}(x_1^n)) \leq \cdots \leq \mu_n(B_{\epsilon/2}(x_N^n)) \leq 1$,

(b) $supp(\mu_n) = \bigcup_{i=1}^N B_{\epsilon/2}(x_i^n)$.

By $(*)$ we find that

$$\begin{aligned} 0 &< \limsup_{n\to\infty} \mu_n(\{x \in X_n : \mu_n(B_\epsilon(x)) \leq \delta\}) \\ &\leq \limsup_{n\to\infty} \sum_{i=1}^n \mu_n(\{x \in B_{\epsilon/2}(x_i^n) : \mu_n(B_\epsilon(x)) \leq \delta\}). \end{aligned}$$

Therefore, there is $i_n \in \{1, \dots, N\}$ such that

$$\limsup_{n\to\infty} \mu_n(\{x \in B_{\epsilon/2}(x_{i_n}^n) : \mu_n(B_\epsilon(x)) \leq \delta\}) =: \eta > 0. \quad (**)$$

Furthermore, for $x \in \{x \in B_{\epsilon/2}(x_{i_n}^n) : \mu_n(B_\epsilon(x)) \leq \delta\}$ we have since $B_{\epsilon/2}(x_{i_n}^n) \subset B_\epsilon(x)$ that $0 < \mu_n(B_{\epsilon/2}(x_{i_n}^n)) \leq \mu_n(B_\epsilon(x)) \leq \delta$. With (a) we find $0 < \mu_n(B_{\epsilon/2}(x_1^n)) \leq \delta$ and finally

$$0 < \eta \leq \limsup_{n\to\infty} \mu_n(B_{\epsilon/2}(x_1^n)) \leq \delta.$$

Note that the latter is independent of δ. Hence, there follows a contradiction with $\delta < \eta$ and we are done. □

Remark 2.42. By the definition of ξ_ϵ, Definition 2.29, and Corollary 2.41, note that for $(\mathcal{X}_n)_{n\in\mathbb{N}}$ and $\mathcal{X} \in \mathbb{M}$ with $\mathcal{X}_n \xrightarrow{n\to\infty} \mathcal{X}$ in $\mathbb{M}$ we have that $\mathcal{X}_n \xrightarrow{n\to\infty} \mathcal{X}$ in $\mathbb{M}_c$ if for all $\epsilon > 0$ there is $N_\epsilon \in \mathbb{N}$ such that

$$\limsup_{n\to\infty} (\pi_{N+1})_* \nu^{\mathcal{X}_n}(B_\epsilon^{N+1}) = 0.$$

Remark 2.43 (Original version of Corollary 2.41). The original result for relative compactness in $\mathbb{M}_c$ from [GPW10, Proposition 6.1] reads as follows.

A set $\Gamma \subset \mathbb{M}_c$ is relatively compact in the Gromov-weak topology on $\mathbb{M}_c$ if the following two conditions are satisfied.

(a) The set $\{w_\mathcal{X} : \mathcal{X} \in \Gamma\}$ is tight in $\mathscr{M}_1(\mathbb{R}_+)$.

(b) For all $\epsilon > 0$ there is $N_\epsilon \in \mathbb{N}$ such that for all $\mathcal{X} \in \Gamma$ and $(X, r, \mu) \in \mathcal{X}$, $(supp(\mu), r)$ can be covered by N_ϵ open balls of radius ϵ.

Example 2.44. Consider the metric measure spaces defined by

$$\mathcal{X}_n := (\{0,1,\dots,n\}, r_{eucl}, Bin(n, \frac{1}{n^2})),$$

where r_{eucl} is the euclidean metric and Bin denotes the binomial distribution. Since $(\mathcal{X}_n)_{n\in\mathbb{N}}$ converges Gromov-weakly to $\mathcal{X} := (\mathbb{N}, r_{eucl}, \delta_0)$, as $n \to \infty$, the set $\Gamma = \{\mathcal{X}_n : n \in \mathbb{N}\}$ is relatively compact in $\mathbb{M}_c$. But (b) in Corollary 2.41 clearly does not hold.

Taking the limit $n \to \infty$, the mass concentrates in zero, therefore, for all $\epsilon > 0$ we can choose $\delta > 0$ small enough such that

$$0 \notin \{k \in \{0,\dots,n\} : Bin(n, \frac{1}{n^2})(B_\epsilon(k)) \le \delta\},$$

i.e.

$$\limsup_{n\to\infty} \mu(\{x \in X : \mu(B_\epsilon(x)) \le \delta\}) = 0.$$

Open questions on relative compactness in $\mathbb{M}_c$. It turned out to be hard to give a more handy characterization of relative compactness in $\mathbb{M}_c$. As indicated in the above example, the limit procedure was hardly transferred to a subset of $\mathbb{M}_c$.

We supposed a characterization for relative compactness of a subset $\Gamma \subset \mathbb{M}_c$ in, for all $\epsilon > 0$,

$$\sup_{\mathcal{X}\in\Gamma} \mu(\{x : \mu(B_\epsilon(x)) \le \delta\}) \xrightarrow{\delta\to 0} 0.$$

The sequence defined by $\mathcal{X}_n := (\{0,n\}, r_{eucl}, (1-\frac{1}{n})\delta_0 + \frac{1}{n}\delta_n)$ converges Gromov-weakly to $\mathcal{X} := (\{0\}, r_{eucl}, \delta_0)$, as $n \to \infty$, but does not satisfy the above condition.

In addition, we considered a modification of the modulus of mass distribution as mentioned in Remark 2.34,

$$\tilde{v}_\delta(\mathcal{X}) := \inf\{\epsilon > 0 : \ \mu(\{x \in X : \ \mu(B_\epsilon(x)) \le \delta\}) = 0\}.$$

We supposed a connection between relative compactness of $\Gamma \subset \mathbb{M}_c$ and the condition $\sup_{\mathcal{X}\in\Gamma} \tilde{v}_\delta \xrightarrow{\delta\to 0} 0$. The following sequence gives a counterexample.

$$\begin{aligned}
\mathcal{X}_1 &:= (\{1\}, r_{eucl}, \delta_1),\\
\mathcal{X}_2 &:= (\{1,2\}, r_{eucl}, \frac{1}{2}\delta_1 + \frac{1}{2}\delta_2),\\
\mathcal{X}_3 &:= (\{1,2,3\}, r_{eucl}, \frac{1}{2}\delta_1 + \frac{1}{4}\delta_2 + \frac{1}{4}\delta_3),\\
&\dots
\end{aligned}$$

that is, the »new« point $n \in X_n$ gets half of the mass of $n-1 \in X_{n-1}$. Clearly, $\mathcal{X}_n$ converges in the Gromov-weak topology to $(\mathbb{N}, r_{eucl}, \sum_{n\geq 1} \frac{1}{2^n}\delta_n)$, as $n \to \infty$, but the above condition does not hold.

2.3 Tightness in $\mathscr{M}_1(\mathbb{M})$ and $\mathscr{M}_1(\mathbb{M}_c)$

This section applies to tightness in the space of probability measures on the Borel-σ-algebra of $\mathbb{M}$ and $\mathbb{M}_c$. Greven et al. extended the characterization of relative compactness in $\mathbb{M}$ to characterize tightness in $\mathscr{M}_1(\mathbb{M})$ with respect to the Gromov-weak topology, [GPW09, Proposition 8.1]. Following the same idea, we give a criterion for tightness in the subspace $\mathscr{M}_1(\mathbb{M}_c)$.

Proposition 2.45 (Characterization of tightness in $\mathscr{M}_1(\mathbb{M})$)**.** *A family $A \subset \mathscr{M}_1(\mathbb{M})$ is tight with respect to the Gromov-weak topology on $\mathbb{M}$ if and only if for all $\epsilon > 0$ there is $\delta > 0$ and $C > 0$ such that*

$$\sup_{\mathbb{P}\in A} \mathbb{P}[v_\delta(\mathcal{X}) + w_{\mathcal{X}}([C,\infty))] < \epsilon. \tag{2.5}$$

Proof. Assume first, that $A \subset \mathscr{M}_1(\mathbb{M})$ is tight and fix $\epsilon > 0$. Then, there exists a compact subset $\Gamma = \Gamma(\frac{\epsilon}{4})$ of $\mathbb{M}$ such that $\inf_{\mathbb{P}\in A} \mathbb{P}(\Gamma) > 1 - \frac{\epsilon}{4}$. Since Γ is compact, by Theorem 2.18, there exists $C(\frac{\epsilon}{4}) > 0$ and $\delta(\frac{\epsilon}{4}) > 0$ such that

$$\begin{aligned}
&\sup_{\mathcal{X}\in\Gamma} w_{\mathcal{X}}([C,\infty)) < \frac{\epsilon}{4},\\
&\sup_{\mathcal{X}\in\Gamma} v_\delta(\mathcal{X}) < \frac{\epsilon}{4}.
\end{aligned}$$

Hence, for all $\mathbb{P} \in A$ we have

$$\begin{aligned}
&\mathbb{P}[v_\delta(\mathcal{X}) + w_{\mathcal{X}}([C,\infty))]\\
&\quad = \mathbb{P}[(v_\delta(\mathcal{X}) + w_{\mathcal{X}}([C,\infty)))\cdot \mathbb{1}_\Gamma] + \mathbb{P}[(v_\delta(\mathcal{X}) + w_{\mathcal{X}}([C,\infty)))\cdot \mathbb{1}_{\Gamma^c}]\\
&\quad \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.
\end{aligned}$$

Now assume (2.5) is true and fix $\epsilon > 0$. By assumption, for $\epsilon_n^2 := (2^{-n}\epsilon)^2$ there is $\delta_n > 0$ and $C_n > 0$ such that

$$\sup_{\mathbb{P}\in A} \mathbb{P}[v_{\delta_n}(\mathcal{X}) + w_{\mathcal{X}}([C_n, \infty))] < \epsilon_n^2.$$

By Chebyshev's inequality, Lemma 4.8, for all $n \in \mathbb{N}$, we get

$$\sup_{\mathbb{P}\in A} \mathbb{P}\{\mathcal{X} \in \mathbb{M} : v_{\delta_n}(\mathcal{X}) + w_{\mathcal{X}}([C_n, \infty)) \geq \epsilon_n\} < \epsilon_n.$$

Now set

$$\Gamma_\epsilon := \bigcap_{n=1}^{\infty} \{\mathcal{X} \in \mathbb{M} : v_{\delta_n}(\mathcal{X}) + w_{\mathcal{X}}([C_n, \infty)) < \epsilon_n\}.$$

By Theorem 2.18, Γ_ϵ is compact with respect to the Gromov-weak topology. We conclude for all $\mathbb{P} \in A$ by

$$\begin{aligned}\mathbb{P}(\overline{\Gamma_\epsilon}) &\geq \mathbb{P}(\Gamma_\epsilon) = 1 - \mathbb{P}(\Gamma_\epsilon^c) \\ &\geq 1 - \sum_{n\in\mathbb{N}} \mathbb{P}\{\mathcal{X} \in \mathbb{M} : v_{\delta_n}(\mathcal{X}) + w_{\mathcal{X}}([C_n, \infty)) \geq 2^{-n}\epsilon\} \\ &\geq 1 - \sum_{n\in\mathbb{N}} 2^{-n}\epsilon \\ &\geq 1 - \epsilon.\end{aligned}$$

□

Lemma 2.46. *Condition* (2.5) *in Proposition 2.45 is equivalent to the following two conditions.*

(a) The family $\{\mathbb{P}[w_{\mathcal{X}}] : \mathbb{P} \in A\}$ *is tight in* $\mathcal{M}_1(\mathbb{R}_+)$.

(b) For all $\epsilon > 0$ *there is* $\delta > 0$ *such that* $\sup_{\mathbb{P}\in A} \mathbb{P}[v_\delta(\mathcal{X})] < \epsilon$.

The latter is also equivalent to

(b') For all $\epsilon > 0$ *there is* $\delta > 0$ *such that*

$$\sup_{\mathbb{P}\in A} \mathbb{P}[\mu(\{x : \mu(B_\epsilon(x)) \leq \delta\})] < \epsilon.$$

Proof. (2.5) $\Rightarrow (a), (b)$. Since $v_\delta \geq 0$ and $w_{\mathcal{X}} \geq 0$, it follows (a) with

$$\mathbb{P}[w_{\mathcal{X}}([C, \infty))] \leq \mathbb{P}[v_\delta(\mathcal{X})] + \mathbb{P}[w_{\mathcal{X}}([C, \infty))] < \epsilon$$

and analogous it follows (b).

$(a),(b) \Rightarrow$ (2.5). Let $\epsilon > 0$. Then we have $\delta > 0$ and $C > 0$ such that $\sup_{\mathbb{P}\in A} \mathbb{P}[w_{\mathcal{X}}([C,\infty))] < \frac{\epsilon}{2}$ and $\sup_{\mathbb{P}\in A} \mathbb{P}[v_\delta(\mathcal{X})] < \frac{\epsilon}{2}$.

For the latter equivalence we use Chebyshev's inequality, Lemma 4.8. Let $\mathbb{P}[v_\delta(\mathcal{X})] < \epsilon^2$. Then, it follows that $\mathbb{P}(v_\delta(\mathcal{X}) < \epsilon) \geq 1-\epsilon$. Since for $\mathcal{X} = (X, r, \mu)$, $v_\delta(\mathcal{X}) < \epsilon$ implies that $\mu(\{x \in X : \mu(B_\epsilon(x)) \leq \delta\}) < \epsilon$, it follows $\mathbb{P}(\mu(\{x \in X : \mu(B_\epsilon(x)) \leq \delta\}) < \epsilon) \geq 1-\epsilon$. Hence, again by Chebyshev's inequality, $\mathbb{P}[\mu(\{x : \mu(B_\epsilon(x)) \leq \delta\})] < \epsilon^2$. □

Recall the map ξ_ϵ, $\epsilon > 0$, from Lemma 2.30. The following condition is sufficient for a subset of $\mathscr{M}_1(\mathbb{M}_c)$ to be tight.

Proposition 2.47 (Criterion for tightness in $\mathscr{M}_1(\mathbb{M}_c)$)**.** *A set $A \subset \mathscr{M}_1(\mathbb{M}_c)$ is tight if it is tight in $\mathscr{M}_1(\mathbb{M})$ and for all $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that*

$$\sup_{\mathbb{P}\in A} \mathbb{P}\{\mathcal{X} \in \mathbb{M}_c : \xi_\epsilon(\mathcal{X}) > N_\epsilon\} < \epsilon. \tag{2.6}$$

Proof. The proof is easy if we follow the idea given in the proof of the characterization of tightness in $\mathcal{M}_1(\mathbb{M})$, Proposition 2.45. For fixed $\epsilon > 0$ we define $\epsilon_n := 2^{-n}(\epsilon/2)$ and by using the assumption a set

$$\Gamma_\epsilon^{\mathbb{M}_c} := \bigcap_{n=1}^{\infty} \{\mathcal{X} \in \mathbb{M}_c : \xi_{\epsilon_n}(\mathcal{X}) \leq N_{\epsilon_n}\}.$$

We are done by taking intersection with the set given in the proof of Proposition 2.45. □

Remark 2.48. Note that condition (2.6) in Proposition 2.47 is equivalent to the condition that for all $\epsilon > 0$, the set of distributions of $\xi_\epsilon(\mathcal{X})$, such that there is $\nu \in A$ with $\mathcal{X}$ is distributed by ν, is tight.

Remark 2.49 (Open question). It turned out to be hard to characterize the compact sets in $\mathbb{M}_c$, see the last paragraph in Section 2.2. Hence, the problem extends to find a characterization of tightness in $\mathscr{M}_1(\mathbb{M}_c)$.

CHAPTER 3

Λ-coalescents and Λ-coalescent measure trees

This chapter deals with the notion of Λ-coalescent. The first three sections give a review over the theory and prepare Section 3.4, where we give a characterization of a Λ-coalescent to come down from infinity by using an assigned metric measure space, the Λ-coalescent measure tree.

By definition, the Λ-coalescent comes down from infinity if it consists only of finitely many lines at any positive time. Indeed, we show in Theorem 3.36 that the coalescent comes down from infinity if and only if the corresponding metric measure space, if it exists, is compact. Moreover, if the coalescent stays infinite, then the corresponding metric measure space is not even locally compact.

We start by setting the basic fundament considering exchangeable random partitions of the natural numbers in Section 3.1. Theorem 3.5 is Kingman's representation which characterizes random partitions. In Section 3.2, we define the Λ-coalescent as a partition-valued stochastic process and show existence and uniqueness in Theorem 3.10. Some important properties about blocks consisting only of one element are collected in Corollary 3.13. Finally, Section 3.3 gives a review over some criteria of a Λ-coalescent to come down from infinity. Mainly, we show in Theorem 3.22 a characterization due to Schweinsberg.

3.1 Exchangeable random partitions

In this section we give a brief summary of the theory of exchangeable random partitions, which is elementary for coalescent theory. The theory is mainly due to Kingman. We show Kingman's representation, Theorem 3.5, an analogue of de Finetti's Theorem for exchangeable random partitions. In addition, we have as a corollary Kingman's correspondence, Corollary 3.8, which permits to think of a random partition as a discrete object, taking values in the set $\mathbb{S}$ of partitions of $\mathbb{N}$, or alternatively as a continuous object, taking values in the set defined by

$$\mathcal{S} := \{f = (f_0, f_1, f_2, \dots) \in (0,1)^{\mathbb{N}} : f_1 \geq f_2 \geq \dots, \sum_{i \geq 0} f_i = 1\}.$$

We follow primarily [Pit05], [Pit99] and [Ber09].

We fix some vocabulary and notations. A *partition* of a set A is a collection $\mathcal{P} = \{\pi_i, i \in I\}$ of pairwise disjoint subsets of A, also called *blocks*, with $A = \bigcup_{i \in I} \pi_i$. We call a block consisting only of one element a *singleton block*. Let $\mathbb{S}$ be the set of all partitions of $\mathbb{N}$, and for all $n \in \mathbb{N}$, let $\mathbb{S}_n$ be the finite set of all partitions of $[n] := \{1, 2, 3, \dots, n\}$. In addition, let $\rho_n : \mathbb{S} \to \mathbb{S}_n$ and for $m \geq n$, $\rho_n^m : \mathbb{S}_m \to \mathbb{S}_n$ be the restriction operators.

We equip $\mathbb{S}_n$ with the discrete topology. Each $\mathcal{P} \in \mathbb{S}$ can be identified with the sequence $(\rho_1\mathcal{P}, \rho_2\mathcal{P}, \dots)$ in $\mathbb{S}_1 \times \mathbb{S}_2 \times \cdots$. Now, we give $\mathbb{S}$ the topology it inherits as a subset of $\mathbb{S}_1 \times \mathbb{S}_2 \times \cdots$. Therefore, $\mathbb{S}$ is compact as a product of compact sets. Furthermore, $\mathbb{S}$ is metrizable by the complete metric $\mathrm{d}_{\mathbb{S}}$ given by

$$\mathrm{d}_{\mathbb{S}}(\mathcal{P}_1, \mathcal{P}_2) := (\max\{n : \rho_n\mathcal{P}_1 = \rho_n\mathcal{P}_2\})^{-1}.$$

Equipped with this metric $\mathbb{S}$ is a Polish space.

Each partition $\mathcal{P} \in \mathbb{S}$ defines an equivalence relation $\sim_{\mathcal{P}}$ on $\mathbb{N}$ by $i \sim_{\mathcal{P}} j$ if and only if there exists a partition element $\pi \in \mathcal{P}$ with $i, j \in \pi$. Given a block $\pi \subset \mathbb{N}$, we define, if it exists, the *asymptotic frequency* of π by

$$f(\pi) := \lim_{n \to \infty} \frac{1}{n} \#\{k \in \{1, \dots, n\} : k \in \pi\}.$$

For instance, consider the blocks of $\mathbb{N}$ consisting either of odd or even numbers. Then the frequencies are both equal to $1/2$.

We turn toward the definition of exchangeable random partitions and remark that each permutation σ of $[n]$ operates on a subset $\pi \subset [n]$ by

$$\sigma\pi := \{\sigma(i) : i \in \pi\}.$$

Hence, we have an operation on partitions $\mathcal{P}^{(n)} = (\pi_1, \pi_2, \dots)$ of $[n]$ if we define

$$\sigma\mathcal{P}^{(n)} := (\sigma\pi_1, \sigma\pi_2, \dots).$$

Definition 3.1 (Exchangeable random partition)**.** A *random partition* $\Pi^{(n)}$ of $[n] := \{1, \dots, n\}$ is a random element of $\mathbb{S}_n$. It is called *exchangeable* if for all permutations σ of $[n]$, $\sigma\Pi^{(n)}$ has the same distribution as $\Pi^{(n)}$. A random partition Π of $\mathbb{N}$ is called *exchangeable* if $\Pi^{(n)}$, the restriction of Π to $[n]$, is exchangeable for all $n \in \mathbb{N}$.

Remark 3.2 (Equivalent definition). The following definition of exchangeable random partitions is equivalent to Definition 3.1. A random partition $\Pi^{(n)}$ of $[n] := \{1, \dots, n\}$ is called exchangeable if for each partition $\{\pi_1, \dots, \pi_k\}$ of $[n]$,

$$\mathbb{P}(\Pi^{(n)} = \{\pi_1, \dots, \pi_k\}) = p(\#\pi_1, \dots, \#\pi_k)$$

for some symmetric function $p : \mathcal{C}_n \to [0, 1]$, where $\mathcal{C}_n$ denotes the set of all compositions of n. This function is called the *exchangeable partition probability function*, [Pit99].

Example 3.3. We consider a random bond percolation (clusters in a random graph) on $\mathbb{Z}^d$. Firstly, we arbitrarily enumerate all vertices of $\mathbb{Z}^d$ by $(v_1, v_2, \dots)$. Then let $i, j \in \mathbb{N}$ be in the same block of $\Pi(\omega)$ iff v_i and v_j are in the same connected component in a realization ω of the bond percolation. The resulting random partition is not exchangeable. On the other hand, take some i.i.d. random vertices $(V_1, V_2, \dots)$ and let $i, j \in \mathbb{N}$ be in the same block iff V_i and V_j are in the same connected component. The second random partition is exchangeable, [Ber09].

The paintbox construction. The following construction of an exchangeable random partition is due to Kingman. According to Theorem 3.5, every exchangeable random partition has the same distribution as one generated this way.

We follow [Ber09] and consider the set defined by

$$\mathcal{S} := \{f = (f_0, f_1, f_2, \dots) \in (0,1)^{\mathbb{N}} : f_1 \geq f_2 \geq \dots, \sum_{i\geq 0} f_i = 1\}.$$

The coordinate f_0 plays a special role, since we do not require $f_0 \geq f_1$.

Take $f \in \mathcal{S}$ and let $U_1, U_2, \dots$ be independent uniformly distributed random variables. We think of a tiling of $(0,1)$ given by

$$\begin{aligned} J_0 &:= (0, f_0), \\ J_1 &:= (f_0, f_0 + f_1), \\ J_2 &:= (f_0 + f_1, f_0 + f_1 + f_2) \\ &\dots \end{aligned}$$

That is, each tile J_i has size f_i. Clearly we can identify the n^{th} tile with the n^{th} coordinate of f. Let, for $0 < u < 1$,

$$I(u) := \inf\{n : \sum_{i=0}^{n} f_i > u\}$$

be the index of the tile which contains u or rather the index of the corresponding coordinate of f.

We define a random partition $\Pi_f = \Pi$ by letting $i, j \in \mathbb{N}$ be in the same block iff $I(U_i) = I(U_j) > 0$ or $i = j$. It is important to note, that if U_i falls into J_0, i.e. in the 0^{th} coordinate of f, then i is guaranteed to form a singleton block in the partition Π. For this reason, f_0 is referred to as the *dust* of f. We say, that the partition Π has *no dust* if $s_0 = 0$, i.e. if Π has no singletons.

On the other hand, if $I(U_i) \geq 1$, then the block containing i is infinitely large. Moreover, the asymptotic frequencies of this blocks exist and are strictly positive, see Corollary 3.6. The partition Π is exchangeable, since $(U_1, \dots, U_n)$ is for all $n \geq 1$ a sequence of exchangeable random variables, Definition 4.24.

Remark 3.4 (Paintbox principle). Kingman suggests the following mental picture: Think of real numbers $0 < u < 1$ as labeling the colors of the spectrum. Imagine coloring objects $1, 2, 3, \dots$ at random by painting object i with color $I(U_i)$. If U_i falls into J_0, then paint i with a unique new color. Hence, we obtain a partition into sets of »identically-colored« objects, [Ald85].

For $f \in \mathcal{S}$, let ρ_f be the law of a random partition Π_f constructed as above.

Theorem 3.5 (Kingman's representation)**.** *Let* Π *be an exchangeable random partition of* $\mathbb{N}$*, then there exists a probability distribution* μ *on* $\mathscr{B}(\mathcal{S})$ *such that*

$$\mathbb{P}(\Pi \in \cdot) = \int_{\mathcal{S}} \rho_f(\cdot)\mu(df).$$

Proof. Let Π be an exchangeable random partition. We sketch the proof given in [Ald85] following [Ber09, Theorem 1.1] and define a random map

$$\begin{aligned} \varphi : \mathbb{N} &\to \mathbb{N} \\ i &\mapsto \min\{n : i \sim_\Pi n\}, \end{aligned}$$

i.e. the minimal integer lying in the same block as i.

Take a sequence $(U_i)_{i\in\mathbb{N}}$ of independent uniformly distributed random variables, independent from φ, and let $X_i := U_{\varphi(i)}$. The sequence $(X_i)_{i\in\mathbb{N}}$ is exchangeable, Definition 4.24. Hence, by de Finetti's Theorem, Theorem 4.26, there exists a measure μ, such that, conditionally the exchangeable σ-algebra, the sequence is i.i.d. with distribution μ. Moreover, we find, that $i, j \in \mathbb{N}$ are in the same block iff $X_i = X_j$.

We denote by F the distribution function of μ. In addition, let

$$q(x) := \inf\{y \in \mathbb{R} : F(y) > x\},$$

then we find that $(X_i)_{i\in\mathbb{N}}$ has the same distribution as $(q(V_i))_{i\in\mathbb{N}}$, where $(V_i)_{i\in\mathbb{N}}$ is a sequence of independent uniformly on $[0, 1]$ distributed random variables.

We find that Π has distribution ρ_f, where $f = (f_0, f_1, f_2, \dots) \in \mathcal{S}$ and $(f_1, f_2, \dots)$ gives an ordered listing of the atoms of μ and $f_0 := 1 - \sum_{i\geq 1} f_i$. □

Corollary 3.6 (Asymptotic frequencies)**.** *Let* Π *be an exchangeable random partition of* $\mathbb{N}$ *and let* $(K_{n,i}, i \geq 1)$ *be the decreasing rearrangement of block sizes of* $\Pi^{(n)}$*, with* $K_{n,i} = 0$ *if* $\Pi^{(n)}$ *has fewer than* i *blocks, where* $\Pi^{(n)}$ *denotes the restriction of* Π *to* $[n]$*. Then for each* $i \geq 1$*, the frequency* $f_i := \lim_{n\to\infty} \frac{1}{n} K_{n,i}$ *exists almost surely.*

Proof. We follow the notation as given in the proof of Theorem 3.5. Since by the Glivenko-Cantelli Theorem, Theorem 4.28, the empirical distributions F_n of the first n values of the sequence given by

$$F_n(x) := \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{X_i \leq x}$$

are converging uniformly almost surely to F, we find that $\frac{1}{n} K_{n,i}$, the i^{th} largest atom of F_n, has almost sure limit f_i, the i^{th} largest atom of F, [Pit05, Theorem 2.2] and [Ald85]. □

Note that by Theorem 3.5, for any exchangeable random partition, the only finite blocks are the singletons, identified with f_0. Overwise, the block is infinite and has by Corollary 3.6 well-defined, strictly positive frequency f_i, $i \geq 1$.

Definition 3.7 (Proper frequencies of a partition)**.** A random partition Π corresponding to $f = (f_0, f_1, f_2, \dots) \in \mathcal{S}$ has *proper frequencies* if

$$\sum_{i \geq 1} f_i = 1, \text{ i.e. } f_0 = 0.$$

To summarize, there is a one to one correspondence between exchangeable random partitions and the set $\mathcal{S}$.

Corollary 3.8 (Kingman's correspondence)**.** *There is a one to one correspondence between the law of exchangeable random partitions* Π *and the distributions* μ *on* $\mathscr{B}(\mathcal{S})$.

3.2 The Λ-coalescent

In this section we introduce the coalescent process with multiple but no simultaneous collisions. The main part here is Theorem 3.10, which gives a characterization of this process by a finite measure Λ on $\mathscr{B}([0,1])$. It is due to Pitman and principally motivated the name Λ-coalescent for this kind of processes. Moreover, we prove some properties of this process in Corollary 3.11 and Corollary 3.13. Finally, we give a quick view over a possible construction of a Λ-coalescent by a Poisson point process as given in [Pit99] and consider a Lévy process interpretation as motivated from Berestycki in [Ber09].

Recall from Section 3.1 the Polish space $\mathbb{S}$ of partitions of $\mathbb{N}$. In addition, let $\rho_n : \mathbb{S} \to \mathbb{S}_n$ and $\rho_n^m : \mathbb{S}_m \to \mathbb{S}_n$ be the restriction operators. Each partition $\mathcal{P} \in \mathbb{S}$ defines an equivalence relation $\sim_{\mathcal{P}}$ on $\mathbb{N}$ by $i \sim_{\mathcal{P}} j$ if and only if there exists a block $\pi \in \mathcal{P}$ with $i, j \in \pi$.

We are looking for an $\mathbb{S}$-valued Markov process $\Pi = (\Pi_t, t \geq 0)$ such that for all $n \in \mathbb{N}$, the restriction $\Pi^{(n)} = \rho_n \Pi$ is an $\mathbb{S}_n$-valued Markov process, exchangeable for any $t \geq 0$ and *consistent* in the sense that for all $1 \leq m \leq n$, the law of $\Pi^{(n)}$ restricted to $[m]$ is that of $\Pi^{(m)}$. The dynamic of the process is as follows. When $\Pi_t^{(n)}$ has b blocks, any fixed k-tuple of blocks of $\Pi_t^{(n)}$ merges to form a single block at rate $\lambda_{b,k}$, for some array of nonnegative real numbers $(\lambda_{b,k} : 2 \leq k \leq b < \infty)$. We call such an array $(\lambda_{b,k})$ *consistent* if the corresponding process is consistent.

Lemma 3.9 (Characterization of consistent arrays). *An array of rates $(\lambda_{b,k})$ is consistent iff for all $2 \leq k \leq b < \infty$,*

$$\lambda_{b,k} = \lambda_{b+1,k} + \lambda_{b+1,k+1}.$$

Proof. It suffices to consider the case $m = n - 1$, i.e. we consider $\Pi^{(n)}$ and $\Pi^{(n-1)}$. Given k blocks among b there are two ways for a Markovian process to coalesce when revealed an extra block $b+1$. Either these k blocks coalesce among themselves without the extra block, or they coalesce with it. □

The following theorem from [Pit99, Theorem 1] ensures that such a process Π as desired exists.

Theorem 3.10 (Λ-coalescent). *Let $(\lambda_{b,k} : 2 \leq k \leq b < \infty)$ be an array of nonnegative real numbers. Then there exists a $\mathbb{S}$-valued process Π as described above if and only if there is a nonnegative finite measure Λ on $\mathscr{B}([0,1])$ such that*

$$\lambda_{b,k} = \int_0^1 x^{k-2}(1-x)^{b-k}\Lambda(dx). \tag{3.1}$$

Since the measure Λ uniquely determines the law of the process Π, we call $\Pi = (\Pi_t, t \geq 0)$ a Λ-*coalescent*. In addition, let $\mathbb{P}^{\Lambda,\mathcal{P}_0}$ be the law of Π with $\Pi_0 = \mathcal{P}_0$ on the space of càdlàg paths with the Skorohod topology. Normally, we start in the trivial partition $\mathcal{P}_0 = \{\{1\}, \{2\}, \{3\}, \dots\}$ of singletons.

Proof of Theorem 3.10. The necessity follows since equation (3.1) gives a bijection between consistent arrays and nonnegative finite measures on $\mathscr{B}([0,1])$: If we define for $i,j = 0,1,2,\dots$,

$$\mu_{i,j} := \lambda_{i+j+2,i+2},$$

then by Lemma 3.9, we have consistency iff for $i,j = 0,1,2,\dots$,

$$\mu_{i,j} = \mu_{i+1,j} + \mu_{i,j+1}.$$

By Theorem 4.27, a version of de Finetti's Theorem for an infinite sequence of exchangeable random variables taking only values 0 and 1, this yields to

$$\begin{aligned}\mu_{0,0} &= 1,\\ \mu_{i,j} &= \int_0^1 x^i(1-x)^j \Lambda(dx)\end{aligned}$$

for some nonnegative finite measure Λ on $\mathscr{B}([0,1])$.

For sufficiency, assume (3.1) holds. Then by Lemma 3.9, it follows easily by linearity of the integral, that the array of rates is consistent. Then, the process $\Pi^{(n)}$ can be constructed following [Kin82b] by an application of the Kolmogorov consistency Theorem, Theorem 4.31. Finally, the desired process Π is obtained by letting Π_t be the unique partition whose restriction to $[n]$ is $\Pi_t^{(n)}$ for every $n \in \mathbb{N}$. □

Corollary 3.11 (Exchangeability of the Λ-coalescent)**.** *For a nonnegative finite measure Λ on $\mathscr{B}([0,1])$ in a Λ-coalescent the partition Π_t is for each $t \geq 0$ an exchangeable random partition of $\mathbb{N}$.*

Proof. This is an immediate consequence of the form of the rates of $\Pi^{(n)}$. If σ is a permutation of $[n]$, the process $\sigma\Pi^{(n)}$ whose blocks are the σ-image of the blocks of $\Pi^{(n)}$ is a copy of $\Pi^{(n)}$, [Pit99, Section 3.2]. □

The next Corollary 3.13 will be frequently used in Section 3.4. We consider for $\pi \subset \mathbb{N}$, the block size $K_{n,\pi} := \#\{k \in \{1,\dots,n\} : k \in \pi\}$. As above in Section 3.1 the family $\{\tilde{f}(\pi) : \pi \in \Pi_t\}$ of frequencies

$$\tilde{f}(\pi) := \lim_{n\to\infty} \frac{1}{n} K_{n,\pi}$$

exists $\mathbb{P}^{\Lambda,\mathcal{P}_0}$-almost surely. We define $f := (f_1, f_2, ...)$ to be the ranked rearrangement of the family $\{\tilde{f}(\pi) : \pi \in \Pi_t\}$. In addition, let

$$(\Pi_t)^j := \{k \in \mathbb{N} : k \sim_{\Pi_t} j\}$$

denote the partition element of Π_t containing j. Recall the notion of dust and of proper frequencies from Section 3.1 and Definition 3.7.

Definition 3.12 (Properties of Λ-coalescent)**.** Let $\Pi = (\Pi_t)_{t\geq 0}$ be a Λ-coalescent corresponding to $f = (f_0, f_1, f_2, \dots) \in \mathcal{S}$.

(a) The Λ-coalescent Π has the *dust-free property* if for all $t > 0$,

$$\mathbb{P}^{\Lambda,\mathcal{P}_0}\{\tilde{f}((\Pi_t)^1) = 0\} = 0.$$

(b) The Λ-coalescent has *proper frequencies* if Π_t has proper frequencies for all $t > 0$, almost surely.

(c) The Λ-coalescent Π *comes down form infinity* if $\mathbb{P}(N(t) < \infty) = 1$ for all $t > 0$. It *stays infinite* if $\mathbb{P}(N(t) = \infty) = 1$ for all $t > 0$.

Note, that the dust-free property means, that, by exchangeability, the Λ-coalescent has for all times $t > 0$ no singleton blocks, $\mathbb{P}^{\Lambda,\mathcal{P}_0}$-almost surely.

Corollary 3.13 (Dust-free)**.** *The total collision rate $\lambda_b^{(i)}$ of the block containing $i \in \mathbb{N}$ with any other blocks among b is bounded above by $\int_0^1 x^{-1}\Lambda(dx)$. Moreover, if the Λ-coalescent has infinitely many blocks, then*

$$\lim_{b\to\infty} \lambda_b^{(i)} = \int_0^1 x^{-1}\Lambda(dx).$$

Furthermore, the following conditions are equivalent.

(a) The dust-free property holds.

(b) The Λ-coalescent has proper frequencies.

(c) $\int_0^1 x^{-1}\Lambda(dx) = \infty$.

Proof. The first assertion is easily proven by the following calculation.

$$\begin{aligned}
\lambda_b^{(i)} &= \sum_{k=2}^{b} \binom{b-1}{k-1} \lambda_{b,k} = \sum_{k=2}^{b} \int_0^1 \binom{b-1}{k-1} x^{k-2}(1-x)^{b-k} \Lambda(dx) \\
&= \int_0^1 x^{-1} \sum_{k=1}^{b} \binom{b-1}{k-1} x^{k-1}(1-x)^{(b-1)-(k-1)} - (1-x)^{b-1} \Lambda(dx) \\
&= \int_0^1 x^{-1} \Big((x + (1-x))^{b-1} - (1-x)^{b-1} \Big) \Lambda(dx) \\
&= \int_0^1 x^{-1} \Big(1 - (1-x)^{b-1} \Big) \Lambda(dx).
\end{aligned}$$

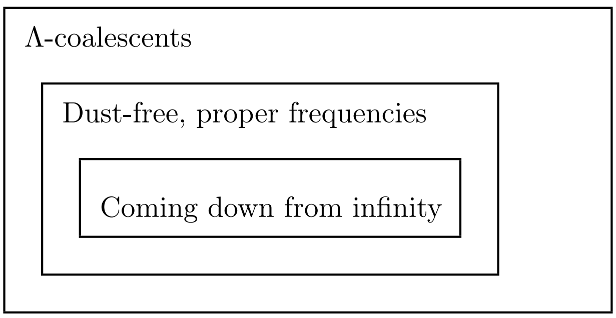
If there are infinitely many blocks, we find by monotone convergence,

$$\lim_{b \to \infty} \lambda_b^{(i)} = \int_0^1 x^{-1} \Lambda(dx).$$

Recall from the definition of proper frequencies, Definition 3.7, and from the comment above this corollary, that (a) and (b) mean the same, namely that there are no singleton blocks.

To prove the equivalence of (c), assume first that the integral is finite. Since this is an upper bound for $\lambda_b^{(i)}$, we may have only finitely many collisions and therefore a strictly positive probability for a singleton block. Conversely, assume (c) holds. If $N(t) := \#\Pi_t = \infty$ then infinitely many collisions have occurred, since the total collision rate is infinite. If $N(t) < \infty$ then by exchangeability, we have neither finite nor singleton blocks. □

Remark 3.14 (Classification of Λ-coalescents). Note, that there is a classification of the Λ-coalescents into (at least) three classes. The class of Λ-coalescents coming down from infinity and the larger class of processes having the dust-free property.



Example 3.15.

(a) For $\Lambda = \delta_0$ the rates are given by $\lambda_{b,k} = 0$ except for $k = 2$, in which case $\lambda_{b,k} = 1$. Thus, the corresponding Λ-coalescent is the *Kingman's coalescent*, in which each pair of blocks merges at rate 1 but no multiple collisions are allowed.

(b) For $\Lambda = \delta_1$ the rates are $\lambda_{b,k} = 0$ expect of $k = b$, in which case $\lambda_{b,k} = 1$. Thus, nothing happens for an exponential distributed time with parameter 1, then all blocks coalesce.

(c) For $\Lambda = U([0,1])$, the uniform distribution, we find that

$$\lambda_{b,k} = \frac{(k-2)!(b-k)!}{(b-1)!}$$

and hence the Λ-coalescent is the *Bolthausen-Sznitman coalescent*. The Bolthausen-Sznitman coalescent firstly appeared in connection to the physics of spin glasses. A *spin glass* is a magnet with a special geometrical property of the crystal structure (frustration), where usually ferromagnetic and antiferromagnetic bonds are randomly distributed. Moreover, the Bolthausen-Sznitman coalescent has applications in certain combinatorial models of branching Brownian motion, [BBS10], random trees and random traveling waves, [Ber09].

(d) Let $0 < \alpha < 2$. For $\Lambda = Beta(2-\alpha, \alpha)$, i.e.

$$\Lambda(dx) := \frac{1}{\Gamma(2-\alpha)\Gamma(\alpha)} x^{1-\alpha}(1-x)^{\alpha-1}dx,$$

the Λ-coalescent is called the *Beta-coalescent* with parameter α. The Beta-coalescent has applications in the genealogy of populations mainly with large variation in the offspring distribution. Moreover, the Beta-coalescent is some kind of interpolation between the Kingman's coalescent and the Bolthausen-Sznitman coalescent. Indeed, if $\alpha = 1$, the Beta-coalescent is just the Bolthausen-Sznitman coalescent. If $\alpha \to 2$, then the Beta-coalescent converges in distribution to the Kingman's coalescent with respect to the Skorohod topology, [Ber09].

Poissonian construction. In the following we describe a construction of the Λ-coalescent as given by Pitman in [Pit99, above Corollary 3]. We have to assume, that Λ has no atom at 0, otherwise see Example 3.16.

Take a sequence $\xi = (\xi_i)_{i\in\mathbb{N}}$ of independent Bernoulli distributed random variables governed by $\mathbb{P}_x$ with $\mathbb{P}_x(\xi_i = 1) = x$. In addition, let N be a Poisson point process on $\mathbb{R}_+ \times \{0,1\}^{\mathbb{N}}$ with intensity $dt \otimes \nu(d\xi)$, where

$$\nu(A) := \int_0^1 x^{-2}\mathbb{P}_x(A)\Lambda(dx),$$

for $A \subset \{0,1\}^{\mathbb{N}}$ product measurable. Recall that a *Poisson point process* N is a σ-finite kernel N with independent increments, i.e. for all $B_1,\dots,B_n \subset \mathbb{R}_+ \times \{0,1\}^{\mathbb{N}}$ disjoint and measurable, $N(B_1),\dots,N(B_n)$ are independent and for $B \subset \mathbb{R}_+ \times \{0,1\}^{\mathbb{N}}$, $N(B)$ is Poisson distributed.

For all $n \in \mathbb{N}$, it is now possible to construct a Markov process $\Pi^{(n)}$. For this, take any partition $\mathcal{P} \in \mathbb{S}$ and set $\Pi_0^{(n)} := \rho_n\mathcal{P}$.

We get the desired dynamic of $\Pi^{(n)}$ as follows. The process jumps at times t of points (t,ξ) of N such that $\sum_{i=1}^n \xi_i \geq 2$. By the construction of the intensity measure, this set is discrete. For such t, if $\Pi_{t-}^{(n)} = \{\pi_1,\dots,\pi_n\}$, then $\Pi_t^{(n)}$ is defined by merging all blocks π_i such that $\xi_i = 1$, i.e. $\mathbb{P}_x(\xi_i = 1) = x$ is the probability of one block to take part in the merging.

Interpreting this way, the intensity measure ν leads directly to the rates $(\lambda_{b,k})$. By a fix success probability x, the probability that k blocks among b merge and $b-k$ blocks stay untouched is given by

$$\mathbb{P}_x(\xi_1 = 1,\dots,\xi_k = 1, \xi_{k+1} = 0,\dots,\xi_b = 0) = x^k(1-x)^{b-k}.$$

Note, that the consistency of the joint distributions is guaranteed by the independence of (ξ_i). Moreover, by the construction of the intensity, $\Pi^{(n)}$ is Markovian. Then, define Π by letting $\Pi^{(n)}$ be the restriction to $[n]$.

Example 3.16.

(a) Let Λ have an atom at 0 with $\Lambda(\{0\}) := \theta$. Then, decomposing Λ into

$$\Lambda = \theta\delta_0 + \tilde{\Lambda},$$

where $\tilde{\Lambda}$ has no atom at 0, the dynamic of the Λ-coalescent is described by a Poisson point process with intensity $dt \otimes \tilde{\nu}(d\xi)$ given by

$$\tilde{\nu} = \int_0^1 x^{-2}\mathbb{P}_x\tilde{\Lambda}(dx)$$

added the dynamic of a Kingman's coalescent, i.e. every pair of blocks merges at rate θ.

(b) Let Λ have an atom at 1 with $\Lambda(\{1\}) := \theta$. Then, we define

$$\tilde{\Lambda} := \Lambda - \theta\delta_1.$$

For the $\tilde{\Lambda}$-coalescent Π and an independent with parameter θ exponential distributed time T, we consider the coalescent process defined by

$$\Pi'_t := \begin{cases} \Pi_t, & \text{if } t < T \\ \{\mathbb{N}\}, & \text{if } t \geq T \end{cases}.$$

Then, Π' is a Λ-coalescent, [Pit99, Example 20].

Lévy process interpretation. For completeness we propose the following interpretation of Theorem 3.10 and the above Poissonian construction as given by Berestycki in [Ber09].

Firstly, we define an operation $\circ$ on $\mathbb{S}$ which turns $(\mathbb{S}, \circ)$ into a monoid. Let $\mathcal{P} = \{\pi_1, \pi_2, \dots\}$ and $\mathcal{P}'$ be two partitions of $\mathbb{N}$. Then, if $i, j \in \mathbb{N}$ are in the same block of $\mathcal{P}'$, π_i and π_j are subsets of a single block of $\mathcal{P} \circ \mathcal{P}'$, i.e. we merge (»coagulate«) all blocks of $\mathcal{P}$ whose labels are in the same block of $\mathcal{P}'$.

We regard a Λ-coalescent Π as a Lévy-process in the monoid $(\mathbb{S}, \circ)$ in the sense that for every $t, s \geq 0$,

$$\Pi_{t+s} = \Pi_t \circ \Pi_s.$$

Recall that a *Lévy-process* is a (normally real-valued) process X with independent and stationary increments. The most common examples are the Brownian motion and the Poisson process. A fundamental result about Lévy-processes is the *Lévy-Itô decomposition*, i.e. any Lévy-process is the sum of a Brownian motion, a deterministic drift and compensated Poisson jumps.

For every Lévy-process, there is a measure ν', the *Lévy measure*, such that the process makes a jump of size x at rate $\nu'(dx)$. We naively write

$$\nu'(dx) = rate : X_t \to X_t + x.$$

The state space of a Lévy-process can be extended to be a group G. The three most interesting cases are locally compact Abelian groups, Lie groups and general locally compact groups. In the first case there exists an analogue Lévy-Itô decomposition, since the Fourier analysis approach to prove the decomposition is possible to be extended, [Hey77, Section 5.6]. In the nonabelian cases, there are still some important results. We refer to [Hun56] and [App04] for more information.

The case where the state space is a monoid such as $\mathbb{S}$ does not seem to be well studied. Nevertheless, we assume that a Λ-coalescent process interpreted as a Lévy-process on the monoid $\mathbb{S}$ also can be described by a measure ν' on the space $\mathbb{S}$, such that the process makes a jump of size $\mathcal{P}$ at rate $\nu'(d\mathcal{P})$. We write as above

$$\nu'(d\mathcal{P}) = rate : \Pi_t \to \Pi_t \circ \mathcal{P}.$$

Thus, the Poissonian construction can be seen as some kind of Lévy-Itô decomposition for a Λ-coalescent.

Next, we have to consider the following random partition $\mathcal{P}_x$. Given $x \in (0,1)$, the integer $i \in \mathbb{N}$ takes part into one special block by success of a Bernoulli distribution with parameter x, else we get a singleton. We consider $\mathcal{P} \circ \mathcal{P}_x$ for any $\mathcal{P} \in \mathbb{S}$. For every block in $\mathcal{P}$ we toss a coin with success probability x, then we merge all the blocks that come success. Hence, we find for the rate $\lambda_{b,k}$ of a k-tuple to merge among b blocks

$$\lambda_{b,k} := \int_0^1 x^k (1-x)^{b-k} \nu'(dx).$$

In the end, we assume that the rate at which at least two blocks among n merge is finite. Hence, we have

$$\int_0^1 \binom{n}{2} x^2 \nu'(dx) < \infty.$$

This yields to a finite measure Λ by letting $\nu'(dx) =: x^{-2}\Lambda(dx)$. We find Theorem 3.10.

3.3 Criteria for coming down from infinity

It is natural to ask if a random partition or a Λ-coalescent has finitely or infinitely many blocks. In this section we focus on the question if a given Λ-coalescent comes down from infinity or not, Definition 3.12. The first results about this are due to Pitman and we start with a quick introduction to these. Then, the main part of this section is Theorem 3.22. It is from the PhD thesis of Schweinsberg and gives a well and more manageable characterization of Λ-coalescent coming down from infinity.

Let Λ be a finite measure on $\mathscr{B}([0,1])$ and $\Pi = (\Pi_t)_{t\geq 0}$ be a Λ-coalescent as defined in Section 3.2. In addition, we set $N(t) := \#\Pi_t$, the number of blocks in the partition Π_t.

We recall the definition of a Λ-coalescent coming down from infinity from Definition 3.12.

Definition 3.17. Let Π be a Λ-coalescent and $N(t) := \#\Pi_t$.

(a) The Λ-coalescent *comes down form infinity* if $\mathbb{P}(N(t) < \infty) = 1$ for all $t > 0$.

(b) The Λ-coalescent *stays infinite* if $\mathbb{P}(N(t) = \infty) = 1$ for all $t > 0$.

Example 3.18.

(a) As we will see later in Example 3.24, for $\Lambda = \delta_0$ or $\Lambda = U([0,1])$, respectively, the Kingman's coalescent comes down form infinity and the Bolthausen-Sznitman coalescent stays infinite.

(b) Assume Λ has an atom at 0, i.e. $\Lambda(\{0\}) := \theta$. Moreover, let $N_\Lambda(t)$ denote the number of blocks in the Λ-coalescent and $N_{\theta\delta_0}(t)$ the number of blocks in the $\theta\delta_0$-coalescent, i.e. a Kingman's coalescent. Then, we find for all $t \geq 0$ that $N_\Lambda(t) \leq N_{\theta\delta_0}(t)$. Hence, the Λ-coalescent comes down from infinity. See also Example 3.16 or the later Corollary 3.25.

(c) Assume Λ has an atom at 1. Then, corresponding to Example 3.15 (b), all blocks coalesce in finite time, i.e. the Λ-coalescent comes down from infinity. Note that [Pit99, Example 20] gives a description of a Λ-coalescent in which Λ has an atom at 1 in terms of a coalescent with the atom at 1 removed, Example 3.16.

Thanks to the above examples, we always can assume that Λ has no atoms at 0 or 1. Moreover, Pitman shows in [Pit99, Proposition 23] that, if $\Lambda \in \mathcal{M}_f([0,1])$ has no atom at 1, then the Λ-coalescent Π either comes down from infinity or stays infinite for all times, almost surely.

Proposition 3.19. *Assume Λ does not have an atom at 1. Then the Λ-coalescent Π either comes down from infinity or stays infinite for all times, almost surely.*

In order to prove this result, we have to recall [Pit99, Theorem 4].

Theorem 3.20. *Let Π be a Λ-coalescent, $i, j \in \mathbb{N}$ in distinct blocks of Π_0 and $\tau_{i,j}$ the collision time of i and j. If the event $\{N(\tau_{i,j}-) = \infty\}$ has strictly positive probability, then given this event a random variable $X_{i,j}$ with distribution Λ is recovered as the almost sure relative frequency of blocks of $\Pi_{\tau_{i,j}-}$ which merge at time $\tau_{i,j}$ to form the block containing both i and j.*

Proof of Proposition 3.19. Let

$$T := \inf\{t : N(t) < \infty\}.$$

We argue by contradiction, therefore, assume that

$$\mathbb{P}(0 < T < \infty) > 0.$$

Then, we find on $\{0 < T < \infty\}$ that $N(T) < \infty$ a.s. In addition, by the definition of T, we have $N(T-) = \infty$ a.s. Hence, T is a collision time. By Theorem 3.20, we recover a random variable X_T as the almost sure relative frequency of blocks of Π_{T-} which merge at time T. Then, by assumption on Λ,

$$\mathbb{P}(X_T = 1) = \Lambda(\{1\}) = 0.$$

But, since T is the time coming down from infinity, there merge infinitely many blocks and just finitely many blocks are left over. Hence, the frequency is 1, almost surely, i.e. $\mathbb{P}(X_T = 1) = 1$. A contradiction. □

Proposition 3.21. *The Λ-coalescent stays infinite if $\int_0^1 x^{-1}\Lambda(dx) < \infty$.*

Proof. By Corollary 3.13, the rate at which a given block takes part in a merger is finite by assumption. Hence, there are singletons which have never taken part in a merger. By exchangeability, the Λ-coalescent stays infinite. □

Recall from Section 3.2, Theorem 3.10, that the rate $\lambda_{b,k}$ at which a k-tuple of blocks, given b blocks, is merging to form a single block, is given by

$$\lambda_{b,k} = \int_0^1 x^{k-2}(1-x)^{b-k}\Lambda(dx).$$

We define

$$\gamma_b := \sum_{k=2}^{b}(k-1)\binom{b}{k}\lambda_{b,k},$$

which is the rate at which the number of blocks is decreasing because merging k blocks into one decreases the number of blocks by $k-1$, i.e.

$$\mathbb{P}[N(t+dt) \mid N(t) = b] = b - \gamma_b dt.$$

In addition, let

$$\eta_b := \sum_{k=2}^{b} k\binom{b}{k}\lambda_{b,k}.$$

In [Sch00b, Theorem 1] we have the following characterization for a Λ-coalescent to come down from infinity.

Theorem 3.22 (Coming down from infinity)**.** *The Λ-coalescent comes down from infinity if and only if*

$$\sum_{b=2}^{\infty}\gamma_b^{-1} < \infty. \tag{3.2}$$

We give a detailed proof of Theorem 3.22 following the original proof of Schweinsberg later in this section mainly in Proposition 3.29 and Proposition 3.31. Firstly, we have a look at some corollaries and examples.

Corollary 3.23. *The Λ-coalescent comes down from infinity if and only if*

$$\sum_{b=2}^{\infty}\eta_b^{-1} < \infty.$$

Roughly speaking, the Λ-coalescent stays infinite whenever the η_b don't grow too rapidly as $b \to \infty$, to be more precise, whenever η_b don't grow faster than $O(b)$.

Proof. Following [Sch00b, Corollary 2], since for all $k \geq 2$ we have that $1 \leq k/(k-1) \leq 2$, we find that for all $b \geq 2$,

$$(2\gamma_b)^{-1} \leq \eta_b^{-1} \leq \gamma_b^{-1}.$$

By Theorem 3.22, we are done. □

Example 3.24.

(a) The Kingman's coalescent has rates $\lambda_{b,k} = 1$ for $k = 2$ and $\lambda_{b,k} = 0$ otherwise. Therefore, we have $\gamma_b = \binom{b}{2}$ and $\sum_{b=2}^{\infty} \gamma_b^{-1} = \binom{b}{2}^{-1} = 1$. So, the Kingman's coalescent comes down from infinity.

This result is surprising, since in a Kingman's coalescent only two blocks are allowed to coalesce. Imagine a coalescent in which many more blocks merge at once, one may think that this coalescent should also come down from infinity. But this is not true in general. Indeed, the Kingman's coalescent is the one in which coalescence is strongest in the sense that among all Λ with $\Lambda([0,1]) = 1$ we have for all $\epsilon > 0$ and for all t sufficiently small,

$$N_\Lambda(t) \geq \frac{2}{t}(1-\epsilon),$$

almost surely, where $2/t$ is the speed of coming down from infinity for the Kingman's coalescent, i.e.

$$\lim_{t \to 0} \frac{N(t)}{2/t} = 1,$$

almost surely. See also Remark 3.32.

(b) By Example 3.15, the rates of the Bolthausen-Sznitman coalescent are given by

$$\lambda_{b,k} = \frac{(k-2)!(b-k)!}{(b-1)!}.$$

Hence, we find $\gamma_b = \sum_{k=2}^{b} \frac{b}{k}$. Since $b \ln b \geq \gamma_b$, we have to show that $\sum_{b \geq 2} 1/(b \ln b) = \infty$. Note, that

$$\sum_{b \geq 2} 1/(b \ln b) \geq \int_{x \geq 2} 1/(x \ln x) dx = \int \frac{d}{dx} \ln \ln x dx.$$

As $\ln \ln x \xrightarrow{x \to \infty} \infty$, we are done.

Corollary 3.25. *Let* Λ_1, Λ_2 *be in* $\mathscr{M}_1([0,1])$ *such that for all* $x \in [0,1]$*,*

$$\Lambda_1([0,x]) \geq \Lambda_2([0,x]).$$

If the Λ_1*-coalescent stays infinite, then the* Λ_2*-coalescent stays infinite. If the* Λ_2*-coalescent comes down from infinity, then the* Λ_1*-coalescent comes down from infinity.*

Proof. We just give a sketch, for more details see [Sch00b, Lemma 10].
Let $\eta_b^{(1)}$ be the rate for the Λ_1-coalescent as defined above and let $\eta_b^{(2)}$ be the rate for the Λ_2-coalescent, respectively. Then it follows from the assumption that $\eta_b^{(1)} \geq \eta_b^{(2)}$. We conclude by Corollary 3.23. □

Now, let for $0 < a < b \leq 1$, $\Lambda_{[a,b]}$ denote the *restriction* of Λ to $[a,b]$.

Lemma 3.26. *Let* $a > 0$*. If the* $\Lambda_{[0,a]}$*-coalescent stays infinite, then the* Λ*-coalescent stays infinite.*

Proof. We follow [Sch00b, Lemma 8] and consider the Poisson constructions of the Λ-coalescent Π, the $\Lambda_{[0,a]}$-coalescent $\Pi_{[0,a]}$ and the $\Lambda_{(a,1]}$-coalescent $\Pi_{(a,1]}$, respectively, as described in Section 3.2. Since the total rate at which merges are occurring converges to the second moment of the corresponding measure, Corollary 3.13, the holding time of the initial state has an exponential distribution with parameter $\int_0^1 x^{-2}\Lambda_{(a,1]}(dx)$, [Pit99, Section 2.1]. In addition, we have

$$\int_0^1 x^{-2}\Lambda_{(a,1]}(dx) \leq a^{-2}\Lambda_{(a,1]}([0,1]) < \infty.$$

Therefore, there is $p \in (0,1]$, such that $\mathbb{P}(\Pi_t = \Pi_{[0,a],t}) \geq p$. By assumption, the $\Lambda_{[0,a]}$-coalescent stays infinite. Hence, we find $\#\Pi_t = \infty$ with probability at least p. It follows that the Λ-coalescent also stays infinite. □

Corollary 3.27.

(a) If there is $\epsilon > 0$ *and* $M < \infty$ *such that for all* $\delta \in [0,\epsilon]$ *we have*

$$\Lambda([0,\delta]) \leq M\delta,$$

then the Λ*-coalescent stays infinite.*

(b) *If there is $\epsilon > 0$, $M > 0$ and $\alpha > 0$ such that for all $\delta \in [0, \epsilon]$ we have*

$$\Lambda([0,\delta]) \geq M\delta^{\alpha},$$

then the Λ-coalescent comes down from infinity.

Proof. The proof is from [Sch00b, Proposition 11, 13].

(a) Let $\Lambda_{[0,\epsilon]}$ be the restriction of Λ to $[0, \epsilon]$. By Lemma 3.26, it suffices to show that the $\Lambda_{[0,\epsilon]}$-coalescent stays infinite. By Example 3.24, for the uniform distribution, the $U([0, 1])$-coalescent stays infinite. Since multiplying U with a constant M leads to multiplying each γ_b by M, the $MU([0, 1])$-coalescent also stays infinite. Then, for all $x \in [0, 1]$,

$$\Lambda_{[0,\epsilon]}([0, x]) \leq Mx = MU([0, 1])([0, x]).$$

We conclude by Corollary 3.25.

(b) By Corollary 3.25, we assume without loss of generality that

$$\Lambda([0,\delta]) = M\delta^{\alpha}.$$

Therefore, we have a Radon-Nikodym derivative $M\alpha x^{\alpha-1}$ of Λ with respect to the Lebesgue measure which makes it possible to find for all $b \geq 2$ a lower bound $Cb^{2-\alpha}$, $C > 0$, for η_b. Hence, $\sum_{b\geq 2} \eta_b^{-1} < \infty$, so the Λ-coalescent comes down from infinity. □

Example 3.28.

(a) For $0 < y < 1$ let $\Lambda := \delta_y$. Take $\epsilon := \frac{y}{2}$. Then, we find for all $\delta \in [0, \epsilon]$ that $\Lambda([0, \delta]) = \delta_{2\epsilon}([0, \delta]) = 0 \leq \delta$. By Proposition 3.27 (a), the Λ-coalescent stays infinite.

(b) Consider the Beta-coalescent as defined in Example 3.15. If $\alpha < 0$, then the coalescent has singletons for all times, therefore it stays infinite. If $\alpha = 1$, we find the Bolthausen-Sznitman coalescent, which stays infinite but has no dust. For $\alpha > 1$, the coalescent comes down from infinity. Finally, if $\alpha \to 2$, we find the Kingman's coalescent, [Ber09].

Proof of Theorem 3.22. We turn toward the proof of Theorem 3.22. For the necessity of equation (A.3), we define the stopping time

$$T_n := \inf\{t : \#\Pi_t^{(n)} = 1\}.$$

Then, we naturally find that $0 = T_1 < T_2 \leq \cdots \leq T_\infty \leq \infty$, [Pit99]. In addition, the Λ-coalescent comes down from infinity iff $T_\infty < \infty$, moreover, iff $\mathbb{P}[T_\infty] < \infty$, [Sch00b, Proposition 5]. Since $(\mathbb{P}[T_n])_{n\geq 1} \uparrow \mathbb{P}[T_\infty]$ we find by monotone convergence, Proposition 4.9, that the Λ-coalescent comes down from infinity iff $(\mathbb{P}[T_n])_{n\geq 1}$ is bounded.

In addition, we define for $N^{(n)}(t) := \#\Pi_t^{(n)}$ stopping times $R_0 := 0$ and for $i \geq 1$,

$$R_i := \begin{cases} \inf\{t : N^{(n)}(t) < N^{(n)}(R_{i-1})\}, & \text{if } N^{(n)}(R_{i-1}) > 1 \\ R_{i-1}, & \text{if } N^{(n)}(R_{i-1}) = 1 \end{cases}.$$

That is the time at which $\Pi^{(n)}$ decreases about at least one block. Let $J_i := N^{(n)}(R_{i-1}) - N^{(n)}(R_i)$ be the decrease of N at this collision.

Proposition 3.29. *If $\sum_{b=2}^{\infty} \gamma_b^{-1} < \infty$, then the Λ-coalescent comes down from infinity.*

Proof. By the above comment, it suffices to show that $(\mathbb{P}[T_n])_{n\geq 1}$ is bounded. We follow [Sch00b, Lemma 6] and define $L_i := R_i - R_{i-i}$. Then, we find on the set $\{N^{(n)}(R_{i-1}) > 1\}$ that

$$\mathbb{P}[L_i \mid N^{(n)}(R_{i-1})] = \lambda^{-1}_{N^{(n)}(R_{i-1})},$$

where

$$\lambda_b := \sum_{k=2}^{b} \binom{b}{k} \lambda_{b,k}$$

is the total rate at which mergers are occurring. Moreover, on the same set we have

$$\mathbb{P}[J_i \mid N^{(n)}(R_{i-1})] = \gamma_{N^{(n)}(R_{i-1})} \lambda^{-1}_{N^{(n)}(R_{i-1})},$$

since

$$\mathbb{P}(J_i = k - 1 \mid N^{(n)}(R_{i-1}) = b) = \binom{b}{k} \frac{\lambda_{b,k}}{\lambda_k}.$$

It follows, that

$$\begin{aligned}
\mathbb{P}[T_n] = \mathbb{P}[R_{n-1}] &= \sum_{i=1}^{n-1} \mathbb{P}[L_i] = \sum_{i=1}^{n-1} \mathbb{P}\Big[\mathbb{P}[L_i \mid N^{(n)}(R_{i-1})]\Big] \\
&= \sum_{i=1}^{n-1} \mathbb{P}[\lambda^{-1}_{N^{(n)}(R_{i-1})} \mathbb{1}_{\{N^{(n)}(R_{i-1})>1\}}] \\
&= \sum_{i=1}^{n-1} \mathbb{P}\Big[\gamma^{-1}_{N^{(n)}(R_{i-1})} \mathbb{P}[J_i \mid N^{(n)}(R_{i-1})] \mathbb{1}_{\{N^{(n)}(R_{i-1})>1\}}\Big] \\
&= \sum_{i=1}^{n-1} \mathbb{P}\Big[\mathbb{P}[\gamma^{-1}_{N^{(n)}(R_{i-1})} J_i \mathbb{1}_{\{N^{(n)}(R_{i-1})>1\}} \mid N^{(n)}(R_{i-1})]\Big].
\end{aligned}$$

Note that $J_i = 0$ on the set $\{N^{(n)}(R_{i-1}) = 1\}$. Hence, we find

$$\begin{aligned}
\mathbb{P}[T_n] &= \sum_{i=1}^{n-1} \mathbb{P}\Big[\mathbb{P}[\gamma^{-1}_{N^{(n)}(R_{i-1})} J_i \mid N^{(n)}(R_{i-1})]\Big] \\
&= \sum_{i=1}^{n-1} \mathbb{P}[\gamma^{-1}_{N^{(n)}(R_{i-1})} J_i] = \mathbb{P}[\sum_{i=1}^{n-1} \gamma^{-1}_{N^{(n)}(R_{i-1})} J_i] \qquad (3.3) \\
&= \mathbb{P}[\sum_{i=1}^{n-1} \sum_{j=0}^{J_i - 1} \gamma^{-1}_{N^{(n)}(R_{i-1})}].
\end{aligned}$$

Finally, since the sequence $(\gamma_b)_{b\geq 2}$ is increasing, [Sch00b, Lemma 3], we conclude

$$\begin{aligned}
\mathbb{P}[T_n] &\leq \mathbb{P}[\sum_{i=1}^{n-1} \sum_{j=0}^{J_i - 1} \gamma^{-1}_{N^{(n)}(R_{i-1})-j}] = \mathbb{P}[\sum_{b=2}^{n} \gamma_b^{-1}] \\
&< \sum_{b=2}^{n} \gamma_b^{-1}.
\end{aligned}$$

By assumption, $\sum_{b=2}^{\infty} \gamma_b^{-1} < \infty$. Hence, $(\mathbb{P}[T_n])_{n\geq 1}$ is bounded and we are done. □

Now, we turn toward the sufficiency in Theorem 3.22, which is Proposition 3.31.

Lemma 3.30. *Let Λ be concentrated on $[0, 1/2]$. If $\sum_{b=2}^{\infty} \gamma_b^{-1} = \infty$, then the Λ-coalescent stays infinite.*

Proof. By the comment at the beginning of this paragraph, it suffices to show that $\mathbb{P}[T_\infty] = \infty$. We follow [Sch00b, Lemma 7] and define for $n, l \in \mathbb{N}$,

$$\begin{aligned} D_l &: = \{2^{l-1} + 1 \le N^{(n)}(t) \le 2^l, \text{ for some t}\}, \\ K &: = \inf\{i : N^{(n)}(R_{i-1}) \le 2^l\}. \end{aligned}$$

Note, that we suppress n at the lefthand side of the above definitions.

By considering the total rate of all collisions that would take the Λ-coalescent given n blocks down to 2^l or fewer blocks and the total rate of all collisions that would take the coalescent down to between $2^{l-1} + 1$ and 2^l blocks and some calculation, see [Sch00b, Lemma 7] for a more detailed analysis, we find,

$$\mathbb{P}(D_l \mid N^{(n)}(K-1)) \ge \frac{1}{2}.$$

In addition, for $n = 2^m$, $m \in \mathbb{N}$ and $j \in \{2, \dots, n\}$, we define

$$L_n(j) := \min\{s \ge j : N^{(n)}(t) = s, \text{ for some t}\}.$$

If $N^{(n)}(R_{i-1}) \ge j \ge N^{(n)}(R_i)$, we find $L_n(j) = N^{(n)}(R_{i-1})$. Hence, by using equation (3.3) from the proof of Proposition 3.29, it follows

$$\begin{aligned} \mathbb{P}[T_n] &= \sum_{i=1}^{n-1} \mathbb{P}[\gamma^{-1}_{N^{(n)}(R_{i-1})} J_i] = \sum_{j=2}^{n} \mathbb{P}[\gamma^{-1}_{L_n(j)}] \\ &= \sum_{l=1}^{m} \sum_{j=2^{l-1}+1}^{2^l} \mathbb{P}[\gamma^{-1}_{L_n(j)}]. \end{aligned}$$

Since $(\gamma_b)_{b \ge 2}$ is increasing, [Sch00b, Lemma 3], and if $j \le 2^l$ then $L_n(j) \le 2^{l+1}$ on D_{l+1}, we find

$$\begin{aligned} \mathbb{P}[T_n] &\ge \sum_{l=1}^{m-1} \sum_{j=2^{l-1}+1}^{2^l} \mathbb{P}[\gamma^{-1}_{L_n(j)}] \ge \sum_{l=1}^{m-1} \sum_{j=2^{l-1}+1}^{2^l} \mathbb{P}(D_{l+1}) \gamma^{-1}_{2^{l+1}} \\ &\ge \sum_{l=1}^{m-1} 2^{l-1} \cdot \frac{1}{2} \cdot \gamma^{-1}_{2^{l+1}} \ge \frac{1}{8} \sum_{l=1}^{m-1} 2^{l+1} \gamma^{-1}_{2^{l+1}}. \end{aligned}$$

Finally, since the sequence $(T_n)_{n \ge 1}$ is monotonically increasing,

$$\begin{aligned} \lim_{n\to\infty} \mathbb{P}[T_n] = \lim_{m\to\infty} \mathbb{P}[T_{2^m}] &\ge \lim_{m\to\infty} \frac{1}{8} \sum_{l=1}^{m-1} 2^{l+1} \gamma^{-1}_{2^{l+1}} \\ &\ge \frac{1}{8} \sum_{l \ge 4} \gamma_l^{-1}. \end{aligned}$$

By assumption, $\sum_{b\geq 2}\gamma_b^{-1}=\infty$. Hence, $\mathbb{P}[T_\infty]=\infty$. □

Proposition 3.31. *If $\sum_{b=2}^{\infty}\gamma_b^{-1}=\infty$, then the Λ-coalescent stays infinite.*

Proof. The proof follows [Sch00b, Lemma 9].
Let $\gamma_{b,[0,1/2]}$ be the analogous quantity for the $\Lambda_{[0,1/2]}$-coalescent. Then, clearly, for all $b\geq 2, \gamma_{b,[0,1/2]}\leq\gamma_b$. Hence, $\sum_{b\geq 2}(\gamma_{b,[0,1/2]})^{-1}=\infty$. By Lemma 3.30, the $\Lambda_{[0,1/2]}$-coalescent stays infinite. We conclude by using Lemma 3.26. □

Remark 3.32. There are several more remarkable results on coming down from infinity. One result worth mentioning is the following characterization of coming down from infinity due to Berestycki, [Ber09, Theorem 4.9]. Define

$$\psi(q)=\int_0^1(e^{-qx}-1+qx)x^{-2}\Lambda(dx),$$

then the Λ-coalescent comes down from infinity iff for some $t>0$,

$$\int_t^\infty\psi(q)^{-1}dq<\infty.$$

Note that if the integral is finite for same $t>0$ it is finite for all $t>0$. Therefore, we define

$$u(t):=\int_t^\infty\psi(q)^{-1}dq$$

and its *cádlág inverse*

$$v(t):=\inf\{s>0:\int_s^\infty\psi(q)^{-1}dq<t\}.$$

The second result we want to mention here is the following on the speed of coming down from infinity from [BL10, Theorem 1], i.e.

$$\lim_{t\to 0}\frac{N(t)}{v(t)}=1\text{ a.s.},$$

where $N(t)$ is the number of blocks in a coalescent at time t. For instance, we have for the Kingman's coalescent, $v(t)=2/t$.

3.4 Coming down from infinity and compactness

In this section we put together Chapter 2 with what we have seen so far in this chapter. Our goal is to give another characterization of a Λ-coalescent to come down from infinity by identifying the process with an element in $\mathbb{M}$, the Λ-coalescent measure tree. First, we have to ensure the existence of such a space in Theorem 3.33 which is due to Greven et al. Then the main part here is Theorem 3.36.

Let $\Pi = (\Pi_t : t \geq 0)$ be a Λ-coalescent. Then for all initial partitions $\mathcal{P} \in \mathbb{S}$ and $\mathbb{P}^{\Lambda,\mathcal{P}}$-almost all sample paths of Π, there is a metric d^Π on $\mathbb{N}$ defined by

$$d^\Pi(i,j) := \inf\{t \geq 0 : i \sim_{\mathcal{P}} j\}.$$

That is the time needed for i and j to coalesce. It is clear that d^Π is an ultra metric. We denote by (L^Π, d^Π) the completion of $(\mathbb{N}, d^\Pi)$. The extension of d^Π to L^Π is also an ultra-metric. Since ultra-metric spaces are associated with tree-like structures, [GPW10, Remark 2.2], we call the metric space (L^Π, d^Π) equipped with a »uniform distribution« the Λ-*coalescent measure tree*. In order to equip (L^Π, d^Π) with this probability measure, we use in the following a limit procedure. In [GPW09], there is a characterization of existence and uniqueness of the Λ-coalescent measure tree.

Define H_n to be the map which takes a realization of the Λ-coalescent and maps it to a metric measure space as follows,

$$H_n : \Pi \mapsto (L^\Pi, d^\Pi, \mu_n^\Pi := \frac{1}{n}\sum_{i=1}^n \delta_i).$$

Moreover, define for given $\mathcal{P}_0 \in \mathbb{S}$ the distribution of H_n by

$$\mathbb{Q}^{\Lambda,n} := (H_n)_* \mathbb{P}^{\Lambda,\mathcal{P}_0}.$$

Recall from Definition 3.12, the notion of the dust-free property.

Theorem 3.33 (The Λ-coalescent measure tree)**.** *The family* $\{\mathbb{Q}^{\Lambda,n} : n \in \mathbb{N}\}$ *converges in the weak topology with respect to the Gromov-weak topology if and only if the dust-free property,*

$$\mathbb{P}^{\Lambda,\mathcal{P}_0}\{\forall j \in \mathbb{N} : \tilde{f}((\Pi_t)^j) = 0\} = 0,$$

holds.

Proof. By Proposition 2.45 and Remark 2.46, we have to show that the family $\{\mathbb{Q}^{\Lambda,n}[w_{\mathcal{X}}] : n \in \mathbb{N}\}$ is tight and that for all $\epsilon > 0$ there is $\delta > 0$ such that $\sup_{n\in\mathbb{N}} \mathbb{Q}^{\Lambda,n}[\mu\{x : \mu(B_\epsilon(x)) \leq \delta\}] < \epsilon$.

By definition, for all $n \in \mathbb{N}$, $\mathbb{Q}^{\Lambda,n}[w_{\mathcal{X}}]$ is exponentially distributed with parameter $\lambda_{2,2}$. Hence, the family is tight. For the latter, by the uniform distribution and exchangeability, we get

$$\begin{aligned}
&\mathbb{Q}^{\Lambda,n}[\mu\{x : \mu(B_\epsilon(x)) \leq \delta\}] \\
&= \mathbb{P}^{\Lambda,\mathcal{P}_0}[\frac{1}{n}\sum_{i=1}^{n} \mu_n^\Pi\{x \in L^\Pi : \mu_n^\Pi(B_\epsilon(x)) \leq \delta \mid x = i\}] \\
&= \mathbb{P}^{\Lambda,\mathcal{P}_0}[\mu_n^\Pi\{x \in L^\Pi : \mu_n^\Pi(B_\epsilon(x)) \leq \delta \mid x = 1\}] \\
&= \mathbb{P}^{\Lambda,\mathcal{P}_0}[\mu_n^\Pi(B_\epsilon(1)) \leq \delta].
\end{aligned}$$

By de Finetti's Theorem, Theorem 4.26, $\mu_n^\Pi(B_\epsilon(1)) \xrightarrow{n\to\infty} \tilde{f}((\Pi_\epsilon)^1)$. Hence,

$$\begin{aligned}
\lim_{\delta\to 0}\lim_{n\to\infty} \mathbb{Q}^{\Lambda,n}[\mu\{x : \mu(B_\epsilon(x)) \leq \delta\}] &= \lim_{\delta\to 0} \mathbb{P}^{\Lambda,\mathcal{P}_0}[\tilde{f}((\Pi_\epsilon)^1) \leq \delta] \\
&= \mathbb{P}^{\Lambda,\mathcal{P}_0}[\tilde{f}((\Pi_\epsilon)^1) = 0]. \qquad \square
\end{aligned}$$

Remark 3.34 (Open question). We attempted to extend this criterion to the subspace of compact metric measure spaces $\mathbb{M}_c$ to find a characterization of a compact limit object. But as we discuss in the paragraph at the end of Section 2.2, it was hard to give a well characterization of relative compactness in $\mathbb{M}_c$. See also Remark 2.49.

Recall from Corollary 3.13, the different equivalent assumptions on a Λ-coalescent to have the dust-free property.

Example 3.35.

(a) For the Kingman's coalescent, i.e. $\Lambda = \delta_0$, we have $\int_0^1 x^{-1}\Lambda(dx) = \int_0^1 x^{-1}\delta_0(dx) = \infty$. Therefore, the Kingman's coalescent measure tree exists.

(b) If $\Lambda = U([0,1])$, the uniform distribution, the Λ-coalescent is the Bolthausen-Sznitman coalescent. Since

$$\int_0^1 x^{-1}\Lambda(dx) = \int_0^1 x^{-1}U([0,1])(dx) = \infty,$$

the Bolthausen-Sznitman coalescent measure tree exists.

(c) The Beta-coalescent as defined in Example 3.15 and considered in Example 3.28 has singletons iff $\alpha < 1$. Therefore the corresponding metric measure tree exists iff $\alpha \geq 1$.

(d) Let $\Lambda = \delta_1$. Then $\int_0^1 x^{-1}\Lambda(dx) = \int_0^1 x^{-1}\delta_1(dx) = 1$. Hence, a corresponding metric measure space does not exist.

Compactness-characterization of a Λ-coalescent. As we see in Section 3.3, Proposition 3.21, the Λ-coalescent stays infinite if

$$\int_0^1 x^{-1}\Lambda(dx) < \infty.$$

Conversely, we show in Corollary 3.13 that the dust-free property and hence the existence of the Λ-coalescent measure tree is equivalent to

$$\int_0^1 x^{-1}\Lambda(dx) = \infty.$$

Therefore, we look for a quality of the Λ-coalescent measure tree to decide if the corresponding Λ-coalescent comes down from infinity or not. For the Kingman's coalescent, which is known to come down from infinity, the Kingman's coalescent measure tree is compact, [Eva00]. The Bolthausen-Sznitman coalescent has infinitely many partitions for all times, therefore, the corresponding metric measure space is not compact.

We show that a Λ-coalescent comes down from infinity if and only if the corresponding metric measure space is compact. The idea is not new but we give here a detailed proof. Moreover, we show that if the Λ-coalescent stays infinite, then the corresponding metric measure space is not even locally compact.

For a Λ-coalescent $\Pi = (\Pi_t)_{t\geq 0}$, we denote the corresponding Λ-coalescent measure tree by $\mathcal{L} = \mathcal{L}^\Pi = \overline{(L^\Pi, d^\Pi, \mu^\Pi)} \in \mathbb{M}$. As usual, let $N(t) := \#\Pi_t$ denote the number of blocks in the partition Π_t.

Theorem 3.36 (Characterization of Λ-coalescent)**.**

(a) The Λ-coalescent comes down from infinity if and only if the corresponding metric measure space is compact.

(b) If the Λ-coalescent stays infinite, then the corresponding metric measure space is not even locally compact.

Proof. (a) Assume first, that the Λ-coalescent comes down from infinity. Using Corollary 2.33, it suffices to show that for all $\epsilon > 0$, $\xi_\epsilon(\mathcal{L}) < \infty$, almost surely. Since $\xi_\epsilon(\mathcal{L}) \leq N(\epsilon) < \infty$ by assumption, we are done.

We prove the converse by contradiction. Hence, assume $\mathcal{L}$ is compact and Π stays infinite for some time $\epsilon > 0$. Recall from Theorem 3.33, that the Λ-coalescent measure tree exists iff the dust-free property,

$$\mathbb{P}^{\Lambda,\mathcal{P}_0}\{\forall j \in \mathbb{N} : \tilde{f}((\Pi_t)^j) > 0\} = 1,$$

holds. Hence, it follows that

$$\hat{\mu}_{\mathcal{X}}(\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \leq \delta\}) > 0.$$

Since $\mathcal{L}$ is compact, we have by Corollary 2.33, that there is a random $\Delta > 0$ with

$$\hat{\mu}_{\mathcal{X}}(\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \leq \Delta\}) = 0,$$

almost surely. In particular, there is $\delta > 0$, such that

$$\hat{\mu}_{\mathcal{X}}(\nu([0,\epsilon)) \leq \Delta) = 0$$

with positive probability. A contradiction.

(b) By Corollary 3.13, the total collision rate of the block containing 1 and any other blocks is infinite. Let $0 < \epsilon < \delta$ and take the ball with radius δ around 1 in L^{Π}. Considering the times between ϵ and δ, there are infinitely many lines coalesce to the line containing 1. Hence, there is a infinite ϵ-separated set in $B_\delta(1)$. It follows that

$$\nu^{\mathcal{L}}(\{R : \xi_\epsilon(\tau_\delta(R)) < \infty\}) = 0,$$

almost surely. So, for any sequence $0 < \epsilon_n < \delta_n$ with $\delta_n \xrightarrow{n\to\infty} 0$, we have

$$\nu^{\mathcal{L}}\Big(\bigcap_{0<\epsilon<\delta_n} \{R : \xi_\epsilon(\tau_{\delta_n}(R)) < \infty\}\Big) \leq \nu^{\mathcal{L}}(\{R : \xi_{\epsilon_n}(\tau_{\delta_n}(R)) < \infty\}) = 0,$$

almost surely. We conclude by Theorem 2.36. □

CHAPTER 4

Foundations

In the last chapter we summarize well known vocabulary and results on metric topology, probability and measure theory. In Section 4.2, we discuss the space of (isometry classes of) compact metric spaces. There is a lot of fitting literature and most statements are known from the first two or three courses in probability theory. But we seek completeness and self-containment of this work, hence we prove most results we have directly referred to in several situations throughout the different chapters.

4.1 Metric spaces

In this section we recall some vocabulary on (ultra) metrics and their induced topology. The characterization of the compact set of a metric spaces, Proposition 4.1, is frequently used throughout this work, [Rin75], [SJAS78] or [Lim04].

A *metric space* is a set X together with a *positive definite symmetric* map $d : X \times X \to \mathbb{R}$ and such that the *triangle inequality* holds, i.e. for all $x, y, z \in X$, we have

(a) $d(x, y) \geq 0$ and $d(x, y) = 0$ iff $x = y$,

(b) $d(x, y) = d(y, x)$,

(c) $d(x,z) \leq d(x,y) + d(y,z)$.

The map d is called a *metric* on X. The metric d is called an *ultra metric*, if it satisfies the *strong triangle inequality*, i.e. for all $x, y, z \in X$,

$$d(x,z) \leq d(x,y) \vee d(y,z),$$

where $p \vee q := max\{p,q\}$.

We define by

$$B_r(x) := \{y \in X : d(x,y) < r\}$$

the *open ball* of radius r with center x and by $\bar{B}_r(x) := \{y \in X : d(x,y) \leq r\}$ the *closed ball*. The *diameter* of a nonempty subset $A \subset X$ is defined by

$$diam(A) := sup\{d(x,y) : x, y \in A\}.$$

The set A is *bounded* if $diam(A) < \infty$. The distance between two nonempty subsets $A, B \subset X$ is $dist(A,B) := \inf\{d(a,b) : a \in A, b \in B\}$. In addition, the distance between a point $x \in X$ and a nonempty subset $A \subset X$ is defined by $dist(x,A) := dist(\{x\}, A)$.

A sequence $(x_n)_{n\in\mathbb{N}}$ is called a *Cauchy-sequence* if for any $\epsilon > 0$ there exists an n_0 such that $d(x_n, x_m) < \epsilon$ for all $n, m \geq n_0$. A metric d is called *complete* if every Cauchy sequence converges. By an abuse of notation, we often say, that the metric space is complete. A point $x \in X$ is called an *accumulation point* of a set if in every neighborhood of x there are infinitely many points of the set.

A metric d induces a topology on X as follows: A set $U \subset X$ is *open* iff for every point $x \in U$ there is an $\epsilon > 0$ such that $B_\epsilon(x) \subset U$. For $A \subset X$ the *closure* $\bar{A}$ of A is the intersection of all closed sets containing A. Analogue, we define the *interior* $\mathring{A}$ of A as the union of all open sets included in A. The metric space X is *separable* if there exists a countable subset $A \subset X$ dense in X, where *dense* means, that every point in X can be approximated arbitrarily close by a sequence in A. We call (X,d) *Polish* if X is separable and complete.

A subset $A \subset X$ is *compact* if any open covering of A has a finite subcovering. Moreover, A is *locally compact* if every point in A has a compact neighborhood. A is called *relatively compact* if the closure $\bar{A}$ is compact.

Moreover, we call A *totally bounded* if for any $\epsilon > 0$, A can be covered by N_ϵ open balls of radius ϵ. For $\epsilon > 0$, a set $S \subset X$ is called an *ϵ-net* for X if $dist(x, S) \leq \epsilon$ for every $x \in X$. Indeed, X is totally bounded if for any $\epsilon > 0$, there is a finite ϵ-net in X. A subset $S \subset X$ is called *ϵ-separated* if $d(x, y) \geq \epsilon$ for any two different points $x, y \in S$. A maximal ϵ-separated set is an ϵ-net.

Proposition 4.1 (Compactness characterization for metric spaces). *Let (X, d) be a metric space. Then the following is equivalent.*

(a) X is compact.

(b) Any sequence in X has a converging subsequence.

(c) X is complete and totally bounded.

Proof. $(a) \Rightarrow (b)$. Let $(x_n)_{n\geq 1}$ be a sequence in X. The proof is by contradiction. Hence, assume $(x_n)_{n\geq 1}$ has no accumulation point. Then there is for all $y \in X$ a neighborhood U_y of y which contains finitely many x_n. Since X is compact, the covering $(U_y)_{y\in X}$ has a finite subcovering $(U_y)_{y\in Y}$, $Y \subset X$ finite. Moreover, $X = \bigcup_{y\in Y} U_y$ contains only finitely many x_n. A contradiction.

$(b) \Rightarrow (c)$. To get completeness, consider a Cauchy-sequence, then by assumption, there is a converging subsequence. Hence, the Cauchy-sequence is convergent itself.

The second part is by contradiction. Assume, X is not totally bounded, then there is $\epsilon > 0$ such that X can not be covered by finitely many ϵ-balls. Therefore, we can construct a sequence $(x_n)_{n\in\mathbb{N}}$ as follows. Take $x_1 \in X$ arbitrarily. Given $x_1, \ldots, x_{n-1} \in X$, take $x_n \in X$, such that $x_n \notin \bigcup_{i=1}^{n-1} B_\epsilon(x_i)$. It follows, that for all $m \neq n$, $d(x_m, x_n) \geq \epsilon$. Hence, $(x_n)_{n\in\mathbb{N}}$ has no convergent subsequence. A contradiction.

$(c) \Rightarrow (a)$. Assume X is not compact, then there is a covering $\mathscr{U}$ of X without finite subcovering. Since X is totally bounded, there is a finite set $Y \subset X$ such that $X = \bigcup_{x\in Y} B_1(x)$. Therefore, there is $x_0 \in Y$, such that $U_0 := B_1(x_0)$ can not be covered by finitely many sets of $\mathscr{U}$. Moreover, $B_1(x_0)$ is also totally bounded. Hence, we find $x_1 \in B_1(x_0)$ such that the ball $U_1 := B_{1/2}(x_1)$ can not be covered by finitely many sets of $\mathscr{U}$. Thus, we find inductively a sequence $(x_n)_{n\in\mathbb{N}}$ such that $x_{n+1} \in B_{1/2^n}$ and such that $U_n := B_{1/2^n}(x_n)$ can not be covered by finitely many sets of $\mathscr{U}$. In addition, for $m > n$,

$$d(x_m, x_n) \leq \sum_{j=n}^{m-1} d(x_j, x_{j+1}) \leq \sum_{j=n}^{m-1} \frac{1}{2^j} \leq \frac{1}{2^{n-1}},$$

i.e. $(x_n)_{n\in\mathbb{N}}$ is a Cauchy-sequence. By assumption, $(x_n)_{n\mathbb{N}}$ is converging to $x \in X$. Since $\mathscr{U}$ is a covering of X there is $V \in \mathscr{U}$ such that $x \in V$. We take $r > 0$ such that $U := B_r(x) \subset V$ and $n \in \mathbb{N}$ such that

$$d(x_n, x) < \frac{r}{2} \text{ and } \frac{1}{2^n} < \frac{r}{2}.$$

For $y \in U_n$, then

$$d(x, y) \leq d(x, x_n) + d(x_n, y) < \frac{r}{2} + \frac{1}{2^n} < r,$$

i.e. $U_n \subset U \subset V$. Note, that $\{V\} \subset \mathscr{U}$ is a finite subcovering of U_n. A contradiction. $\square$

Proposition 4.2. *Let for all $n \in \mathbb{N}$, (X_n, r_n) be a Polish space. Then $X := \prod_{n\in\mathbb{N}} X_n$ is Polish with respect to the product topology. In particular, $\mathbb{R}_+^{\binom{n}{2}}$ and $\mathbb{R}_+^{\binom{\mathbb{N}}{2}}$ are Polish.*

Proof. First note, that the metric on X defined by

$$r(x, y) := \sum_{n\in\mathbb{N}} \frac{1}{2^n}(r_n(x_n, y_n) \wedge 1)$$

is complete and generates the product topology, where $\wedge$ means the minimum. Moreover, let $Y_n \subset X_n$ be a countable dense subset. Then the set

$$Y := \{y \in \prod_{n\in\mathbb{N}} Y_n : y_i \neq y_i' \text{ for finitely many } i \in \mathbb{N}\}$$

is countable and dense in X. $\square$

4.2 The space of compact metric spaces

We denote by $\mathbb{X}_c$ the *space of (isometry classes of) compact metric spaces*. See [BBI01] for more details. For $\mathcal{X} = (X, r_X), \mathcal{Y} = (Y, r_Y) \in \mathbb{X}_c$ the

Gromov-Hausdorff metric is given by

$$\mathrm{d_{GH}}(\mathcal{X},\mathcal{Y}) := \inf_{(\varphi_X,\varphi_Y,Z)} d_H^{(Z,r_Z)}(\varphi_X(X),\varphi_Y(Y)),$$

where the infimum is taken over all isometric embeddings φ_X and φ_Y from X and Y, respectively, into some common metric spaces (Z, r_Z).

The *Hausdorff metric* for closed subsets A, B of a metric space (Z, r_Z) is given by

$$d_H^{(Z,r_Z)}(A,B) := \inf\{\epsilon > 0 : A \subset B^\epsilon, B \subset A^\epsilon\},$$

where $A^\epsilon := \{z \in Z : r_Z(z,A) < \epsilon\}$ and $B^\epsilon := \{z \in Z : r_Z(z,B) < \epsilon\}$ is the ϵ-*neighborhood* of A and B, respectively.

Definition 4.3 (Correspondence, Distortion)**.**

(a) A *correspondence* between two metric spaces (X, r_X) and (Y, r_Y) is a subset $R \subset X \times Y$ such that for all $x \in X$, there is at least one $y \in Y$ with $(x, y) \in R$, and for all $y' \in Y$, there exists at least one $x' \in X$ with $(x', y') \in R$.

(b) The *distortion* dis(R) of a correspondence R is defined by

$$\mathrm{dis(R)} := \sup\{\mid r_X(x,x') - r_Y(y,y') \mid : (x,y),(x',y') \in R\}.$$

Remark 4.4 (Extension of metrics via relations). Let (X_1, r_{X_1}) and (X_2, r_{X_2}) be two metric spaces and $R \subset X_1 \times X_2$ a nonempty correspondence. Then there is a extension to a metric $r^R_{X_1 \sqcup X_2}$ on the disjoint union $X_1 \sqcup X_2$ defined by $r^R_{X_1 \sqcup X_2}(x_1, x_2) := r_{X_i}(x_1, x_2)$ if $x_1, x_2 \in X_i$, $i = 1, 2$, and

$$r^R_{X_1 \sqcup X_2}(x_1,x_2) := \inf\{r_{X_1}(x_1,x_1') + \frac{1}{2}dis(R) + r_{X_2}(x_2,x_2') : (x_1',x_2') \in R\}.$$

There is a statement analogous to Lemma 2.14, which concerns convergence in the Gromov-Prohorov metric.

Lemma 4.5. *Let* $\mathcal{X} = (X, r_X), \mathcal{X}_1 = (X_1, r_{X_1}), \mathcal{X}_2 = (X_2, r_{X_2}), \dots$ *be in* $\mathbb{X}_c$*. Then*

$$d_{GH}(\mathcal{X}_n,\mathcal{X}) \xrightarrow{n\to\infty} 0,$$

iff there is a compact metric space (Z, r_Z) *and isometric embeddings* $\varphi, \varphi_1, \varphi_2, \dots$ *of* $(X, r), (X_1, r_1), (X_2, r_2), \dots,$ *respectively, into* (Z, r_Z) *such that*

$$d_H^{(Z,r_Z)}(\varphi_n(X_n),\varphi(X)) \xrightarrow{n\to\infty} 0.$$

Proof. First note, that there is an equivalent formulation of the Gromov-Hausdorff metric based on correspondences. Indeed, by [BBI01, Theorem 7.3.25], we have

$$d_{GH}((X, r_X), (Y, r_Y)) = \frac{1}{2} \inf_R \text{dis(R)},$$

where the infimum is taken over all correspondences R between X and Y.

Therefore, if $d_{GH}(X_n, X) \xrightarrow{n\to\infty} 0$, we find correspondences R_n between X_n and X such that $\text{dis(R)} \xrightarrow{n\to\infty} 0$. We may define recursively metrics r_{Z_n} on $Z_n := \bigsqcup_{k=0}^n X_k$, where $\bigsqcup$ means the disjoint union, based on this correspondences. In the limit $n \to \infty$, we obtain a metric space, which completion is denoted by Z. Since $X, X_1, X_2, \dots$ are compact, we have for all $\epsilon > 0$, finite ϵ-nets of $X, X_1, X_2, \dots$, respectively. Then for well chosen ϵ_n, the union of this nets is a net for Z. Hence, Z is totally bounded, i.e. compact. The converse it clear by definition, [GPW09, A1]. □

Lemma 4.6. *Let $\mathcal{X}, \mathcal{Y} \in \mathbb{X}_c$ and for $\epsilon > 0$, $S_\mathcal{X} = (x_i)_{i\le N}$ and $S_\mathcal{Y} = (y_i)_{i\le N}$ be a ϵ-net for X and Y, respectively. If for all $i, j \in \{1, \dots, N\}$,*

$$| \, r_X(x_i, x_j) - r_Y(y_i, y_j) \, | < \epsilon,$$

then

$$d_{GH}(\mathcal{X}, \mathcal{Y}) < 3\epsilon.$$

Proof. The correspondence $\{(x_i, y_i) : i \le N\}$ between $S_\mathcal{X}$ and $S_\mathcal{Y}$ has distortion less then ϵ. Hence, we find as in the proof of the last lemma, that

$$d_{GH}(S_\mathcal{X}, S_\mathcal{Y}) < \frac{\epsilon}{2}.$$

Moreover, since $S_\mathcal{X}$ and $S_\mathcal{Y}$ are ϵ-nets of X and Y, respectively, we have

$$d_{GH}(\mathcal{X}, S_\mathcal{X}) \le \epsilon \text{ and } d_{GH}(S_\mathcal{Y}, \mathcal{Y}) \le \epsilon.$$

By the triangle inequality for d_{GH}, we are done, [BBI01, Proposition 7.4.11]. □

Proposition 4.7 (Criterion for relative compactness in $\mathbb{X}_c$)**.** *A set $\Gamma \subset \mathbb{X}_c$ is relatively compact if it is* uniformly totally bounded*, i.e.*

(a) there is $D > 0$ such that $\sup_{\mathcal{X}\in\Gamma} diam(X) \le D$,

(b) for all $\epsilon > 0$ there is $N_\epsilon \in \mathbb{N}$ such that every $\mathcal{X} \in \Gamma$ can be covered by at most N_ϵ balls of radius ϵ.

Proof. We define $N(1) := N_1$ and inductively for $k \geq 2$,

$$N(k) := N(k-1) + N_{1/k}.$$

Consider a sequence $(\mathcal{X}_n)_{n\geq 1}$ in Γ. For each $n \in \mathbb{N}$, we define the union S_n of $(1/k)$-nets in X_n, such that for every $k \in \mathbb{N}$, the first $N(k)$ points of S_n form a $(1/k)$-net in X_n. We get $S_n = (x_{i,n})_{i\in\mathbb{N}}$.

Clearly, we have by assumption, that $r_{X_n}(x_{i,n}, x_{j,n}) \leq D$, i.e. the distances belong to a compact interval. Therefore, we can use the Cantor diagonal procedure to get a subsequence of $(\mathcal{X}_n)_{n\geq 1}$ in which $r_{X_n}(x_{i,n}, x_{j,n})$ converges for all $i, j \in \mathbb{N}$ as $n \to \infty$.

We use a new countable space $\bar{X} := (x_i)_{i\in\mathbb{N}}$ to construct a limit object $\mathcal{X}$ for our subsequence, which we denote also by $(\mathcal{X}_n)_{n\geq 1}$. Firstly, we define a semi-metric $r_{\bar{X}}$ on $\bar{X}$ by

$$r_{\bar{X}}(x_i, x_j) := \lim_{n\to\infty} r_{X_n}(x_{i,n}, x_{j,n}).$$

We define an equivalence relation on $\bar{X}$ by $x_i \sim x_j$ iff $r_{\bar{X}}(x_i, x_j) = 0$. Then, we denote by X the completion of the quotient space $\bar{X}/\sim$ with respect to the induced metric $r_{\bar{X}/\sim}$.

Finally, we have to show, that $\mathcal{X}_n \xrightarrow{n\to\infty} X$. Consider for all $k \in \mathbb{N}$ the $(1/k)$-net in X given by $S^{(k)} = \{[x_i] : 1 \leq i \leq N(k)\}$. Indeed, each set $S_n^{(k)} = \{x_{i,n} : 1 \leq i \leq N(k)\}$ is a $(1/k)$-net in X_n, respectively. Hence, for all $x_{i,n} \in S_n$ there is $j \leq N(k)$, such that $r_{X_n}(x_{i,n}, x_{j,n}) \leq 1/k$. Note, that for every fixed $i \in \mathbb{N}$, $N(k)$ does not depend on n. Thus, there is $j_k \leq N(k)$, such that $r_{X_n}(x_{i,n}, x_{j_k,n}) \leq 1/k$ for infinitely many $n \in \mathbb{N}$. In addition, we find $r_X([x_i], [x_{j_k}]) \leq 1/k$.

To get convergence, note that, $S_n^{(k)} \xrightarrow{n\to\infty} S^{(k)}$ in the Gromov-Hausdorff topology. Indeed, by construction, the sets $S_n^{(k)}$ are finite and the distances are converging. We conclude by Lemma 4.6, since for all $k \in \mathbb{N}$ we find $n \in \mathbb{N}$, such that $d_{GH}(\mathcal{X}_n, \mathcal{X}) < 3/k$, [BBI01, Theorem 7.4.15]. □

4.3 Measure and probability theory

In this section we recall some well known definitions and facts about general measure and probability theory needed in various situations throughout this work. See [Kal02], [Dur05], [Bau02], [EK86] or [Els02], [Kle08]. We use a metric space (E, r) or a (background) measure space $(\Omega, \mathscr{F}, \mu)$. We start with the well known Chebyshev inequality.

Lemma 4.8 (Markov's inequality, Chebyshev's inequality)**.** *Let X be a random variable and $f : \mathbb{R}_+ \to \mathbb{R}_+$ increasing. For all $\epsilon > 0$ such that $f(\epsilon) > 0$ we have the Markov inequality, i.e.*

$$\mathbb{P}(\mid X \mid \geq \epsilon) \leq \frac{\mathbb{P}[f(\mid X \mid)}{f(\epsilon)}.$$

In particular, for a square-integrable random variable X and $f(x) = x^2$ we get Chebyshev's inequality .

Proof. We have

$$\begin{aligned}\mathbb{P}[f(\mid X \mid)] &\geq \mathbb{P}[f(\mid X \mid)\mathbb{1}_{f(|X|) \geq f(\epsilon)}] \geq \mathbb{P}[f(\epsilon)\mathbb{1}_{f(|X|) \geq f(\epsilon)}] \\ &\geq f(\epsilon)\mathbb{P}(\mid X \mid \geq \epsilon),\end{aligned}$$

[Kle08, Satz 5.11]. □

Proposition 4.9 (Monotone convergence theorem)**.** *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of measurable functions on $(\Omega, \mathscr{F}, \mu)$ with $0 \leq f_n \uparrow f$. Then we have*

$$\mu[f_n] \uparrow \mu[f].$$

Proof. For each $n \in \mathbb{N}$ we approximate f_n by some simple measurable functions $g_{n,k}$, i.e. $0 \leq g_{n,k} \uparrow f_n$. Then the functions $h_{n,k} := g_{1,k} \vee \cdots \vee g_{n,k}$, where $\vee$ denotes the maximum, have the same properties and are nondecreasing in both indices. Then it follows,

$$f \geq \lim_{k\to\infty} h_{k,k} \geq \lim_{k\to\infty} h_{n,k} = f_n \uparrow f.$$

Hence, $0 \leq h_{k,k} \uparrow f$. Furthermore,

$$\mu[f] = \lim_{k\to\infty} \mu[h_{k,k}] \leq \lim_{k\to\infty} \mu[f_k] \leq \mu[f],$$

[Kal02, Theorem 1.19]. □

Lemma 4.10 (Fatou)**.** *Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of nonnegative measurable functions on $(\Omega, \mathscr{F}, \mu)$. Then we have*

$$\mu[\liminf_{n\to\infty} f_n] \leq \liminf_{n\to\infty} \mu[f_n].$$

On weak convergence.

Definition 4.11 (Weak convergence)**.** Let $\mu, \mu_1, \mu_2, \ldots \in \mathscr{M}_1(E)$. We say, that the sequence $(\mu_n)_{n\in\mathbb{N}}$ *converges weakly* to μ if for all $f \in C_b(E)$,

$$\int f d\mu_n \xrightarrow{n\to\infty} \int f d\mu.$$

We denote weak convergence in $\mathscr{M}_1(E)$ by $\Rightarrow$.

The weak convergence induces the weak topology on $\mathscr{M}_1(E)$. If E is separable the weak topology is metrizable by, for example, the Prohorov metric, Definition 2.11.

The following theorem is used frequently.

Theorem 4.12 (Portmanteau Theorem)**.** *Let $\mu, \mu_1, \mu_2, \ldots \in \mathscr{M}_1(E)$. Then it is equivalent.*

(a) $\mu_n \xRightarrow{n\to\infty} \mu$.

(b) For all $U \subset E$ open we have $\mu(U) \leq \liminf_{n\to\infty} \mu_n(U)$.

(c) For all $A \subset E$ closed we have $\limsup_{n\to\infty} \mu_n(A) \leq \mu(A)$.

(d) For all $B \subset E$ measurable with $\mu(\partial B) = 0$, where $\partial B := \bar{B} \setminus \mathring{B}$, we have $\lim_{n\to\infty} \mu_n(B) = \mu(B)$.

Proof. Assume (a) and let $U \subset E$ be open. We consider a continuous function f such that $0 \leq f \leq \mathbb{1}_U$. Then we find $\mu_n[f] \leq \mu_n(U)$ and obtain (b) if we let $n \to \infty$ and then $f \uparrow \mathbb{1}_U$. We get the equivalence of (b) and (c) by taking complements. Now assume (b) and (c) and let $B \subset E$ be measurable. We have

$$\mu(\mathring{B}) \leq \liminf_{n\to\infty} \mu_n(B) \leq \limsup_{n\to\infty} \mu_n(B) \leq \mu(\bar{B}).$$

Hence, we find (d), if $\mu(\partial B) = 0$.

Next, assume (d) and fix $A \subset E$ closed. We consider the ϵ-neighborhood $A^\epsilon := \{x \in E : r(x, A) \leq \epsilon\}$. Then the sets $\partial A^\epsilon \subset \{x \in E : r(x, A) = \epsilon\}$ are disjoint and we find $\mu(\partial A^\epsilon) = 0$. Since $\mu_n(A) \leq \mu(A^\epsilon)$, it follows (c) as we let $n \to \infty$ and then $\epsilon \to 0$.

Finally, assume (b) and let $f \geq 0$ be a continuous function. By Fatou's Lemma, Lemma 4.10,

$$\mu[f] = \int_{\mathbb{R}_+} \mu(f > t)dt \leq \int_{\mathbb{R}_+} \liminf_{n\to\infty} \mu_n(f > t)dt$$
$$\leq \liminf_{n\to\infty} \int_{\mathbb{R}_+} \mu_n(f > t)dt = \liminf_{n\to\infty} \mu_n[f].$$

Now take $f \in C_b(E)$ with upper bound $c < \infty$. Applying the above to $c \pm f$ yields to (a), [Kal02, Theorem 4.25]. □

Lemma 4.13.

(a) The set $A := \{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,\epsilon)) \leq \delta\}$ *is closed in* $\mathscr{M}_1(\mathbb{R}_+)$.

(b) The set $\{\nu \in \mathscr{M}_1(\mathbb{R}_+) : \nu([0,\epsilon]) < \delta\}$ *is open in* $\mathscr{M}_1(\mathbb{R}_+)$.

Proof. For (a), let $(\nu_n)_{n\in\mathbb{N}}$ be a sequence in A which converges weakly to $\nu \in \mathscr{M}_1(\mathbb{R}_+)$ as $n \to \infty$. As $[0,\epsilon) \subset \mathbb{R}_+$ is open, by the Portmanteau Theorem, Theorem 4.12, we have

$$\nu([0,\epsilon)) \leq \liminf_{n\to\infty} \mu_n([0,\epsilon)) \leq \delta.$$

It follows (b) by similar argument. We omit the details. □

Proposition 4.14 (Continuous Mapping Theorem). *Let* (E_1, d_1) *and* (E_2, d_2) *be two metric spaces,* $\varphi : E_1 \to E_2$ *measurable and* U_φ *the set consisting of all points of discontinuity of* φ. *Let* $\mu, \mu_1, \mu_2, ... \in \mathscr{M}_1(E_1)$ *such that* $\mu(U_\varphi) = 0$. *If* $\mu_n \Rightarrow \mu$ *then* $\varphi_*\mu_n \Rightarrow \varphi_*\mu$ *as* $n \to \infty$.

On tightness and relative compactness in $\mathscr{M}_1(E)$.

Definition 4.15. A family $F \subset \mathscr{M}_1(E)$ is called *tight* if for all $\epsilon > 0$ there is a compact set $K \subset E$ such that

$$\inf_{\mu\in F} \mu(K) \geq 1 - \epsilon.$$

Proposition 4.16 (Tightness and relative compactness, Prohorov). *Let* E *be Polish and* $F \subset \mathscr{M}_1(E)$. *Then* F *is tight if and only if* F *is relatively compact in* $\mathscr{M}_1(E)$ *with respect to the weak topology.*

Proof. Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in F. By assumption, there are compact set $K_1 \subset K_2 \subset \ldots \subset E$ such that

$$\inf_{n\in\mathbb{N}} \mu_n(K_j) \geq 1 - \frac{1}{j}.$$

For a dense subset $\{x_1, x_2, \ldots\} \subset E$, we consider the countable semiring defined by

$$\mathscr{K} := \Big\{ \bigcup_{k=1}^{N} K_{j_k} \cap \bar{B}_{\epsilon_k}(x_k) : N, j_k \in \mathbb{N}, \epsilon_k \in \mathbb{Q} \Big\}.$$

Note, that $\mathscr{K}$ is stable under intersections and generates $\mathscr{B}(E)$. Since $\mathscr{K}$ is countable, we find a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ of $(\mu_n)_{n\in\mathbb{N}}$ such that for all $K \in \mathscr{K}$, $(\mu_{n_k}(K))_{k\in\mathbb{N}}$ is converging. We define our desired limit object on the generator $\mathscr{K}$ by

$$\mu(K) := \lim_{k\to\infty} \mu_{n_k}(K).$$

Then, by Proposition 4.30, we obtain a measure on $\mathscr{B}(E)$. Indeed, μ is a probability measure, since

$$1 \geq \mu(E) = \sup_{j\in\mathbb{N}} \mu(K_j) = \sup_{j\in\mathbb{N}} \lim_{k\to\infty} \mu_{n_k}(K_j) \geq \sup_{j\in\mathbb{N}} 1 - \frac{1}{j} = 1.$$

Moreover, for all $U \subset E$ open, we have

$$\mu(U) = \sup_{K\in\mathscr{K}, K\subset U} \mu(K) = \sup_{K\in\mathscr{K}, K\subset U} \lim_{k\to\infty} \mu_{n_k}(K) \leq \liminf_{k\to\infty} \mu_{n_k}(U).$$

By the Portmanteau Theorem, Theorem 4.12, it follows $\mu_{n_k} \xRightarrow{k\to\infty} \mu$.

Conversely, let again $\{x_1, x_2, \ldots\} \subset E$ be dense. If we define

$$A_{n,N} := \bigcup_{i=1}^{N} B_{1/n}(x_i),$$

then we have for all $n \in \mathbb{N}$, that $A_{n,N} \uparrow E$ as $N \to \infty$. Now let

$$\delta := \sup_{n\in\mathbb{N}} \inf_{N\in\mathbb{N}} \sup_{\mu\in F} \mu(A_{n,N}^c).$$

Then, there is $n \in \mathbb{N}$, such that for all $N \in \mathbb{N}$ there exists $\mu_N \in F$ with $\mu_N(A_{n,N}^c) \geq \delta/2$. Since F is relatively compact, there is a subsequence

$(\mu_{N_k})_{k\geq 1}$ of $(\mu_N)_{N\geq 1}$ which converges to $\mu \in \mathscr{M}_{\leq 1}(E)$. By the Portmanteau Theorem, Theorem 4.12, we find for all $N \in \mathbb{N}$,

$$\mu(A_{n,N}^c) \leq \liminf_{k\to\infty} \mu_{N_k}(A_{n,N}^c) \leq \liminf_{k\to\infty} \mu_{N_k}(A_{n,N_k}^c) \leq \frac{\delta}{2}.$$

Further, $A_{n,N}^c \downarrow \emptyset$ as $N \to \infty$. Hence, $\mu(A_{n,N}^c) \xrightarrow{N\to\infty} 0$.

Let $\epsilon > 0$. By the above, for all $n \in \mathbb{N}$, there is $N_n' \in \mathbb{N}$, such that for all $\mu \in F$,

$$\mu(A_{n,N_n'}^c) < \epsilon/2^n.$$

The set defined by

$$A := \bigcap_{n=1}^{\infty} A_{n,N_n'}$$

is totally bounded by construction. Hence, relatively compact. Finally, we find for all $\mu \in F$,

$$\mu(\bar{A}^c) \leq \mu(A^c) \leq \sum_{n=1}^{\infty} \mu(A_{n,N_n'}) \leq \epsilon,$$

i.e. F is tight, [Kle08, Satz 13.29]. □

On separating classes of functions.

Definition 4.17 (Algebra, separating classes of functions)**.**

(a) A set $\mathscr{A} \subset \mathcal{C}(E)$ is called an *algebra* if $\mathscr{A}$ is a vector space and there is a bilinear map $\mathscr{A} \times \mathscr{A} \to \mathscr{A}$ which is associative and distributive.

(b) A set $\mathscr{A} \subset \mathcal{C}_b(E)$ *separates points* in E, if for all $x.y \in E$ with $x \neq y$ there is $f \in \mathscr{A}$ such that $f(x) \neq f(y)$.

(c) A set $\mathscr{A} \subset \mathcal{C}_b(E)$ is *separating* in $\mathcal{M}_1(E)$ if for $\mu, \nu \in \mathcal{M}_1(E)$ with $\mu[f] = \nu[f]$ for all $f \in \mathscr{A}$ it follows that $\mu = \nu$.

Lemma 4.18. *Let $\mathscr{A} := \mathcal{C}_b(E)$. Then $\mathscr{A}$ separates points and is separating in $\mathcal{M}_1(E)$.*

Proof. Let $x, y \in E$ with $x \neq y$. Then the function defined by

$$z \mapsto (r(x,z) \wedge 1)$$

is continuous and bounded. Here $\wedge$ means the minimum. Moreover, the function separates x and y. Furthermore, let $\mu, \nu \in \mathcal{M}_1(E)$ with $\mu \neq \nu$. Then there is an open ball $B \subset E$ such that $\mu(B) \neq \nu(B)$. Now, we consider a sequence $(f_n)_{n \in \mathbb{N}}$ in $\mathcal{C}_b(E)$ such that $f_n \uparrow \mathbb{1}_B$. We argue by contradiction. Hence, assume that for all $n \in \mathbb{N}$, $\mu[f_n] = \nu[f_n]$. Then

$$\mu(B) = \lim_{n\to\infty} \mu[f_n] = \lim_{n\to\infty} \nu[f_n] = \nu(B).$$

□

Proposition 4.19 (Stone-Weierstrass). *Let E be compact and $\mathscr{A} \subset \mathcal{C}_b(E)$ an algebra which separates points. Then $\mathscr{A}$ is dense in $\mathcal{C}_b(E)$ with respect to the uniform norm $\| f \| := \sup_{x \in E} f(x)$.*

Proposition 4.20. *Let (E, r) be complete and separable and $\mathscr{A}$ an algebra which separates points in E. Then $\mathscr{A}$ is separating in $\mathcal{M}_1(E)$.*

Proof. Take $\mu, \nu \in \mathcal{M}_1(E)$ such that for all $f \in \mathscr{A}$, $\mu[f] = \nu[f]$. Let $g \in \mathcal{C}_b(E)$ and $\epsilon > 0$. Note, that there is a compact set $K = K_\epsilon \subset E$ such that $\mu(K) > 1-\epsilon$ and $\nu(K) > 1-\epsilon$. Then by Stone-Weierstrass, Proposition 4.19, there exists $g_\epsilon \in \mathscr{A}$ such that

$$\sup_{x \in K} \mid g(x) - g_\epsilon(x) \mid < \epsilon.$$

Then by the triangle inequality, it follows

$$\begin{aligned}
\mid \mu[ge^{-\epsilon g^2}] - \nu[ge^{-\epsilon g^2}] \mid \leq & \mid \mu[ge^{-\epsilon g^2}] - \mu[ge^{-\epsilon g^2}, K] \mid \\
& + \mid \mu[ge^{-\epsilon g^2}, K] - \mu[g_\epsilon e^{-\epsilon g_\epsilon^2}, K] \mid \\
& + \mid \mu[g_\epsilon e^{-\epsilon g_\epsilon^2}, K] - \mu[g_\epsilon e^{-\epsilon g_\epsilon^2}] \mid \\
& + \mid \mu[g_\epsilon e^{-\epsilon g_\epsilon^2}] - \nu[g_\epsilon e^{-\epsilon g_\epsilon^2}] \mid \\
& + \mid \nu[g_\epsilon e^{-\epsilon g_\epsilon^2}] - \nu[g_\epsilon e^{-\epsilon g_\epsilon^2}, K] \mid \\
& + \mid \nu[g_\epsilon e^{-\epsilon g_\epsilon^2}, K] - \nu[ge^{-\epsilon g^2}, K] \mid \\
& + \mid \nu[ge^{-\epsilon g^2}, K] - \nu[ge^{-\epsilon g^2}] \mid .
\end{aligned}$$

The four terms of the form $|\mu[ge^{-\epsilon g^2}] - \mu[ge^{-\epsilon g^2}, K]|$ are bounded above by

$$|\mu[ge^{-\epsilon g^2}] - \mu[ge^{-\epsilon g^2}, K]| \leq \sup_{x\geq 0} xe^{-\epsilon x}\frac{1}{\sqrt{\epsilon}}\mu(K^c) \leq \sup_{x\geq 0} xe^{-\epsilon x}\sqrt{\epsilon},$$

since all mass expect of ϵ is put on K.

The two terms of the form $|\mu[ge^{-\epsilon g^2}, K] - \mu[g_\epsilon e^{-\epsilon g_\epsilon^2}, K]|$ converge to 0, since g_ϵ approximates g on K. Moreover, the term $|\mu[g_\epsilon e^{-\epsilon g_\epsilon^2}] - \nu[g_\epsilon e^{-\epsilon g_\epsilon^2}]|$ is 0 by assumption. Finally, considering the limit $\epsilon \to 0$, we find

$$\mid \mu[g] - \nu[g] \mid = \lim_{\epsilon\to 0} \mid \mu[ge^{-\epsilon g^2}] - \nu[ge^{-\epsilon g^2}] \mid = 0.$$

By Lemma 4.18, $\mathcal{C}_b(E)$ is separating in $\mathcal{M}_1(E)$ and we are done, [EK86, Theorem 4.5]. □

On semi-continuity. In Section 2.2, Lemma 2.30, we consider a lower semi-continuous function ξ_ϵ. For a function $f : E \to \mathbb{R}$ to be continuous at a point $a \in E$ it is necessary and sufficient that given $h < f(a)$ there is a neighborhood V of a such that for each $b \in V$, $h < f(b)$ and given $k > f(a)$ there is a neighborhood W of a such that for each $b \in V$, $k > f(b)$.

Definition 4.21 (Semi-continuity)**.**

(a) A function $f : E \to \mathbb{R}$ is called *lower semi-continuous* at a point $a \in E$, if for each $h < f(a)$ there is a neighborhood V of a such that for all $b \in V$, $h < f(b)$. A function f is called lower semi-continuous if it is lower semi-continuous at each $a \in E$.

(b) A function $f : E \to \mathbb{R}$ is called *upper semi-continuous* at a point $a \in E$, if for each $k > f(a)$ there is a neighborhood W of a such that for all $b \in W$, $k > f(b)$. A function f is called upper semi-continuous if it is upper semi-continuous at each $a \in E$.

Lemma 4.22 (Characterization of lower semi-continuous functions)**.** *Let $f : E \to \mathbb{R}$. Then it is equivalent.*

(a) The map f is lower semi-continuous.

(b) For all $h \in \mathbb{R}$, the set $f^{-1}((-\infty, h])$ is closed.

(c) For all $a \in E$, $\liminf_{x\to a} f(x) \geq f(a)$.

Proof. For the equivalence of (a) and (b) note that by definition, if f is lower semi-continuous at a, for each $h < f(a)$, the set $f^{-1}((h, \infty))$ must be a neighborhood of a.

$(a) \Rightarrow (c)$. Given any $h < f(a)$, there is a neighborhood V of a such that for all $b \in V$, $h < f(b)$. Therefore, $h \leq \inf_{x \in V} f(x) \leq \liminf_{x \to a} f(x)$. Hence, $f(a) \leq \liminf_{x \to a} f(x)$.

$(c) \Rightarrow (a)$. For each $h < f(a)$, there is a neighborhood V of a such that $h \leq \inf_{x \in V} f(x)$. Hence, f is lower semi-continuous at a, [Bou66, IV 6.2, Proposition 1, 3]. □

Lemma 4.23. *Let $f : E \to \mathbb{R}$ be lower semi-continuous. Then f is measurable with respect to the Borel-σ-algebra $\mathscr{B}(E)$.*

Proof. Note that sets of the form $(-\infty, y]$ generate the Borel-σ-algebra $\mathscr{B}(\mathbb{R})$. Since $\{x \in E : f(x) \leq y\} = f^{-1}((-\infty, y])$, f is measurable if the set $\{x \in E : f(x) \leq y\}$ is closed. We conclude by Lemma 4.22. □

On exchangeability. An important property of a Λ-coalescent process is the exchangeability for all times. Therefore, we give here the most important results on exchangeable random variable. Moreover, exchangeable random parons are discussed is Section 3.1.

Definition 4.24 (Exchangeable random variables)**.** A sequence of random variables $(X_n)_{n \in \mathbb{N}}$ is called *exchangeable* if for each n and permutation σ of $[n]$

$$(X_1, \ldots, X_n) \stackrel{d}{=} (X_{\sigma(1)}, \ldots, X_{\sigma(n)}),$$

where $\stackrel{d}{=}$ means equality in distribution.

Definition 4.25 (Exchangeable σ-field)**.** Let $\mathcal{E}_n$ be the σ-field generated by events that are invariant under permutations that leaves $n+1, n+2, \ldots$ fix. Then, $\mathcal{E} := \bigcap_{n \in \mathbb{N}} \mathcal{E}_n$ is called the *exchangeable σ-field.*

Theorem 4.26 (de Finetti)**.** *If the sequence $(X_n)_{n \in \mathbb{N}}$ is exchangeable then conditional on the exchangeable σ-field $\mathcal{E}$, $X_1, X_2, \ldots$ are independent and identically distributed.*

Proof. We define

$$A_n(\varphi) := \frac{1}{(n)_k} \sum_i \varphi(X_{i_1}, X_{i_2}, \dots, X_{i_k}),$$

where the sum is over all sequences of distinct integers $1 \leq i_1, \dots, i_k \leq n$ and

$$(n)_k := n(n-1)\cdots(n-k+1)$$

is the number of such sequences. Since the sequence $(X_n)_{n\in\mathbb{N}}$ is exchangeable, all terms in the following sum are equal. Indeed, we find

$$\begin{aligned} A_n(\varphi) = \mathbb{P}[A_n(\varphi) \mid \mathcal{E}_n] &= \frac{1}{(n)_k} \sum_i \mathbb{P}[\varphi(X_{i_1}, \dots, X_{i_k}) \mid \mathcal{E}_n] \\ &= \mathbb{P}[\varphi(X_{i_1}, \dots, X_{i_k}) \mid \mathcal{E}_n]. \end{aligned}$$

In Addition, $\mathcal{E}_n \downarrow \mathcal{E}$. Hence,

$$A_n(\varphi) = \mathbb{P}[\varphi(X_{i_1}, \dots, X_{i_k}) \mid \mathcal{E}_n] \xrightarrow{n\to\infty} \mathbb{P}[\varphi(X_{i_1}, \dots, X_{i_k}) \mid \mathcal{E}].$$

Now, we consider two bounded functions f and g on $\mathbb{R}^{k-1}$ and $\mathbb{R}$, respectively. Then

$$\begin{aligned} (n)_{k-1} A_n(f) \cdot n A_n(g) &= \sum_i f(X_{i_1}, \dots, X_{i_{k-1}}) \sum_m g(X_m) \\ = \sum_i f(X_{i_1}, \dots, X_{i_{k-1}}) g(X_{i_k}) &+ \sum_i \sum_{j=1}^{k-1} f(X_{i_1}, \dots, X_{i_{k-1}}) g(X_{i_k}) \end{aligned}$$

Therefore, if we let

$$\begin{aligned} \varphi(x_1, \dots, x_k) &= f(x_1, \dots, x_{k-1}) g(x_k), \\ \varphi_j(x_1, \dots, x_k) &= f(x_1, \dots, x_{k-1}) g(x_j), \end{aligned}$$

we find

$$A_n(\varphi) = \frac{n}{n-k+1} A_n(f) A_n(g) - \frac{1}{n-k+1} \sum_{j=1}^{k-1} A_n(\varphi_j).$$

Then, applying the above to all functions φ, f, g and φ_j, we find

$$\mathbb{P}[f(X_1, \dots, X_{k-1}) g(X_k) \mid \mathcal{E}] = \mathbb{P}[f(X_1, \dots, X_{k-1}) \mid \mathcal{E}] \; \mathbb{P}[g(X_k) \mid \mathcal{E}].$$

We conclude inductively and find

$$\mathbb{P}\Big[\prod_{j=1}^{k} f_j(X_j) \mid \mathcal{E}\Big] = \prod_{j=1}^{k} \mathbb{P}\Big[f_j(X_j) \mid \mathcal{E}\Big],$$

[Dur05, (6.6)]. □

Theorem 4.27. *Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of exchangeable random variables taking only the values 0 and 1. Then there is a nonnegative finite measure Λ on $\mathscr{B}([0,1])$ such that*

$$\mathbb{P}(X_1 = 1, \dots, X_k = 1, X_{k+1} = 0, \dots, X_n = 0) = \int_0^1 x^k (1-x)^{n-k} \Lambda(dx).$$

Proof. Define $p_{k,n} := \mathbb{P}(X_1 = 1, \dots, X_k = 1, X_{k+1} = 0, \dots, X_n = 0)$. Then we find

$$p_{n-1,n} = p_{n-1,n-1} - p_{n,n},$$

and recursively

$$p_{n-2,n} = p_{n-2,n-1} - p_{n-1,n},$$
$$p_{k,n} = p_{k,n-1} - p_{k-1,n}.$$

Finally, the sequence $(p_{n,n})_{n\in\mathbb{N}}$ is nonnegative with $p_{0,0} = 1$. Therefore, there is a nonnegative finite measure Λ with moment sequence $(p_{n,n})$, [Fel71, VII.4]. □

Theorem 4.28 (Glivenko-Cantelli)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be independent and identically distributed with distribution F and let $F_n(x) := \frac{1}{n}\sum_{m=1}^{n} \mathbb{1}_{X_m \le x}$. Then*

$$\sup_x \mid F_n(x) - F(x) \mid \xrightarrow{n\to\infty} 0 \ a.s.$$

Proof. We fix $x \in \mathbb{R}$. Since $(X_n)_{n\in\mathbb{N}}$ is i.i.d. the same is true for $(\mathbb{1}_{X_n<x})_{n\in\mathbb{N}}$. Moreover, $\mathbb{P}[\mathbb{1}_{X_n<x}] = \mathbb{P}(X_n < x) = F(x-)$. By the strong law of large numbers we have

$$F_n(x-) = \frac{1}{n}\sum_{m=1}^{n} \mathbb{1}_{X_m<x} \xrightarrow{n\to\infty} F(x-) \text{ a.s.}$$

For $1 \leq j \leq k-1$ we define $x_{j,k} := \inf\{y : F(y) \geq j/k\}$. Note that $(F_n(x))_{n\in\mathbb{N}}$ and $(F_n(x-))_{n\in\mathbb{N}}$ are converging pointwise. Therefore, we find $N_k(\omega) \in \mathbb{N}$ such that for all $n \geq N_k(\omega)$,

$$\left.\begin{array}{l} |\ F_n(x_{j,k}) - F(x_{j,k})\ | \\ |\ F_n(x_{j,k}-) - F(x_{j,k}-)\ | \end{array}\right\} < \frac{1}{k}.$$

The same is true for $j = 0, k$ if we define $x_{0,k} := -\infty$ and $x_{k,k} := \infty$. For $1 \leq j \leq k$ and $n \geq N_k(\omega)$ let $x \in (x_{j-1,k}, x_{j,k})$. By the monotonicity of F_n and F and since $F(x_{j,k}-) - F(x_{j-1,k}) \leq 1/k$, we have

$$F_n(x) \leq F_n(x_{j,k}-) \leq F(x_{j,k}-) + 1/k \leq F(x_{j-1,k}) + 2/k \leq F(x) + 2/k,$$
$$F_n(x) \geq F_n(x_{j-1,k}) \geq F(x_{j-1,k}) - 1/k \geq F(x_{j,k}-) - 2/k \geq F(x) - 2/k,$$

i.e.

$$\sup_x |\ F_n(x) - F(x)\ | \leq \frac{2}{k},$$

[Dur05, (7.4)]. □

On the existence of stochastic processes.

Definition 4.29 (Consistent measures)**.** A family $\{\mathbb{P}_J : J \subset I \text{ finite }\}$ of measures is called *consistent*, if $\mathbb{P}_J$ is a probability measure on the product σ-field $\mathscr{F}^J$ and if for all $H \subset J$,

$$\mathbb{P}_H = (\vartheta^J_H)_* \mathbb{P}_J,$$

where $\vartheta^J_H : \Omega^J \to \Omega^H$ is the projection operator.

Proposition 4.30 (Carathéodory 1914)**.** *Let μ be a finitely additive and countably subadditive set function on a semiring $\mathscr{H}$ such that $\mu(\emptyset) = 0$. Then μ extends to a measure on the σ-algebra generated by $\mathscr{H}$.*

Proof. See for example [Kal02, Theorem 2.5]. □

Theorem 4.31 (Kolmogorov consistency theorem)**.** *Let (E, r) be Polish and $\{\mathbb{P}_J : J \subset I \text{ finite }\}$ a consistent family with respect to the Borel-σ-algebra on E. Then there is a unique probability measure $\mathbb{P}$ on $\mathscr{B}(E)^I$ such that $\mathbb{P}_J$ is the marginal distribution for $J \subset I$ finite. Moreover, there exists a probability space and a stochastic process with state space E and distribution $\mathbb{P}$.*

Proof. We define $\mathscr{C}_J := \vartheta_J^{-1}(\mathscr{B}(E)^J)$ the σ-field of $J-$cylinder and

$$\mathscr{C} := \bigcup_{J \subset I \text{ finite}} \mathscr{C}_J.$$

Hence, $\mathscr{C}$ generates $\mathscr{B}(E)^I$. Therefore, if we assume there is such a measure $\mathbb{P}$, since for $C = \vartheta_J^{-1}(B)$ where $B \in \mathscr{B}(E)^J$ we have $\mathbb{P}(C) = \mathbb{P}_J(B)$, the measure $\mathbb{P}$ is unique.

For existence of $\mathbb{P}$ we use Proposition 4.30. We define

$$\mathbb{P}_I(C) := \mathbb{P}_J(B),$$

where $C = \vartheta_J^{-1}(B)$ and $B \in \mathscr{B}(E)^J$. By the consistency of the family of measures, the set function $\mathbb{P}_I$ is well defined. Then $\mathbb{P}_I$ is finitely additive and we have $\mathbb{P}_I(\emptyset) = 0$. To get countable subadditivity, it can be shown, that $\mathbb{P}_I$ is regular (or alternatively continuous from above). We leave out the long calculation. Finally, if we let $\Omega := E^I$, $\mathscr{F} := \mathscr{B}(E)^I$, $\mathbb{P} := \mathbb{P}_I$ and $X : \Omega \to E$, $\omega = (\omega_t)_{t \in I} \mapsto \omega_t$, we find the desired process, [Bau02, 35.3 Satz] or [Dur05, (7,1)]. $\square$

Definition 4.32 (Projective limit)**.** The probability measure $\mathbb{P}$ in the above theorem is called the projective limit of the family $\{\mathbb{P}_J : J \subset I \text{ finite }\}$. We write

$$\mathbb{P}_I = \varprojlim_{J \subset I \text{ finite}} \mathbb{P}_J.$$

APPENDIX A

Compact metric measure spaces and Λ-coalescents coming down from infinity

The following includes the essential results of the above theses. This paper is a joint work with Peter Pfaffelhuber and is submitted to the Latin American Journal of Probability and Mathematical Statistics, ALEA (2011).

Abstract. We study topological properties of random metric spaces which arise by Λ-coalescents. These are stochastic processes, which start with an infinite number of lines and evolve through multiple mergers in an exchangeable setting. We show that the resulting Λ-coalescent measure tree is compact iff the Λ-coalescent comes down from infinity, i.e. only consists of finitely many lines at any positive time. If the Λ-coalescent stays infinite, the resulting metric measure space is not even locally compact.

Our results are based on general notions of compact and locally compact (isometry classes of) metric measure spaces. In particular, we give characterizations for general (random) metric measure spaces to be (locally) compact using the Gromov-weak topology.

A.1 Introduction

Metric structures arise frequently in probability theory. Prominent examples are random trees (e.g. [Ald85] [EO94], [Gal99], [Ber09]) where the distance between two points is given by the length of the shortest path connecting the points. A class of random trees is given by coalescent processes, where a subset of an infinite number of *lines* can merge and the distance of two leaves is proportional to the coalescence time ([Kin82a], [Pit99], [Ald99], [Sch00b], [Eva00]). The complexity of this class of processes is properly described by the concepts of Λ-coalescent, where any set of lines can merge to a single line (a multiple collision, [Pit99]) and Ξ-coalescents, where any set of lines can merge to several lines at the same time (a simultaneous multiple collision, [Sch00a]). The resulting metric space has so far mostly been studied in the simplest case, where only binary mergers are allowed, the Kingman-coalescent ([Kin00], [EO94]).

Analyzing metric structures requires geometrical and topological foundations. In the context of Riemannian geometry, such foundations have already been laid by Gromov, summarized in his book ([Gro99], see also [Ver00], [BBI01]). These authors study convergence of (isometry classes of) compact metric spaces by the notion of Gromov-Hausdorff convergence. In addition, Gromov introduced a topology on the space of (isometry classes of) metric measure spaces (mm-spaces, for short), which are metric spaces equipped with a measure. We will call this the Gromov-weak topology in the sequel (see also [GPW09]).

In probability theory, results on weak convergence and stochastic process theory require that the underlying space is Polish. In addition, a characterization of the compact sets is required in order to show tightness. These concepts have been worked out based on Gromov's notions by [EPW06] and [GPW09].

The goal of the present paper is as follows: we concentrate on the spaces of *locally compact* and *compact* mm-spaces and give a characterization of these (see Theorems A.11 and A.16). In addition, we apply these general results to random mm-spaces (Λ-coalescent measure trees) which arise in connection to Λ-coalescents. Recall that Λ-coalescents fall into one of two categories, depending on Λ. Either a Λ-coalescent comes down from infinity, meaning that it can be started with an infinite number of lines and only finitely many are left at any positive time, or it stays infinite for all times (see [Pit99, Proposition 23]). The proof of the following result is given in Section A.3.

Theorem A.1 (Coming down from infinity and compactness)**.** *Let Λ be a finite measure on $[0,1]$ and $(\Pi_t)_{t\geq 0}$ the corresponding Λ-coalescent. Moreover, $\mathcal{L}$ is the associated Λ-coalescent measure tree, taking values in the space of mm-spaces. Then the following is equivalent.*

(a) $(\Pi_t)_{t\geq 0}$ comes down from infinity, i.e. $\#\Pi_t < \infty$ almost surely, for all $t > 0$.

(b) $\mathcal{L}$ is compact, almost surely.

If (a) (or (b)) does not hold, $\mathcal{L}$ is not even locally compact.

We proceed as follows: In Section A.2 we develop our general theory on compact and locally compact isometry classes of metric measure spaces. Section A.3 contains a short introduction to Λ-coalescent measure trees. Finally, the proof of Theorem A.1 is given in Section A.3. We remark that the application of (locally) compact mm-spaces is not restricted to trees. For example, it is possible to study large random planar maps, as given in [Gal07], or random Graphs (e.g. the Erdős-Renyi random graph, [ABBG10]), by our notions.

A.2 Metric measure spaces

We start with some notation. Our main results, the characterization of compact and locally compact mm-spaces, is given in Theorems A.11 and A.16.

Remark A.2 (Notation). As usual, given a topological space $(X, \mathscr{O}_X)$, we denote by $\mathscr{M}_1(X)$ the space of all probability measures on the Borel-σ-algebra $\mathscr{B}(X)$. The *support* of $\mu \in \mathscr{M}_1(X)$, $\text{supp}(\mu)$, is the smallest closed set $X_0 \subseteq X$ such that $\mu(X \setminus X_0) = 0$. The *push-forward* of μ under a measurable map φ from X into another topological space, $(Z, \mathscr{O}_Z)$, is the probability measure $\varphi_*\mu \in \mathscr{M}_1(Z)$ defined for all $A \in \mathscr{B}(Z)$ by $\varphi_*\mu(A) := \mu(\varphi^{-1}(A))$. We denote weak convergence in $\mathscr{M}_1(X)$ by $\Rightarrow$.

Definition A.3 (Metric measure and mm-spaces)**.**

(a) A *metric measure space* is a triple (X, r, μ) such that $X \subseteq \mathbb{R}$ and (X, r) is a complete and separable metric space which is equipped with a probability measure μ on $\mathscr{B}(X)$. We say that (X, r, μ) and (X', r', μ') are *measure-preserving isometric* if there exists an isometry φ between $\text{supp}(\mu) \subseteq X$ and $\text{supp}(\mu') \subseteq X'$ such that $\mu'|_{\text{supp}(\mu')} = \varphi_*(\mu|_{\text{supp}(\mu)})$. It is clear that the property of being measure-preserving isometric is an equivalence relation.

(b) The equivalence class of the metric measure space (X, r, μ) is called the mm-space of (X, r, μ) and is denoted $\overline{(X, r, \mu)}$. The set of mm-spaces is denoted $\mathbb{M}$ and generic elements are $\mathcal{X}, \mathcal{Y}, \ldots$

(c) An mm-space $\mathcal{X} \in \mathbb{M}$ is *(locally) compact* if there is $(X, r, \mu) \in \mathcal{X}$ such that (X, r) is (locally) compact. The space of (locally) compact mm-spaces is denoted $\mathbb{M}_c$ ($\mathbb{M}_{lc}$).

Following [GPW09], we equip $\mathbb{M}$ with the Gromov-weak topology as follows.

Definition A.4 (Gromov-weak topology)**.** For a metric space (X, r) define

$$R^{(X,r)} : \begin{cases} X^{\mathbb{N}} & \to \mathbb{R}_+^{\binom{\mathbb{N}}{2}} \\ (x_i)_{i\in\mathbb{N}} & \mapsto (r(x_i, x_j))_{1\le i<j} \end{cases}$$

the map which sends a sequence of points in X to its distance matrix and for an mm-space $\mathcal{X} = \overline{(X, r, \mu)}$ we define the *distance matrix distribution* by

$$\nu^{\mathcal{X}} := (R^{(X,r)})_*\mu^{\otimes\mathbb{N}} \in \mathscr{M}_1(\mathbb{R}_+^{\binom{\mathbb{N}}{2}}),$$

where $\mu^{\otimes\mathbb{N}}$ is the infinite product measure of μ, where $\mathbb{R}_+^{\binom{\mathbb{N}}{2}}$ is equipped with the product σ-field. We say that a sequence $\mathcal{X}_1, \mathcal{X}_2, \ldots \in \mathbb{M}$ converges *Gromov-weakly* to $\mathcal{X} \in \mathbb{M}$ if

$$\nu^{\mathcal{X}_n} \xRightarrow{n\to\infty} \nu^{\mathcal{X}}.$$

Note that $\nu^{\mathcal{X}}$ does not depend on the representative $(X, r, \mu) \in \mathcal{X}$, hence is well-defined.

Remark A.5 (When is a random mm-space compact?). Recall from Theorem 1 of [GPW09] that the space $\mathbb{M}$, equipped with the Gromov-weak topology, is Polish. Hence, $\mathbb{M}$ allows to use standard tools from probability, e.g. from the theory of weak convergence.

In order to show that a random variable taking values in $\mathbb{M}$ is supported by the space of locally compact or compact mm-spaces, there are two strategies, formulated here in the case of compact mm-spaces:

Either, consider the Gromov-weak topology on $\mathbb{M}_c$. Defining an approximating sequence in $\mathbb{M}_c$ and showing that the sequence is tight in $\mathbb{M}_c$ ensures compactness of the limiting object. Note that any mm-space can be approximated by finite (hence compact) mm-spaces, so $\mathbb{M}_c$ is not closed in $\mathbb{M}$. So, this approach amounts to knowing the compact sets in $\mathbb{M}_c$. See Proposition 6.2 of [GPW10] for an example.

Our application to the Λ-coalescent measure tree in Section A.3 relies on a different approach. It is possible to give handy characterizations of compact mm-spaces; see Theorem A.11. Hence, if we are given a random variable taking values in $\mathbb{M}$ through a sequence of mm-spaces, it is possible to check directly if the limiting object is compact.

Definition A.6 (Distance distribution, Moduli of mass distribution)**.** Let $\mathcal{X} \in \mathbb{M}$. We set $\underline{\underline{r}} := (r_{ij})_{1\le i<j} \in \mathbb{R}_+^{\binom{\mathbb{N}}{2}}$.

(a) Let $r : \mathbb{R}_+^{\binom{\mathbb{N}}{2}} \to \mathbb{R}_+$ be given by $r(\underline{\underline{r}}) := r_{12}$. Then, the *distance distribution* is given by $w_{\mathcal{X}} := r_*\nu^{\mathcal{X}}$, i.e.

$$w_{\mathcal{X}}(\cdot) := \nu^{\mathcal{X}}\big\{\underline{\underline{r}} : r_{12} \in \cdot\big\}.$$

(b) For $\varepsilon > 0$, define $s_\varepsilon : \mathbb{R}_+^{\binom{\mathbb{N}}{2}} \to \mathbb{R}_+$ by

$$s_\varepsilon\big(\underline{\underline{r}}\big) := \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} 1_{\{r_{1j}\le\varepsilon\}}$$

if the limit exists (and zero otherwise). Note that $s_\varepsilon(\underline{\underline{r}})$ exists for $\nu^{\mathcal{X}}$-almost all $\underline{\underline{r}}$ by exchangeability and de Finetti's Theorem. For $\delta > 0$, the *moduli of mass distribution* are

$$v_\delta(\mathcal{X}) := \inf\{\varepsilon > 0 : \ \nu^{\mathcal{X}}\big\{\underline{\underline{r}} : s_\varepsilon(\underline{\underline{r}}) \le \delta\big\} \le \varepsilon\}$$

and

$$\widetilde{v}_\delta(\mathcal{X}) := \inf\{\varepsilon > 0 : \ \nu^{\mathcal{X}}\{\underline{\underline{r}} : s_\varepsilon(\underline{\underline{r}}) \leq \delta\} = 0\}.$$

Example A.7 (Representatives of $\mathcal{X}$). Let $\mathcal{X} = \overline{(X, r, \mu)}$. Without loss of generality we assume that $\text{supp}(\mu) = X$. Since $\nu^{\mathcal{X}} = (R^{(X,r)})_*\mu^{\otimes\mathbb{N}}$, we have that

$$w_{\mathcal{X}}(\cdot) = \mu^{\otimes 2}\{(x, y) : r(x, y) \in \cdot\}.$$

Moreover,

$$\nu^{\mathcal{X}}\{\underline{\underline{r}} : s_\varepsilon(\underline{\underline{r}}) \in \cdot\} = \mu\{x : \mu(B_\varepsilon(x)) \in \cdot\} \tag{A.1}$$

by construction, where $B_\varepsilon(x)$ is the closed ball of radius ε around x. This implies that

$$v_\delta(\mathcal{X}) \leq \varepsilon \qquad \Longleftrightarrow \qquad \mu\{x : \mu(B_\varepsilon(x)) \leq \delta\} \leq \varepsilon.$$

In particular, $v_\delta(\mathcal{X}) \leq \varepsilon$ means, that thin points (in the sense that $\mu(B_\varepsilon(x)) \leq \delta$) are rare (i.e. carry mass at most ε). Moreover,

$$\widetilde{v}_\delta(\mathcal{X}) \leq \varepsilon \qquad \Longleftrightarrow \qquad \mu\{x : \mu(B_\varepsilon(x)) \leq \delta\} = 0.$$

This means that there are μ-almost surely no points which are too thin (in the sense that $\mu(B_\varepsilon(x)) \leq \delta$).

Definition A.8 (Size of ε-separated set)**.** Let $\underline{\underline{r}} \in \mathbb{R}_+^{\binom{\mathbb{N}}{2}}$. For $\varepsilon > 0$, define the *maximal size of an ε-separated set* by

$$\xi_\varepsilon(\underline{\underline{r}}) := \sup\big\{N \in \mathbb{N} : \ \exists k_1 < ... < k_N : (r_{k_i,k_j})_{1\leq i<j\leq N} \in (\varepsilon, \infty)^{\binom{N}{2}}\big\}.$$

Lemma A.9 (ξ_ε is constant, $\nu^{\mathcal{X}}$-almost surely)**.** *Let $\mathcal{X} \in \mathbb{N}$ and $\varepsilon > 0$. Then, ξ_ε is constant, $\nu^{\mathcal{X}}$-almost surely and equals*

$$\xi_\varepsilon(\mathcal{X}) := \inf\big\{N \in \mathbb{N} : \nu^{\mathcal{X}}\big(\rho_N^{-1}\big((\varepsilon, \infty)^{\binom{N}{2}}\big) > 0\big\},$$

where $\rho_N : \mathbb{R}^{\binom{\mathbb{N}}{2}} \to \mathbb{R}^{\binom{N}{2}}$ is the projection on the first $\binom{N}{2}$ coordinates.

Proof. Assume $\mathcal{X} = \overline{(X, r, \mu)}$. Let $x_1, x_2, ... \in X$ be such that

$$\xi_\varepsilon((r(x_i, x_j))_{1\leq i<j}) = N.$$

Then, N is the maximal size of an ε-separated set in (X, r), $\mu^{\otimes\mathbb{N}}$-almost surely. All results follow, since $\nu^{\mathcal{X}} = (R^{(X,r)})_*\mu^{\otimes\mathbb{N}}$ and since $\nu^{\mathcal{X}}$ is exchangeable. □

Remark A.10 (Tightness in $\mathbb{M}$). Recall from Theorem 2 in [GPW09] that for any $\mathcal{X} \in \mathbb{M}$, it holds that $v_\delta(\mathcal{X}) \xrightarrow{\delta\to 0} 0$. Moreover, a set $\Gamma \subseteq \mathbb{M}$ is pre-compact iff $\{w_\mathcal{X} : \mathcal{X} \in \Gamma\}$ is tight (as a family in $\mathscr{M}_1(\mathbb{R}_+)$) and $\sup_{\mathcal{X}\in\Gamma} v_\delta(\mathcal{X}) \xrightarrow{\delta\to 0} 0$.

This leads to a characterization of tightness for a family of random mm-spaces, see [GPW09], Theorem 3: Here, (the distributions of) a family $\{\mathcal{X} : \mathcal{X} \in \Gamma\}$ of $\mathbb{M}$-valued random variables is tight iff $\{\langle w_\mathcal{X}\rangle : \mathcal{X} \in \Gamma\}$ is tight (where $\langle w_\mathcal{X}\rangle$ is the first moment measure of $(w_\mathcal{X})_*\mathbf{P} \in \mathscr{M}_1(\mathscr{M}_1(\mathbb{R}_+))$ and $\sup_{\mathcal{X}\in\Gamma} \mathbf{E}[v_\delta(\mathcal{X})] \xrightarrow{\delta\to 0} 0$. Given a sequence of random mm-spaces, we can use these results in order to obtain limiting objects, at least along subsequences.

Now we come to a characterization of compact mm-spaces.

Theorem A.11 (Compact mm-spaces)**.** *Let $\mathcal{X} \in \mathbb{M}$. The following conditions are equivalent.*

(a) The mm-space $\mathcal{X}$ is compact, i.e. $\mathcal{X} \in \mathbb{M}_c$.

(b) For all $\varepsilon > 0$, it holds that $\xi_\varepsilon(\mathcal{X}) < \infty$.

(c) For all $\varepsilon > 0$, there is $\delta > 0$ such that $\widetilde{v}_\delta(\mathcal{X}) \le \varepsilon$.

The following characterization of random, almost surely compact mm-spaces is immediate.

Corollary A.12 (Random compact mm-spaces)**.** *Let $\mathcal{X}$ be a random variable taking values in $\mathbb{M}$. The following conditions are equivalent.*

(a) The mm-space $\mathcal{X}$ is compact, almost surely, i.e. $\mathbf{P}(\mathcal{X} \in \mathbb{M}_c) = 1$.

(b) For all $\varepsilon > 0$, it holds that $\mathbf{P}(\xi_\varepsilon(\mathcal{X}) < \infty) = 1$.

(c) For all $\varepsilon > 0$, there is a random variable $\Delta > 0$ with

$$\mathbf{P}(\widetilde{v}_\Delta(\mathcal{X}) \le \varepsilon) = 1.$$

Remark A.13 (Size of ε-separated set and size of ε-covering). The following observation will be used in the proof of Theorem A.11: Let (X, r) be a metric space and $\varepsilon > 0$, let ξ_ε be the maximal size of an ε-separated set and N_ε be the minimal number of ε-balls needed to cover (X, r). Then

$$N_\varepsilon \le \xi_\varepsilon \le N_{\varepsilon/2}.$$

In order to see this, let $x_1, ..., x_{\xi_\varepsilon}$ be a maximal ε-separated set. Then, $X = \bigcup_{i=1}^{\xi_\varepsilon} B_\varepsilon(x_i)$, since otherwise, we find $x \in X \setminus \Big(\bigcup_{i=1}^{\xi_\varepsilon} B_\varepsilon(x_i)\Big)$ and hence, the set is not maximal. This shows $N_\varepsilon \leq \xi_\varepsilon$. For the second inequality, it is clear that $B_{\varepsilon/2}(x_1), ..., B_{\varepsilon/2}(x_{\xi_\varepsilon})$ are disjoint. Hence, any set of centers of $\varepsilon/2$-balls which cover (X, r) must hit each $B_{\varepsilon/2}(x_i)$ at least once. As a consequence, $\xi_\varepsilon \leq N_{\varepsilon/2}$.

Proof of Theorem A.11. Let $\mathcal{X} = \overline{(X, r, \mu)}$. We use the notation laid out in Remark A.7. In particular, recall (A.1).

(1) $\Rightarrow$ (2): Let $\mathcal{X}$ be compact and $\varepsilon > 0$. Then (X, r) is totally bounded and there is $N_{\varepsilon/2} \in \mathbb{N}$ such that (X, r) can be covered by $N_{\varepsilon/2}$ balls of radius $\varepsilon/2$. Then we find $\xi_\varepsilon(\mathcal{X}) \leq N_{\varepsilon/2} < \infty$ by the last remark.

(2) $\Rightarrow$ (3): Let $\varepsilon > 0$. The space (X, r) can be covered by $\xi_{\varepsilon/2}(\mathcal{X}) < \infty$ balls of radius $\varepsilon/2$, again by the last remark. Let $x_1, \dots, x_{\xi_{\varepsilon/2}}$ be centers of such balls and $\delta := \min\{\mu(B_{\varepsilon/2}(x_i)) : \mu(B_{\varepsilon/2}(x_i)) > 0\}$. Then $\delta > 0$. Now take any $x \in X$ and choose $i \in \{1, \dots, \xi_{\varepsilon/2}\}$ such that $x \in B_{\varepsilon/2}(x_i)$. Then we have

$$\mu(B_\varepsilon(x)) \geq \mu(B_{\varepsilon/2}(x_i)) \geq \delta.$$

Hence,

$$\nu^{\mathcal{X}}\{\underline{\underline{r}} : s_\varepsilon(\underline{\underline{r}}) \leq \delta\} = \mu\{x \in X : \mu(B_\varepsilon(x)) \leq \delta\} = 0.$$

(3) $\Rightarrow$ (1): It suffices to show that (X, r) is totally bounded. Let $\varepsilon > 0$. By assumption, there is $\delta > 0$ such that

$$\nu^{\mathcal{X}}\{\underline{\underline{r}} : s_\varepsilon(\underline{\underline{r}}) \leq \delta\} = \mu\{x \in X : \mu(B_\varepsilon(x)) \leq \delta\} = 0.$$

We show that there is a finite maximal 2ε-separated set in X. For this, take a maximal 2ε-separated set $S \subseteq X$ (and without loss of generality assume that $\mathrm{supp}(\mu) = X$). Then, using the last remark,

$$1 = \mu(X) = \mu\Big(\bigcup_{x \in S} B_{2\varepsilon}(x)\Big) \geq \mu\Big(\bigcup_{x \in S} B_\varepsilon(x)\Big) = \sum_{x \in S} \mu(B_\varepsilon(x)) \geq |S| \cdot \delta,$$

since $\mu(B_\varepsilon(x)) > \delta$ holds μ-almost surely by assumption. Now,

$$|S| \leq 1/\delta < \infty$$

and $\varepsilon > 0$ was arbitrary, so (X, r) is totally bounded. □

Next, we come to a characterization of locally compact mm-spaces. Again some notation is needed.

Definition A.14 (δ-restriction)**.** Let $\underline{\underline{r}} := (r_{ij})_{1\le i<j} \in \mathbb{R}_+^{\binom{\mathbb{N}}{2}}$. Set $\widehat{\tau}_\delta(0) := 1$ and

$$\widehat{\tau}_\delta(i+1) := \inf\{j > \widehat{\tau}_\delta(i) : r_{1j} \le \delta\}.$$

Then,

$$\tau_\delta(\underline{\underline{r}}) := (r_{\widehat{\tau}_\delta(i),\widehat{\tau}_\delta(j)})_{1\le i<j}$$

is called the δ-*restriction* of $\underline{\underline{r}}$.

Remark A.15 (δ-restriction for distance matrices.). Let $\mathcal{X} = \overline{(X, r, \mu)} \in \mathbb{M}$ and $x_1, x_2, ... \in X$. We note that $x_k \in \widehat{\tau}_\delta(\mathbb{N})$ iff $r(x_1, x_k) \le \delta$. Hence, $\tau_\delta((r(x_i, x_j))_{1\le i\le j})$ is the distance matrix distribution for points among $x_2, x_3, ...$ which have distance at most δ to x_1. So,

$$\begin{aligned}(\tau_\delta)_*\nu^{\mathcal{X}}(\cdot) &= \nu^{\mathcal{X}}\{\tau_\delta(\underline{\underline{r}}) \in \cdot\} = \nu^{\mathcal{X}}\{\underline{\underline{r}} \in \cdot | r_{12}, r_{13}, ... \le \delta\} \\ &= \mu^{\otimes\mathbb{N}}\{(r(x_i, x_j))_{1\le i<j} \in \cdot | r(x_1, x_j) \le \delta \text{ for all } j = 2, 3, ...\}\end{aligned}$$

Clearly, $(\tau_\delta)_*\nu^{\mathcal{X}}$ is exchangeable, since $\nu^{\mathcal{X}}$ is exchangeable.

Theorem A.16 (Locally compact mm-spaces)**.** *Let $\mathcal{X} \in \mathbb{M}$. The following conditions are equivalent.*

(a) The mm-space $\mathcal{X}$ is locally compact, $\mathcal{X} \in \mathbb{M}_{lc}$.

(b) It holds that

$$\nu^{\mathcal{X}}\Big(\bigcap_{0<\eta<\delta} \Big\{\underline{\underline{r}} : \xi_\eta(\tau_\delta(\underline{\underline{r}})) < \infty\Big\}\Big) \xrightarrow{\delta\to 0} 1.$$

Proof. Let $\mathcal{X} = \overline{(X, r, \mu)}$. Then, $\mathcal{X}$ is locally compact iff for μ-almost all $x \in X$ there is $\delta > 0$, such that the ball $B_\delta(x)$ can be covered by a finite number of balls with radius η, for all $0 < \eta < \delta$. Hence,

$$\begin{aligned}
1 =& \mu\Big(\bigcup_{\delta>0}\bigcap_{0<\eta<\delta}\{x : B_\varepsilon(x) \text{ can be covered by finitely many balls of radius } \eta\}\Big)\\
=& \lim_{\delta\to 0}\mu\Big(\bigcap_{0<\eta<\delta}\{x : \text{ the maximal } \eta\text{-separated set in } B_\delta(x) \text{ is finite}\}\Big)\\
=& \lim_{\delta\to 0}\mu^{\otimes\mathbb{N}}\Big(\bigcap_{0<\eta<\delta}\{(x_1,x_2,...) : \xi_\eta((r_{x_i,x_j})_{2\le i<j}) < \infty |\\
& \qquad r(x_1,x_2), r(x_1,x_3), ... < \delta\}\Big)\\
=& \lim_{\delta\to 0}\mu^{\otimes\mathbb{N}}\Big(\bigcap_{0<\eta<\delta}\{(x_1,x_2,...) : \xi_\eta(\tau_\delta((r_{x_i,x_j})_{1\le i<j})) < \infty\}\Big)\\
=& \lim_{\delta\to 0}\nu^{\mathcal{X}}\Big(\bigcap_{0<\eta<\delta}\Big\{\underline{\underline{r}} : \xi_\eta(\tau_\delta(\underline{\underline{r}})) < \infty\Big\}\Big).
\end{aligned}$$

□

A.3 Λ-coalescents

We come to the application of the general results from the last section to metric spaces which arise in the context of coalescents which allow for multiple mergers. The proof of Theorem A.1 is given in the next section. Introduced by [Pit99], Λ-coalescents are usually described by Markov processes taking values in partitions of $\mathbb{N}$, which become coarser as time evolves, almost surely, and are exchangeable. More exactly, we define $(\Pi_t)_{t\ge 0} = (\Pi_t^\Lambda)_{t\ge 0}$, starting in the trivial partition of $\mathbb{N}$. For a finite measure Λ on $[0,1]$, set

$$\lambda_{b,k} = \int_0^1 x^{k-2}(1-x)^{b-k}\Lambda(dx). \tag{A.2}$$

Among any set of b partition elements in Π_t, each subset of size k merges to one partition element at rate $\lambda_{b,k}$. It is easy to check that such a process is well-defined (i.e. the $\lambda_{b,k}$'s are consistent) and leads to an exchangeable partition of $\mathbb{N}$ for all $t \ge 0$. In our analysis we restrict ourselves to measures Λ which do not have an atom at 1; see Example 20 in [Pit99] for a discussion of this case.

One intuitive way to construct a Λ-coalescent (given Λ has no atom at 0) is as follows: consider a Poisson-process with intensity measure $\frac{\Lambda(dx)}{x^2}\cdot dt$ on $[0,1]\times\mathbb{R}_+$. At any Poisson point (x,t), mark all partition elements, which are available by time t with probability x and merge all marked partition elements.

The set of Λ-coalescents falls into (at least) three classes. The class of Λ-coalescents coming down from infinity (see Property 1 in Theorem A.1), the larger class of processes having the *dust-free*-property, i.e. $f(\Pi_t^1) > 0$ for all $t > 0$, almost surely, where $f(\Pi_t^j)$ is the frequency of the partition element containing j at time t, $j \in \mathbb{N}$. All other Λ-coalescents contain *dust*, which is a positive frequency of natural numbers forming their own partition element.

Starting with [Sch00b], sharp conditions for a Λ-coalescents coming down from infinity have been given. Precisely, it was stated that a Λ-coalescent comes down from infinity iff

$$\sum_{b=2}^{\infty}\Big(\sum_{k=2}^{b} k\binom{b}{k}\lambda_{b,k}\Big)^{-1} < \infty. \tag{A.3}$$

It has been shown by [BG06] that this is equivalent to

$$\int_t^{\infty} \psi(q)^{-1} dq < \infty.$$

for some $t > 0$ where

$$\psi(q) = \int_0^1 (e^{-qx} - 1 + qx)x^{-2}\Lambda(dx).$$

The larger class of coalescents having the dust-free property is characterized by the requirement that

$$\int_0^1 x^{-1}\Lambda(dx) = \infty, \tag{A.4}$$

see Theorem 8 in [Pit99].

Let $\Pi^\Lambda := \Pi = (\Pi_t : t \geq 0)$ be the Λ-coalescent. Then for almost all sample paths of Π^Λ, there is a metric r^Π on $\mathbb{N}$, associated to Π, defined by

$$r^\Pi(i,j) := \inf\{t \geq 0 : i, j \text{ in the same partition element of } \Pi_t\},$$

that is the time needed for i and j to coalesce. We denote by (L^Π, r^Π) the completion of $(\mathbb{N}, r^\Pi)$. In order to equip (L^Π, r^Π) with a probability measure, we use a limit procedure. Set

$$H^n(\Pi) := \overline{\Big(L^\Pi, r^\Pi, \tfrac{1}{n}\sum_{i=1}^{n}\delta_i\Big)}.$$

Then, the family of $\mathbb{M}$-valued random variables $(H^n(\Pi))_{n=1,2,...}$ converges in distribution with respect to the Gromov-weak topology iff Π^Λ is dust-free, i.e. (A.4) holds (see Theorem 5 in [GPW09]). Since coalescent processes are associated with tree-like structures, we call the limiting mm-space $\mathcal{L} = \overline{(L^\Pi, r^\Pi, \mu^\Pi)}$ the Λ-*coalescent measure tree.*

Proof of Theorem A.1

Let $N(t) := \#\Pi_t$ denote the number of blocks in the partition Π_t and note that $\xi_\varepsilon(\mathcal{L}) \leq N(\varepsilon)$ where $\xi_\varepsilon(\mathcal{L}) < N(\varepsilon)$ is only possible if there are partition elements in Π_ε which carry no mass in $\mathcal{L}$.

(1) $\Rightarrow$ (2): Using Corollary A.12, we must show that for all $\varepsilon > 0$, we have $\xi_\varepsilon(\mathcal{L}) < \infty$ almost surely. This follows directly from the fact that $\xi_\varepsilon(\mathcal{L}) \leq N(\varepsilon)$ and the assumption that Π comes down from infinity.

(2) $\Rightarrow$ (1): The proof is by contradiction. Assume $\mathcal{L}$ is compact and Π stays infinite for some time $\varepsilon > 0$. Since Π_ε contains no dust, we have that $f((\Pi_\varepsilon^j)) > 0$ for all $j = 1, 2, ...$, almost surely. Since there are infinitely many lines up to time ε, we find partition elements of arbitrarily small mass. This implies that $\nu^{\mathcal{L}}\{\underline{\underline{r}} : s_\varepsilon(\underline{\underline{r}}) \leq \delta\} > 0$ almost surely, for all $\delta > 0$. On the other hand, since $\mathcal{L}$ is compact, there is a random variable $\Delta > 0$ such that $\nu^{\mathcal{L}}\{\underline{\underline{r}} : s_\varepsilon(\underline{\underline{r}}) \leq \Delta\} = 0$, almost surely by Corollary A.12. In particular, there is $\delta > 0$ such that

$$\nu^{\mathcal{L}}\{\underline{\underline{r}} : s_\varepsilon(\underline{\underline{r}}) \leq \delta\} = 0$$

with positive probability, which gives a contradiction.

Last, assume that $\mathcal{L}$ does not come down from infinity and recall that Λ cannot have an atom at 0 in this case. It has been shown in Proposition 23 of [Pit99] that the total coalescence rate of all lines is infinite for all times, almost surely. This is easy to see from the construction of Λ-coalescence using the Poisson process with intensity $\Lambda(dx)/x^2$, since the total coalescence rate of the partition element containing 1, given that there are infinitely many lines, is

$$\int_0^1 x \frac{\Lambda(dx)}{x^2} = \int_0^1 x^{-1} \Lambda(dx) = \infty,$$

since the dust-free property, (A.4), holds by assumption.

Let $0 < \eta < \delta$ and consider the δ-ball around 1 in L^Π. Since the coalescence rate is infinite and an infinite number of lines coalesce to the line containing 1 between times η and δ, there is an infinite η-separated set in $B_\delta(\{1\})$. Hence,

$$\nu^{\mathcal{L}}\{\underline{\underline{r}} : \xi_\eta(\tau_\delta(\underline{\underline{r}})) < \infty\} = 0,$$

almost surely. Hence, for any sequences $0 < \eta_n < \delta_n$ with $\delta_n \xrightarrow{n\to\infty} 0$, we find that

$$\nu^{\mathcal{L}}\Big(\bigcap_{0<\eta<\delta_n} \{\underline{\underline{r}} : \xi_\eta(\tau_{\delta_n}(\underline{\underline{r}})) < \infty\} \leq \nu^{\mathcal{L}}(\{\underline{\underline{r}} : \xi_{\eta_n}(\tau_{\delta_n}(\underline{\underline{r}})) < \infty\} = 0,$$

almost surely. By Theorem A.16, $\mathcal{L}$ cannot be locally compact.

APPENDIX B

Lebenslauf

Florian Holger Biehler

Merzhauserstraße 156

79100 Freiburg i. Brsg.

E-Mail: florian.holger.biehler@gmx.de

Tel. (privat): 01578/4883143

Persönliche Daten:

Geboren am	22. September 1984 in Donaueschingen / deutsch
Familienstand	ledig

Bildungsgang:

09/95 - 08/04	Besuch des Fürstenberg-Gymnasiums in Donaueschingen Kernkompetenz: Mathematik, Deutsch, Englisch Profilfach: Geschichte, Neigungsfach: Biologie Abiturnote: 2,2
10/05	Beginn des Mathematikstudiums an der Albert-Ludwigs-Universität Freiburg, Nebenfach Physik
10/07	Vordiplom LA I-II,Algebra: 1,3, Analysis I-III: 1,7, Stochastik,WT: 1,7, Physik: 2,3
09/08 - 02/09	Auslandsaufenthalt an der ENS Lyon, Frankreich, ERASMUS
danach	Vertiefungsgebiet Stochastik und besonderes Interesse an reiner Mathematik
12/10 - 06/11	Diplomarbeit: "Compact metric measure spaces and Λ-coalescents" bei Prof. Dr. Peter Pfaffelhuber
11/11	Voraussichtlicher Abschluss: Diplom Prüfungen voraussichtlich u.a. über stochastische Prozesse, stochastische Integration, Martingalprobleme, Funktionalanalysis und Topologie

Universitäres und außeruniversitäres Arbeiten:

seit 10/07	Wissenschaftlicher Mitarbeiter an der Albert-Ludwigs-Universität Freiburg: Tutorate zu Linearer Algebra und Wahrscheinlichkeitstheorie
02/11 - 06/11	Mitarbeit bei interConcept Medienagentur, München Erstellung von Themeneinheiten mit Schulwissen für ein Internetportal
05/11	"Compact metric measure spaces and Λ-coalescents coming down from infinity" in Zusammenarbeit mit Peter Pfaffelhuber, submitted to ALEA, http://arxiv.org/pdf/1105.2409

Sonstige Tätigkeiten:

09/03	Probestudium Mathematik, LMU-Mathe-Sommer 2003, München
10/04 - 06/05	Wehrdienst, 9. LWAusbRgt3 in Germersheim und Einsatzführungskompanie 11 in Meßstetten
04/05	Studienvorbereitung Mathematik, SRH Business Academy GmbH, Heidelberg

Zusätzliche Qualifikationen:

Sprachen:	Englisch - sichere Beherrschung in Wort und Schrift Französisch - GER B1
EDV-Kenntnisse:	Sicherer Umgang mit LaTeX und den üblichen Office-Programmen

---

Freiburg i. Brsg., Mai 2011 **Florian H. Biehler**

Bibliography

[ABBG10] L. Addario-Bery, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. *Probab. Theory Relat. Fields*, online first, 2010.

[Ald85] D. J. Aldous. *Exchangeability and related topics, Lecture Notes in Mathematics*. Springer, 1985.

[Ald93] D. J. Aldous. The continuum random tree iii. *Ann. Prop.*, 21, 248-289, 1993.

[Ald99] D. J. Aldous. Deterministic and stochastic models for coalescent (aggregation and coagulation): a review of the mean-field theory for probabilists. *Bernoulli*, 5 (1), 3-48, 1999.

[App04] D. Applebaum. Lévy processes: from probability to finance and quantum groups. *Notices of the AMS*, 2004.

[Bau02] H. Bauer. *Wahrscheinlichkeitstheorie, 5. Auflage*. De Gruyter Lehrbuch, 2002.

[BBI01] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*. Graduate studies in mathematics. AMS, Boston, 2001.

[BBS10] J. Berestycki, N. Berestycki, and J. Schweinsberg. The genealogy of branching brownian motion with absorbtion. *submitted*, arXiv:1001.2337v1, 2010.

[Ber09] N. Berestycki. Recent progress in coalescent theory. *Sociedade Brasileira de Matemática, Ensaios Matemáticos*, 16, 1-193, 2009.

[BG06] J. Bertoin and J.-F. Le Gall. Stochastic flows associated to coalescent processes iii: limit theorems. *Illinois J. Math*, 50, 147-181, 2006.

[BL10] N. Berestycki and V. Limic. The lambda-coalescent speed of coming down from infinity. *Ann of Prob*, 38, Number 1, 2010.

[Bou66] N. Bourbaki. *Elements of Mathematics, General Topology, Part 1.* Hermann, 1966.

[BS98] E. Bolthausen and A. S. Sznitman. On ruelles probability cascades and an abstract cavity method. *Comm. Math. Phys.*, 197 247-276, 1998.

[DGP11a] A. Depperschmidt, A. Greven, and P. Pfaffelhuber. Marked metric measure spaces. *Elec. Comm. Probab.*, 2011.

[DGP11b] A. Depperschmidt, A. Greven, and P. Pfaffelhuber. Tree-valued fleming-viot dynamics with mutation and selection. *Arxiv math.PR/1101.0759, submitted to Ann. Appl. Probab.*, 2011.

[Dur05] R. Durrett. *Probability: Theory and Examples, third edition.* Thomson Brooks/Cole, 2005.

[EK86] S. N. Ethier and T. G. Kurtz. *Markov Processes, Characterization and Convergence.* John Wiley and Sons, INC, Publication, 1986.

[Els02] J. Elstrodt. *Maß- und Integrationstheorie, 3., erweiterte Auflage.* Springer, 2002.

[EO94] S. N. Evans and N. O'Connell. Weighted occupation time for branching particle systems and a representation for the supercritical superprocess. *Canad. Math. Bull.*, 37 (2), 187-196, 1994.

[EPW06] S. N. Evans, J. Pitman, and A. Winter. Rayleight processes, real trees, and root growth with re-grafting. *Prop. Theo. Rel. Fields.*, 134 (1), 81-126, 2006.

[Eva00] S. N. Evans. Kingman's coalescent as a random metric space. *Ann of Prop*, 2000.

[Fel71] W. Feller. *An Introduction to Probability Theory and its Applications II.* J. Wiley and Sons., Inc., 1971.

[Gal99] J.-F. Le Gall. *Spatial branching processes, random snakes and partial differential equations.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1999.

[Gal07] J.-F. Le Gall. The topological structure of scaling limits of large planar maps. *Invent. Math.*, 169, 621-670, 2007.

[GPW09] A. Greven, P. Pfaffelhuber, and A. Winter. Convergence in distribution of random metric measure spaces (the lambda-coalescent measure tree). *Prob. Theo. Rel. Fields*, 2009.

[GPW10] A. Greven, P. Pfaffelhuber, and A. Winter. Tree-valued resampling dynamics (martingale problems and applications). *submitted to Prob. Theo. Rel. Fields*, 2010.

[Gro99] M. Gromov. *Metric structures for Riemannian and Non-Riemannian spaces.* Birkhäuser, Basel, 1999.

[Hey77] H. Heyer. *Probability Measures on Locally Compact Groups.* Springer-Verlag, Berlin-Heidelberg, 1977.

[Hun56] G. A. Hunt. Semigroup of measures on lie groups. *Transactions of the American Mathematical Society*, 81, No. 2, pp. 264-293, 1956.

[Kal02] O. Kallenberg. *Foundations of Modern Probability, second edition.* Springer, 2002.

[Kin82a] J. F. C. Kingman. The coalescent. *Stochastic Process. Appl. 13 235-248*, 1982.

[Kin82b] J. F. C. Kingman. On the genealogy of large populations. *Applied Probability Trust, Sheffield, England*, In Essays in Statistical Science 27-43, 1982.

[Kin00] J. F. C. Kingman. Origins of the coalescent: 1974-1982. *Genetics 156: 1461-1463*, 2000.

[Kle08] A. Klenke. *Wahrscheinlichkeitstheorie, 2., korrigierte Auflage.* Springer, 2008.

[Lim04] B. V. Limaye. *Functional analysis, seconde edition.* New Age International (P) Ltd., 2004.

[LS06] V. Limic and A. Sturm. The spatial lambda-coalescent. *Electron. J. Probab. 11*, 2006.

[Pit99] J. Pitman. Coalescents with multiple collisions. *Ann. Prop.*, 27(4), 1870-1902, 1999.

[Pit05] J. Pitman. *Combinatorial Stochastic Processes, Lecture Notes in Mathematics.* Springer, 2005.

[Rin75] W. Rinow. *Topologie.* VEB Deutscher Verlag der Wissenschaft, Berlin, 1975.

[Sch00a] J. Schweinsberg. Coalescent with simultaneous multiple collisions. *Electron. J. Probab. 5*, 2000.

[Sch00b] J. Schweinsberg. A necessary and sufficient condition for lambda-coalescent to come down from infinity. *Elec. Comm. Prob.*, 5, 1-11, 2000.

[SJAS78] L. A. Steen and Jr J. A. Seebach. *Counterexamples in topology, seconde edition.* Springer-Verlag, 1978.

[Ver00] A. M. Vershik. The universal urysohn space, gromov metric triples and random matrices on the natural numbers. *Elec. Comm. Prob.*, 5, 1-11, 2000.

[Win07] A. Winter. *Tree-valued Markov limit dynamics.* PhD thesis, Universität Erlangen-Nürnberg, 2007.

Nomenclature

η_b rate more manageable and alternative to γ_b, page 49

γ_b rate at which the number of blocks is decreasing, page 49

$\hat{\mu}_{\mathcal{X}}$ random distance distribution, page 14

$\mathbb{M}$ space of metric measure spaces, page 6

$\mathbb{M}_c$ space of compact metric measure spaces, page 21

$\mathbb{M}_{lc}$ space of locally compact metric measure spaces, page 21

$\mathbb{P}[\cdot]$ expectation operator, page 6

$\mathbb{S}$ set of all partitions of $\mathbb{N}$, page 34

$\mathbb{S}_n$ set of all partitions of $\{1, \ldots, n\}$, page 34

$\mathbb{X}_c$ space of compact metric spaces, page 64

$\mathcal{S}$ $:= \{(f_0, f_1, f_2, \ldots) \in (0,1)^{\mathbb{N}} : f_1 \geq f_2 \geq \ldots, \sum_{i \geq 0} f_i = 1\}$, page 36

$\mathcal{E}$ exchangeable σ-field, page 75

$\mathscr{A}$ algebra of all polynomials on $\mathbb{M}$, page 8

$\mathscr{B}(X)$ Borel-σ-algebra on metric space X, page 6

$\mathscr{M}_f(X)$ space of finite measures on $\mathscr{B}(X)$, page 6

[n]	$:= \{1, \ldots, n\} \subset \mathbb{N}$, page 34
$\nu^{\mathcal{X}}$	distance matrix distribution, page 7
d^{Π}	metric on $\mathbb{N}$ associated to a Λ-coalescent Π, page 57
$\mathrm{d}_{\mathbb{S}}$	complete metric on $\mathbb{S}$, page 34
d_{H}	Hausdorff metric for closed subsets of a metric space, page 65
d_{GH}	Gromov-Hausdorff metric on $\mathbb{X}_c$, page 65
$\mathrm{d}_{\mathrm{GPr}}$	Gromov-Prohorov metric on $\mathbb{M}$, page 9
d_{Pr}	Prohorov metric on $\mathscr{M}_1(Z)$, page 9
diam(A)	diameter of a set A, page 62
dis(R)	distortion of a correspondence R, page 65
$\mathrm{R}^{(\mathrm{X,r})}$	distance matrix of a given metric space (X, r), page 7
supp(μ)	support of a measure μ, page 6
$\stackrel{d}{=}$	equality in distribution, page 75
π_n^m	restriction operator from $\mathbb{R}_+^{\binom{m}{2}}$ to $\mathbb{R}_+^{\binom{n}{2}}$, where $m \geq n$, page 7
ρ_n^m	restriction operator from $\mathbb{S}_m$ to $\mathbb{S}_n$, where $m \geq n$, page 34
$\Rightarrow$	weak convergence, page 69
$\sim_{\mathcal{P}}$	equivalence relation on $\mathbb{N}$ associated to a partition $\mathcal{P}$ of $\mathbb{N}$, page 34
$\tau_\delta(R)$	δ-restriction of $R = (r_{i,j})_{1 \leq i < j} \in \mathbb{R}_+^{\binom{\mathbb{N}}{2}}$, page 24
$\tilde{v}_\delta(\mathcal{X})$	modification of the modulus of mass distribution, page 24
$\varprojlim$	projective limit, page 79
$\vee$	maximum, page 62
$\wedge$	minimum, page 64
ξ_ϵ	size of an ϵ-separated set, page 21

$N(t)$ number of blocks in a Λ-coalescent at time t, page 47

$v_\delta(\mathcal{X})$ modulus of mass distribution, page 12

$w_{\mathcal{X}}$ distance distribution, page 12

Index

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass alle Stellen der Arbeit, die wörtlich oder sinngemäß aus anderen Quellen übernommen wurden, als solche kenntlich gemacht sind und dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt wurde.

Freiburg, den 30. Mai 2011 Biehler