



HAL
open science

Regularity of generalized stochastic processes

Brice Hannebicque

► **To cite this version:**

Brice Hannebicque. Regularity of generalized stochastic processes. Probability [math.PR]. Université Paris-Saclay, 2021. English. NNT : 2021UPASM007 . tel-03260689

HAL Id: tel-03260689

<https://theses.hal.science/tel-03260689>

Submitted on 15 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Regularity of generalized stochastic processes

Régularité de processus stochastiques généralisés

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n°574, Mathématiques Hadamard (EDMH)
Spécialité de doctorat : mathématiques fondamentales
Unité de recherche : Fédération de Mathématiques de CentraleSupélec
(CNRS FR-3487), 91192, Gif-sur-Yvette, France
Réfèrent : CentraleSupélec

**Thèse présentée et soutenue à Paris-Saclay
le 23 mars 2021 par**

Brice HANNEBICQUE

Composition du jury :

Pauline LAFITTE Professeur, Université Paris-Saclay	Présidente
Robert J. ADLER Professeur émérite, Technion (Israel Institute of Technology)	Rapporteur et examinateur
Antoine AYACHE Professeur, Université de Lille	Rapporteur et examinateur
Nicolas CURIEN Professeur, Université Paris-Saclay	Examineur
Yimin XIAO Professeur, Michigan State University	Examineur
Érick HERBIN Professeur, Université Paris-Saclay	Directeur
Ely MERZBACH Professeur émérite, Bar-Ilan University	Invité

Nous sommes comme des nains juchés sur les épaules de géants.
Bernard de Chartres, d'après Jean de Salisbury, *Metalogicon* (Livre III).

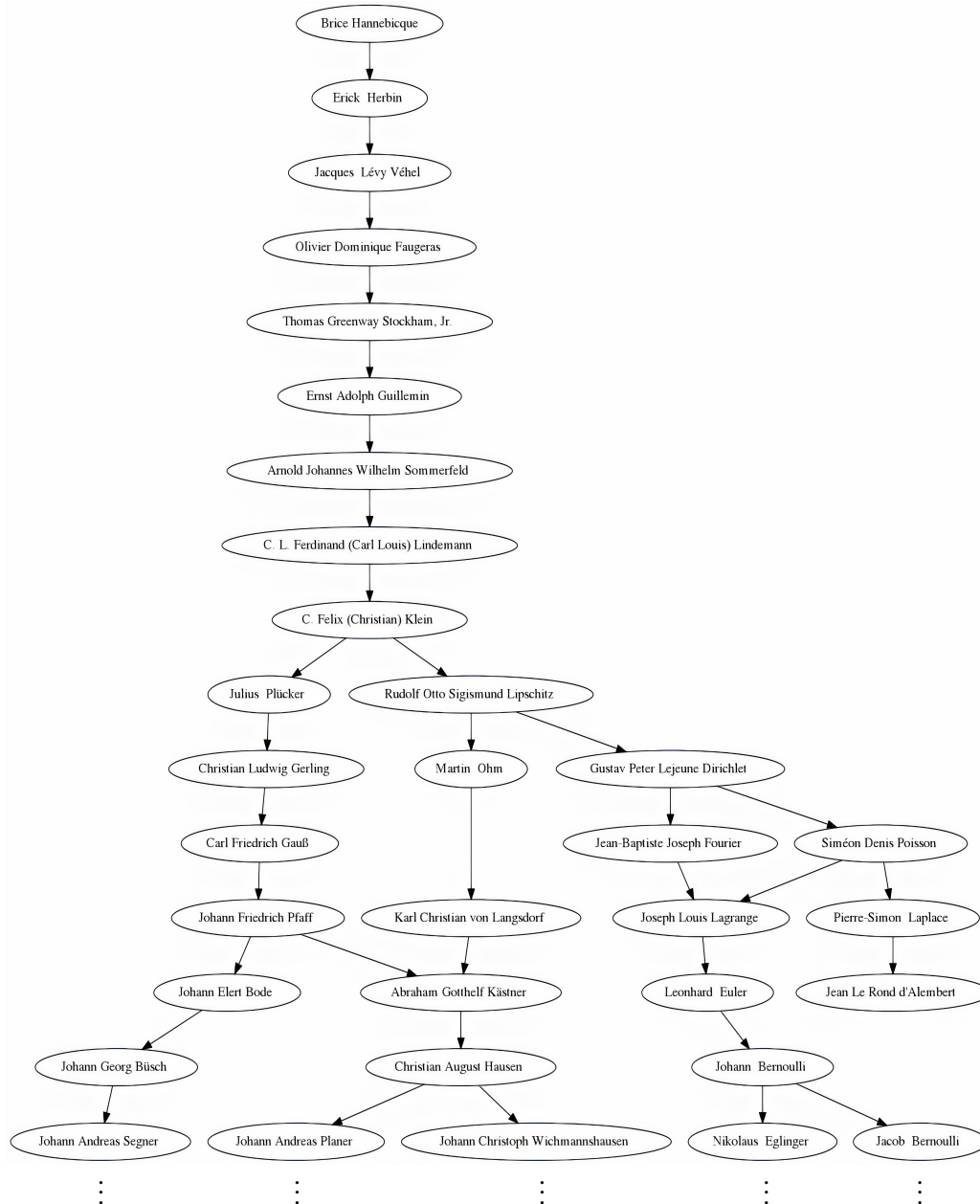


Figure 1: Les ancêtres mathématiques¹ de l'auteur sur 16 générations.

¹Chaque parent dans l'arbre a été directeur de thèse de son enfant (source : www.mathgenealogy.org).

Remerciements

Certain·e·s seront remercié·e·s plusieurs fois, c'est parce qu'ils·elles le méritent. Certain·e·s ont le même nom, c'est comme ça. Certain·e·s ne seront même pas mentionné·e·s, c'est parce que ma gratitude est tellement immense qu'elle ne peut pas tenir sur ces quelques pages². Lorsqu'ils sont appelés par leurs prénoms, les gens sont listés par ordre alphabétique, n'y voyez qu'une convention et non une préférence hiérarchisée dénuée de sens.

Commençons par les personnes impliquées au plus proche du processus de la thèse et de sa soutenance. À Érick Herbin, pour m'avoir pris sous votre aile depuis ma plus tendre enfance mathématique, j'ai beaucoup appris de vous et grandement apprécié notre dynamique durant toutes ces années. À Pauline Lafitte, pour avoir été disponible et à l'écoute tout du long, ton encre a coulé bien plus d'une fois pour moi, que ce soit pour d'innombrables lettres de recommandation ou pour le rapport final de mon jury où tu m'as fait l'honneur d'en être la Présidente. À Robert J. Adler, pour m'avoir fait l'honneur de rapporter ma thèse, vos écrits et votre humour m'accompagneront encore longtemps. À Antoine Ayache, pour m'avoir fait l'honneur de rapporter ma thèse, pour nos échanges et ta grande bienveillance. À Nicolas Curien, pour m'avoir fait l'honneur d'être examinateur, j'ai beaucoup appris du temps où j'étais parmi vos élèves. À Yimin Xiao, pour m'avoir fait l'honneur d'être examinateur, votre intérêt envers mes travaux m'a beaucoup touché. À Ely Merzbach, pour m'avoir fait l'honneur d'être invité, votre théorie a été mon plus fidèle compagnon tout au long de cette thèse. À Pierre, pour avoir streamé ma soutenance sur Twitch avec brio. À Christophe, Emma, Érick et Mahmoud, pour avoir relu sans rechigner un manuscrit mal dégrossi (ou au moins une portion), vos retours m'ont été fort précieux.

Grande a été ma chance d'être doctorant à l'EDMH et d'avoir un bureau au sein du laboratoire MICS où le cadre était idéal pour ma thèse, un idéal qui ne l'aurait pas autant été sans la contribution des personnes à venir. À Paul-Henry, pour ton accueil au sein du labo, du temps du parcours recherche, du stage de master et enfin pour la thèse. À Sylvie, notre fée aux pulls anachroniques et à la bécane jaune (dans ses rêves), qui me voyait bien en petit Schwarzie. À Fabienne, pour ta gentillesse et ta disponibilité. À Suzanne Thuron, pour votre serviabilité, votre efficacité et votre bonne humeur. À Farid, avec qui j'ai fini nombre de mes soirées au labo, pipelette intarissable et grand bonhomme jovial. À un autre agent d'entretien dont je n'ai jamais su le nom, j'espère que votre projet de vie se réalisera. Aux innombrables autres petites mains qui facilitent notre quotidien et avec qui mes échanges ont été plus brefs, mais je l'espère non moins agréablement reçus.

²et éventuellement aussi parce que je suis étourdi... mais non moins reconnaissant !

Plus largement, un bon nombre de personnes ont plus ou moins directement contribué à cette thèse, les voici. À Adrien, Alexandre G., Alexandre R., Antonin, Arthur, Aymeric, Christophe, Emma, Emmanuel, Emmanuelle, Érick, Fragkiskos, Gurvan, Houda, Léo, Léopold, Ludovic, Mahmoud, Pauline, Quentin, Romain, Sarah, Stéphane et Yoann pour au moins une³ discussion mathématique, impromptue ou non, qui m'aura marqué. À Mathilde, pour la Figure 1 de ce manuscrit. À Rémi et Guillaume, dont le bureau est bizarrement situé sur une étonnante singularité où le temps se déforme au fil des discussions rôlistes. À Gautier, Rémi, Stefania et Sylvain pour nos échanges musicaux. À la Franc-Pâtisserie, Mathilde, Myriam — sans oublier ta vorace moitié Manu m'ayant dépêtré de la jungle qu'est \LaTeX — et Jun, pour nos confections culinaires en tous genres. À Andreas, Fan, Laura, Ruiwen, Sophie et Xueyan, pour nos échanges gastronomiques extra-confrérie. À Gautier et Yoann pour les roustes au babyfoot⁴. Au chocolat aux noisettes de Gautier, dont la rareté ne faisait qu'accroître le plaisir de mettre la main (et la dent) dessus. Aux membres du groupe Magis pour avoir attisé ma flamme. À mes camarades du laboratoire que je n'ai pas encore remerciés, pour l'excellente et saine ambiance qui y règne : Agathe, Blandine, Chloé, Dimitri, Erwan, Elvire, Enzo, Julie, Léo, Maria P, Maria V, Mihir, Othmane, Théo et Wallid. À Damien et Irène, dont la présence rayonnante a éclairé nombre de mes journées ces derniers temps. À mes élèves, vous êtes la source de mes exaspérations, mais aussi de mes joies les plus grandes. À Vincent, pour ton amitié depuis la nuit des temps, ou au moins du nôtre. À mes amis plus généralement, qui m'ont épaulé par leur présence et envers qui j'ai une dette de gratitude : Ambroise, la smala Bonnet, Edwige, Gaétan, Guillemette, Jérôme, Louis-Marie, Mallory, Marie, Matthieu, Nathalie, Romain, Shuhui, Yuguang et ⁵.

Comme tout chercheur en herbe, j'eus l'occasion de participer à des *rallies* entre chercheurs. Je chéris ces épisodes comme des points particulièrement heureux et productifs de ma thèse. Je tiens donc à remercier tout particulièrement les organisateurs et les participants des Journées de Probabilités et de l'école d'été de probabilités de Saint-Flour de l'année 2019.

Une thèse, c'est également une "traversée du désert" comme le dit si justement Érick⁶. Cette traversée est heureusement ponctuée de quelques oasis, à la manière du célèbre épisode du marche-pied [81] que je veux brièvement rapporter ici.

Cette sieste dans le canapé à Bordeaux après une journée de calculs infructueux et avant un dîner entre amis qui n'allait pas me permettre de rallonger mes efforts outre mesure, réveillé d'une semi-torpeur à dix minutes de leur arrivée par une soudaine inspiration, je griffonne frénétiquement les détails du calcul qui s'avéra le plus délicat de ma thèse juste avant que mes convives ne sonnent à la porte.

Au Pérou en compagnie de Renaud, ce mail anodin sur l'abstract d'un prochain séminaire à l'autre bout du monde, qui m'expliquait que j'avais déjà effleuré la solution à mon problème il y a quelques mois déjà alors que je l'avais abandonné, il fallait juste que je reprenne mes notes, les mette au clair, relise une paire de références et la solution allait m'apparaître, ce que je fis dans les mois qui suivirent.

Cette journée de co-travail en compagnie de Zénon où les dernières pièces de la dernière preuve de ce manuscrit⁷ se sont agencées avant le grand saut dans la rédaction.

³et pour certain-e-s, beaucoup plus !

⁴Bien qu'on sache très bien qui est le meilleur de vous deux...

⁵Ceux-elles non cité-e-s ailleurs, si vous n'apparaissez pas dans cette liste et que vous sentez devoir en faire partie, c'est que vous avez certainement raison : imprimez une version de ce manuscrit et ajoutez votre nom après le "et " laissé à cet effet.

⁶D'autres personnes ont déjà employé cette terminologie il y a une paire de millénaires, mais les lois sur le plagia étaient encore un peu floues à l'époque.

⁷techniquement située au beau milieu de celui-ci

Merci donc à mon canapé, mes mails, Renaud et Zénon pour avoir été mes "marche-pieds" hors du "désert".

À ma famille pour son support affectueux et inébranlable, mes parents en particulier, sans qui je ne serai tout bonnement pas là.

Mot de bienvenue

Cher·ère lecteur·rice, voici quelques mots pour adoucir votre entrée dans ce manuscrit. Plonger dans les pensées mathématiques d'autrui n'est que rarement chose aisée, d'où mon envie d'essayer d'alléger votre peine, ne serait-ce qu'un peu.

Vous trouverez page [xix](#) l'ensemble des notations et conventions qui seront utilisées. Même si j'essaierai autant que possible de rappeler ce qu'il est bon d'avoir en tête à un moment donné, les hyperliens vers les notions pertinentes pourront toujours y être retrouvés.

Par ailleurs, comme il n'y a pas de conclusion définitive à ce travail, il n'y a pas de conclusion générale à ce manuscrit non plus. Vous trouverez cependant des perspectives de recherche (en français et en anglais) à la fin de chaque chapitre.

Passons maintenant à une introduction portant sur la structure du manuscrit ainsi que ses principaux résultats.

Comme le titre de cette thèse le suggère, l'objet central de ce manuscrit est un processus stochastique $X = \{X_t : t \in \mathcal{T}\}$ indexé par un ensemble général \mathcal{T} . L'objectif du Chapitre 1 est de partir d'hypothèses minimales sur \mathcal{T} et de développer un cadre général riche en exemples et en applications. La structure dont est équipé \mathcal{T} est un triplet $(\mathcal{T}, \preceq, \mathcal{D})$ où \preceq est une relation d'ordre partiel mimant un écoulement de temps et \mathcal{D} un sous-ensemble dénombrable dense⁸. Un temps (long) est passé en Section 1.2.2 pour donner des exemples et montrer comment en construire de plus variés à partir d'une poignée d'exemples simples. Les deux qu'on pourra avoir en tête à tout moment sont les suivants :

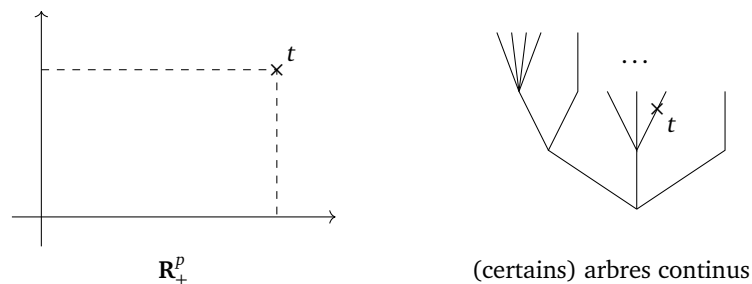


Figure 2: Deux exemples fondamentaux d'espaces \mathcal{T} indexant nos processus d'intérêt.

⁸Chose fort appréciée des probabilistes, car elle leur permet de transformer certaines intersections indénombrables d'événements en intersections dénombrables, bien plus digests.

- ◊ Le cas multiparamétrique $\mathcal{T} = \mathbf{R}_+^p$ (Exemple 1.2.10) où \preceq est la relation d'ordre partielle comparant coordonnée à coordonnée, *i.e.*

$$\forall s = (s_1, \dots, s_p), t = (t_1, \dots, t_p) \in \mathbf{R}_+^p, \quad s \preceq t \iff \forall i \in \llbracket 1, p \rrbracket, s_i \leq t_i. \quad (0.0.1)$$

Le sous-ensemble dénombrable dense \mathcal{D} peut alors être choisi comme étant l'ensemble \mathbf{Q}_+^p ou plus malin encore⁹, l'ensemble $\{k2^{-n} : k, n \in \mathbf{N}\}^p$ des vecteurs de \mathbf{R}_+^p à coordonnées dyadiques. Il s'agit probablement de la première généralisation au-delà de \mathbf{R}_+ qui a intéressé les probabilistes. Celle-ci a fait l'objet de nombreuses investigations, qui ont abouti sur une solide théorie qui leur est propre [56]. Il sera intéressant de voir que si certaines de leurs méthodes peuvent s'adapter, d'autres échouent et l'intuition donnée par notre cadre permet de découvrir des résultats multiparamétriques jusque-là encore inconnus (Exemple 3.5.16).

- ◊ Le cas où \mathcal{T} est un arbre continu (ou \mathbf{R} -arbre, cf. Exemple 1.2.25) obtenu comme recollement d'une famille au plus dénombrable de segments, qui feront office d'arêtes, et enraciné en un point $\rho \in \mathcal{T}$. La relation d'ordre \preceq est alors définie par

$$\forall s, t \in \mathcal{T}, \quad s \preceq t \iff \llbracket \rho, s \rrbracket \subseteq \llbracket \rho, t \rrbracket$$

où $\llbracket \rho, s \rrbracket$ (resp. $\llbracket \rho, t \rrbracket$) est l'unique segment géodésique entre la racine ρ et s (resp. t). Pour le sous-ensemble dénombrable dense \mathcal{D} , il suffit de choisir de prendre l'image par l'application de recollement d'un sous-ensemble dense sur chaque arête (c'est là où intervient la nécessité d'avoir une quantité au plus dénombrable d'arêtes).

Même s'il s'agit de nos deux exemples phares, le texte en est émaillé d'autres pour varier les plaisirs et exhiber une pluralité de comportements.

Par ailleurs, le Théorème 1.2.5 montre qu'étudier $X = \{X_t : t \in \mathcal{T}\}$ revient à étudier un autre processus $X = \{X_A : A \in \mathcal{A}\}$ indexé par une classe \mathcal{A} de sous-ensembles spécifiques de \mathcal{T} . Ces ensembles sont définis très naturellement à partir de la relation d'ordre \preceq :

$$\mathcal{A} = \{A(t) : t \in \mathcal{T}\} \cup \{\emptyset\} \quad \text{où} \quad A(t) = \{s \in \mathcal{T} : s \preceq t\}. \quad (0.0.2)$$

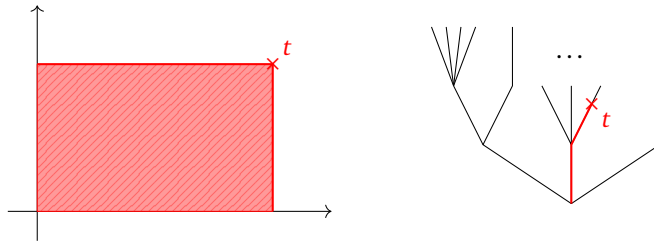


Figure 3: Illustrations d'un $A(t)$ dans les deux cas de la Figure 2.

Cela nous permet de faire le parallèle avec la théorie des processus indexés par des ensembles développée par Ivanoff, Merzbach et leurs co-auteurs [47]. Nous leur empruntons de nombreux concepts, notamment sur deux sujets :

⁹Plus malin pour des raisons qui apparaîtront plus clairement par la suite, mais disons pour l'instant que les dyadiques sont bien mieux "organisés" que les rationnels.

- ◊ Dans la Section 1.2.3, la définition d'un certain nombre de collections d'ensembles \mathcal{D} contenant \mathcal{A} et permettant de considérer plusieurs notions d'accroissements pour le processus X . En particulier, nous mettons en évidence (Proposition 1.2.44 et 1.6.4) la nécessité de la condition dite *SHAPE* (Définition 1.2.1) pour l'existence du processus d'accroissement ΔX de X ¹⁰. Les diverses extensions (et leur ordre d'introduction) de \mathcal{A} et X sont résumées dans les Figures 4 et 5 : vous êtes invité·e à vous y reporter en cas de doute durant votre lecture.

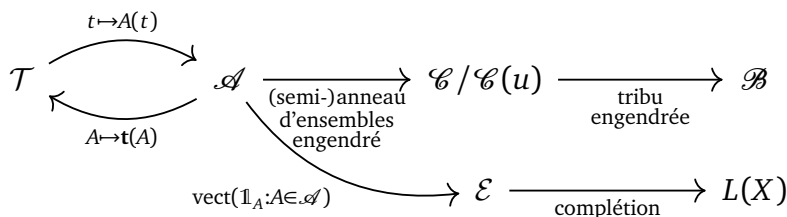


Figure 4: Principales collections (de points, d'ensembles ou de fonctions) indexant les processus d'intérêt dans cette thèse.

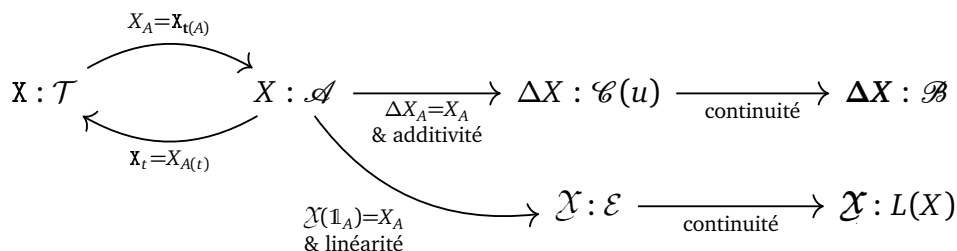


Figure 5: Le processus X (ou X), ses extensions et les collections sur lesquelles elles sont définies.

- ◊ Dans la Section 1.3.2, la définition de la notion de *flots* qui sont des fonctions croissantes¹¹ $\phi : [0, 1] \rightarrow \mathcal{A}$ (ou toute autre collection \mathcal{D} contenant \mathcal{A}) L'intérêt pour l'intuition est immédiat car un processus X peut alors se *projeter* le long de ϕ pour donner le processus $X^\phi = \{X_{\phi(t)} : t \in [0, 1]\}$. Ainsi, on peut ramener dans certains cas l'étude de X à celle de ses projections (Théorème 1.3.13), ce qui se révélera crucial dans le Chapitre 2 pour établir des formules de représentation lorsque X est à accroissements échangeables (Théorèmes 2.5.30 et 2.5.31).

Une comparaison détaillée de la théorie développée ici avec celle d'Ivanoff et Merzbach est menée en Section 1.6. On y montre notamment que notre cadre est plus général, sans toutefois minimiser les avantages offerts par la première théorie, notamment en ce qui concerne l'étude des martingales indexées par des ensembles où notre cadre est moins adapté, même si quelques résultats subsistent (cf. Section 2.3.3).

Outre les hypothèses décrites précédemment, nous équipons \mathcal{T} dans la suite du Chapitre 1 de trois nouvelles structures, plus modulaires dans le sens où elles se seront pas nécessaires à

¹⁰i.e. l'existence d'une extension additive à l'anneau d'ensembles engendré par \mathcal{A}

¹¹ \mathcal{A} étant une collection d'ensembles, elle est munie de la relation d'ordre naturelle qu'est \subseteq .

l'ensemble des résultats du manuscrit. Nous vous invitons à les voir comme des "patches" que vous pourriez vouloir ajouter à \mathcal{T} au gré de vos besoins.

- ◊ Le premier "patch" est la présence d'une mesure \mathbf{m} sur $\mathcal{B} = \sigma(\mathcal{A})$ (Section 1.3) avec une certaine propriété de point milieu qui donne une qualité continue à \mathcal{T} jusqu'alors non requise par notre théorie, contrairement à celle d'Ivanoff et Merzbach. Elle nous permet de construire en Section 1.3.2 des flots particuliers, dits *géodésiques* : ce sont des flots ϕ continus à droite tels que $\mathbf{m} \circ \phi$ est une fonction affine. Comme mentionné précédemment, ils seront utiles pour faire le lien avec la théorie unidimensionnelle. Une autre utilisation de la mesure \mathbf{m} est de donner en Section 2.5 une notion de "taille" aux accroissements, et donc permet de considérer des notions de stationnarité des accroissements dans ce cadre très général.
- ◊ Le second est une distance $d_{\mathcal{T}}$ sur \mathcal{T} (ou $d_{\mathcal{A}}$ sur \mathcal{A} comme rappelle la Figure 6) introduite en Section 1.4. Celle-ci permet enfin de rendre justice au premier mot dans la titre de la thèse : régularité. En effet, étudier la régularité fine des trajectoires du processus $X = \{X_t : t \in \mathcal{T}\}$ — objectif principal du Chapitre 3 — nécessite au bas mot d'avoir une distance sur \mathcal{T} .

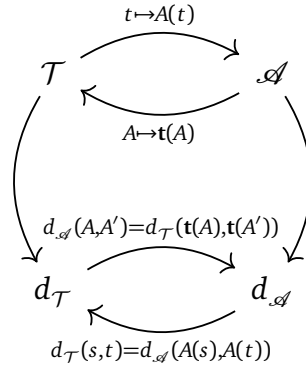


Figure 6: Correspondance entre les espaces métriques $(\mathcal{T}, d_{\mathcal{T}})$ and $(\mathcal{A}, d_{\mathcal{A}})$.

Une nouveauté notable par rapport au cadre d'Ivanoff et Merzbach est la notion de *vicinity* $\mathcal{V}(A, \rho)$ introduite à la Section 1.4.3 et donnée pour $A \in \mathcal{A}$ et $\rho > 0$ par

$$\mathcal{V}(A, \rho) = \bigcup_{\substack{A' \in \mathcal{A}: \\ d_{\mathcal{A}}(A, A') < \rho}} A \Delta A'.$$

Bien que sa définition puisse paraître obscure à première vue, elle constitue la bonne région à regarder lorsqu'on étudie des quantités de la forme $X_A - X_{A'}$ lorsque A' est proche de A . Cette notion sera mise à profit pour démontrer des inégalités maximales pour des martingales (Section 2.3.3), une loi du 0-1 (Section 2.3.4) et donner des bornes plus pertinentes sur la régularité hölderienne ponctuelle des processus étudiés dans le Chapitre 3 (Section 3.5).

- ◊ Le troisième est une notion pour \mathcal{T} (ou \mathcal{A}) d'être de *dimension finie* qui est développée en Section 1.5. Reprenant l'ensemble dénombrable \mathcal{D} permettant d'approcher tout élément de \mathcal{T} , il est possible de construire une suite $(\mathcal{A}_n)_{n \in \mathbf{N}}$ de discrétisations finies de \mathcal{A} . Pour chaque $n \in \mathbf{N}$, on est capable de plonger \mathcal{A}_n dans $\mathbf{R}_+^{p_n}$ en respectant la structure d'ordre¹² où $p_n \in \mathbf{N}$

¹²L'ordre sur \mathcal{A}_n est donné par l'inclusion \subseteq tandis que l'ordre sur $\mathbf{R}_+^{p_n}$ est l'ordre partiel défini en (0.0.1).

est la *dimension d'ordre* de \mathcal{A}_n . De plus, \mathcal{A}_n induit une partition de \mathcal{T} dont les éléments sont de la forme $A \setminus \bigcup_{A' \in \mathcal{A}_n: A' \subset A} A'$ pour tout $A \in \mathcal{A}_n$ mais dont l'écriture peut être simplifiée en enlevant les doublons en $A \setminus \bigcup_{i=1}^q A_i$ où $q = q(A)$ est choisi le plus petit possible. En notant $q_n = \max_{A \in \mathcal{A}_n} q(A)$, on obtient une autre manière de décrire la dimension de \mathcal{A}_n ¹³. Le nombre $\sup_{n \in \mathbb{N}} \max\{p_n, q_n\}$ donne à son tour une notion de dimension pour \mathcal{A} qui correspond intuitivement au nombre minimal de paramètres requis pour décrire \mathcal{A} ainsi que son approximation $(\mathcal{A}_n)_{n \in \mathbb{N}}$ comme schématisé dans la Figure 7.

$$\mathcal{A} \xrightarrow{\text{discrétisation}} \mathcal{A}_n \xrightarrow{\text{nombre de paramètres}} \dim \mathcal{A}$$

Figure 7: La dimension de \mathcal{A} exprime à quel point \mathcal{A} est bien approchée par ses discrétisations.

Nous montrons que cette définition implique que \mathcal{A} est nécessairement une classe de Vapnik-Červonenkis, permettant ainsi un nombre d'arguments à base d'entropie métrique pour obtenir la continuité des trajectoires de certains processus. Ceci dit, notre définition ne s'arrête pas là car elle nous permet non seulement d'étudier plus finement dans le Chapitre 3 l'aspect "càdlàg"¹⁴ de processus présentant des sauts mais aussi leur régularité hölderienne ponctuelle.

Le Chapitre 2 se concentre sur l'étude des propriétés de ΔX en tant que "prémessure stochastique". Plus précisément, on commence par présenter en Section 2.2 les travaux de Kwapien, Rajput, Rosiński, Urbanik, Woyczyński et leurs co-auteurs afin de donner des conditions nécessaires et suffisantes pour que ΔX puisse s'étendre en une mesure stochastique ΔX sur les boréliens. Cela permet de considérer l'intégrale \mathcal{X} de fonctions déterministes $f : \mathcal{T} \rightarrow \mathbf{R}$ contre X (cf. partie droite de la Figure 5), intégrale dont l'effet régularisant est étudié dans le Chapitre 3.

Certaines propriétés distributionnelles de ΔX sont ensuite étudiées. Nous commençons par le cas des accroissements indépendants (Section 2.3), qui se traduit peu ou prou¹⁵ par la propriété suivante :

$$\forall C, C' \in \mathcal{C}, \quad C \cap C' = \emptyset \implies \Delta X_C \perp \Delta X_{C'}.$$

Ce cas est bien compris largement grâce aux travaux de Rajput et Rosiński [83] qui montrent que dans ce cas, X doit nécessairement être un processus infiniment divisible possédant une représentation de Lévy-Khintchine. Notre contribution dans les Sections 2.3.1 et 2.3.2 est largement de l'ordre de l'exposition, à l'exception du calcul d'un certain nombre d'exposants de Lévy-Khintchine que nous n'avons pas trouvés ailleurs dans la littérature. On en profite ensuite pour démontrer quelques inégalités maximales (Section 2.3.3) et une loi du 0-1 (Section 2.3.4) qui se révéleront utiles pour une étude de la régularité hölderienne menée dans le Chapitre 3.

En guise d'interlude en Section 2.4, on présente les processus de Lévy indexés par des ensembles possédant deux propriétés remarquables : leurs accroissements sont indépendants dans le sens mentionné ci-dessus en plus d'être stationnaires au sens suivant¹⁶ :

$$\forall C, C' \in \mathcal{C}, \quad m(C) = m(C') \implies \Delta X_C \stackrel{\text{loi}}{=} \Delta X_{C'}.$$

¹³Les Exemples 1.5.6 et 1.5.7 montreront qu'aucune de ces deux notions n'implique l'autre.

¹⁴Continues à "droite" avec limites à "gauche", même si "droite" et "gauche" doivent être interprétées au travers du prisme de la relation d'ordre \preceq dont \mathcal{T} est muni.

¹⁵Il faudrait techniquement prendre une famille finie quelconque d'éléments de \mathcal{C} deux à deux disjoints.

¹⁶À nouveau, il faudrait prendre une famille finie quelconque de \mathcal{C} entre autres précautions.

Ces processus ont été introduits et étudiés par Herbin et Merzbach [41] où il est montré qu'une représentation de Lévy-Itô a lieu. Nous précisons ici ce résultat (Corollaire 2.4.9) et fournissons une preuve différente qui permet de préparer d'autres résultats de représentation par la suite (Théorèmes 2.5.30 et 3.3.8 entre autres).

Cette notion de stationnarité basée sur une mesure m mérite cependant que la Section 2.5 lui soit entièrement dédiée. Une vision globale sur le sujet est développée en Section 2.5.1 et donne naissance à plusieurs notions tour à tour étudiées :

- ◊ être à \mathcal{A} -accroissements stationnaires (Section 2.5.2) qui est certainement la bonne notion pour appréhender la stationnarité du mouvement brownien fractionnaire indexé par des ensembles (sifBm),
- ◊ être à accroissements échangeables (Section 2.5.3) et être échangeable au sens fonctionnel (Section 2.5.4) qui mènent pour leur part à des théorèmes de représentation (Théorèmes 2.5.30 et 2.5.31) qui généralisent ceux de Bühlmann et Kallenberg [55, Theorems 1.19 and 3.15] pour les processus indexés par \mathbf{R}_+ . Nous montrons qu'un processus $X = \{X_A : A \in \mathcal{A}\}$ est à accroissements échangeables si et seulement si il peut s'écrire sous la forme suivante :

$$\forall A \in \mathcal{A}, \quad X_A = bm(A) + \sigma \widehat{W}_A + \widetilde{Q}_A$$

où $(b, \sigma, J) : \Omega \rightarrow \mathbf{R} \times \mathbf{R}_+ \times \mathcal{M}(\mathbf{R}^*)$ est un triplet aléatoire, J une mesure de Lévy ponctuelle aléatoire, \widehat{W} un pont brownien indépendant indexé par des ensembles et \widetilde{Q} un processus ponctuel compensé d'intensité J ¹⁷ indépendant de ce qui précède conditionnellement à J .

Le Chapitre 3 est quant à lui consacré à l'étude des trajectoires des processus généralisés. Nous développons tout d'abord un espace de fonctions "continues" $\mathcal{C}^\Phi(\mathcal{A})$ (Section 3.2) et "càdlàg" $\mathcal{D}^\Phi(\mathcal{A})$ (Section 3.3) adaptés à notre cadre. En particulier, l'hypothèse de dimension finie pour \mathcal{A} donne une condition suffisante pour l'appartenance d'un grand nombre de processus à ces espaces (Proposition 3.2.2 et Théorème 3.3.8). Sous ces hypothèses et pour un processus de Lévy X indexé par des ensembles et une fonction $f : \mathcal{T} \rightarrow \mathbf{R}$ localement X -intégrable, la primitive Y de f par rapport à X donnée par

$$\forall A \in \mathcal{A}, \quad Y_A = \mathcal{X}(f \mathbb{1}_A) = \int_A f dX$$

est presque sûrement à trajectoires dans $\mathcal{D}^\Phi(\mathcal{A})$.

Dans la suite, nous investiguons plus en détails la régularité de Y . À cet effet, la Section 3.4 expose et étend le cadre posé par Herbin et Richard [42] pour l'étude de la régularité hölderienne ponctuelle des processus généralisés. Pour une fonction $h : \mathbf{R}_+ \rightarrow \mathbf{R}$, la définition de son exposant ponctuel en $t \in \mathbf{R}_+$ ne fait aucun doute¹⁸ : il s'agit du meilleur $\alpha \in \mathbf{R}_+$ tel que pour tout $s \in \mathbf{R}_+$ suffisamment proche de t , on ait $|h(s) - h(t)| \leq |s - t|^\alpha$. Cependant, pour notre cadre général, les choses ne sont pas aussi simples lorsqu'on veut parler de la régularité ponctuelle d'une fonction $h : \mathcal{A} \rightarrow \mathbf{R}$ en $A = A(t) \in \mathcal{A}$: de quels accroissements veut-on parler ? À cette question nous proposons deux réponses :

¹⁷Il s'agit d'un processus similaire à la partie poissonnienne d'un processus de Lévy dans sa représentation de Lévy-Itô. En particulier, c'est une limite de processus ponctuels compensés (cf. Exemple 2.5.23).

¹⁸Techniquement, nous commettons un impair de taille ici puisque nous oublions d'enlever à h une partie polynomiale représentant son développement de Taylor. Cependant, nous avons été bien en peine de trouver un *ersatz* de polynôme satisfaisant pour notre cadre.

- ◇ Les différences $h(A) - h(A')$ où A' est proche de A , qui mènent à la notion d'*exposant d'Hölder ponctuel* $\alpha_h(A)$ donnée en (3.4.1). Probablement la définition la plus instinctive, mais peut-être pas la plus naturelle. En effet, regarder ce genre d'accroissements fait intervenir des points $t \in \mathcal{T}$ "loins" de A : ce sont précisément ceux appartenant à la *vicinity* $\mathcal{V}(A, \rho)$ introduite précédemment et potentiellement beaucoup plus grosse que la boule $B_{\mathcal{A}}(A, \rho)$ ¹⁹. Dans un cas simple, le Théorème 3.5.2 implique presque sûrement :

$$\alpha_{\mathcal{V}}(A) = \alpha_X(A) + \alpha_{f, \mathcal{V}}(t) \mathbb{1}_{f(t) \neq 0} \quad (0.0.3)$$

où $\alpha_{f, \mathcal{V}}(t)$ est une forme d'exposant d'Hölder ponctuel pour f , mais où la distance $d_{\mathcal{T}}$ est remplacée par une autre grandeur \mathfrak{d} , appelée *divergence*, reliée à $\mathcal{V}(A, \rho)$ et donnée dans la Définition 1.4.6.

- ◇ Les accroissements $\Delta h(C)$ où $C \in \mathcal{C}$ est inclus dans une petite boule centrée en A ²⁰, qui mènent à la notion (nouvelle) d'*exposant d'Hölder $d_{\mathcal{T}}$ -localisé* $\alpha_{h, d_{\mathcal{T}}}(A)$ donnée en (3.4.7). Comme mentionné précédemment, cet exposant a l'avantage de ne faire intervenir que les points $t \in \mathcal{T}$ "proches" de A . Dans un cas simple, le Théorème 3.5.12 implique presque sûrement :

$$\alpha_{\mathcal{Y}, d_{\mathcal{T}}}(A) = \alpha_{X, d_{\mathcal{T}}}(A) + \alpha_f(t) \mathbb{1}_{f(t) \neq 0} \quad (0.0.4)$$

où $\alpha_f(t)$ est l'exposant d'Hölder ponctuel "usuel" de f en t .

Quoiqu'il en soit, (0.0.3) et (0.0.4) confirment bien qu'intégrer, même contre une mesure stochastique, est une opération régularisante en un certain sens. Mentionnons également que les Théorèmes 3.5.2 et 3.5.12 ne se réduisent pas à ces seules estimées et laissent entrevoir un mélange des régularités de f et X plus subtil qu'il n'y paraît dans des cas plus complexes.

Si ces quelques lignes ont plus attisé qu'assouvi votre curiosité, je vous souhaite une belle lecture !

¹⁹Léger abus ici, on devrait plutôt parler d'une boule dans \mathcal{T} (et non \mathcal{A}), mais ce n'est pas un soucis grâce à la correspondance donnée par (0.0.2).

²⁰Second abus identique au précédent.

A word of welcome

Dear reader, here are a few words to ‘ease your way’ into reading this thesis. Delving into another’s mathematics is not always an easy task, hence the need I felt to lighten your burden, even in some small way.

You will find page [xix](#) all the notations and conventions that I will be using. Even though I will try as much as possible to remind the reader of what they need to have in mind at a given point, all the links to the relevant concepts may be found over there.

Also, there is no definite conclusion to this work, but the reader may find perspectives at the end of each chapter.

The next few figures and their arrows in this introduction are meant to indicate the dependency between several concepts that will be introduced later on. I proceed to give a brief introduction about them.

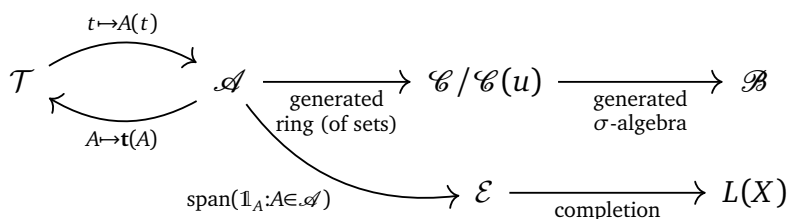


Figure 8: Main collections (points, sets or functions) indexing the processes of interest in this thesis.

As the title of the thesis suggests, my main object of interest has been stochastic processes $X = \{X_t : t \in \mathcal{T}\}$ indexed by quite a general set \mathcal{T} . A bit of time on examples is taken at the beginning (Section 1.2.2) to demonstrate the variety of situations the theory encompasses. In Chapter 1, we will see that, given only a couple of simple axioms on \mathcal{T} , it becomes equivalent to study processes $X = \{X_A : A \in \mathcal{A}\}$ where \mathcal{A} is some specific collection of subsets of \mathcal{T} called *indexing collection*. This ties in nicely with the theory of set-indexed processes exposed by Ivanoff and Merzbach [47] from which we borrow quite a few concepts. In particular, the class \mathcal{A} may be extended to bigger ones — and the process X along with it — to consider increments and linear extensions. All the steps are illustrated in Figures 8 and 9 and treated in Chapter 1 but for the rightmost part of Figure 9 — *i.e.* the extensions needing continuity — which is carried out

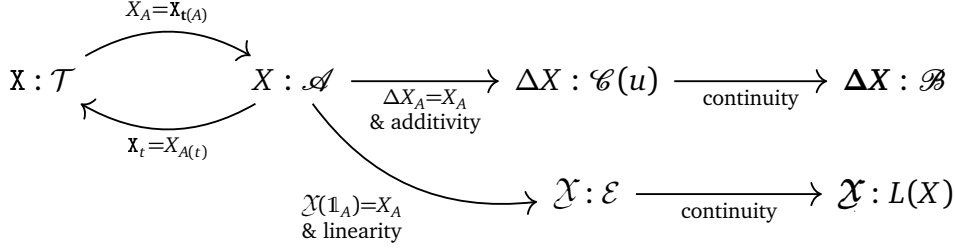


Figure 9: The basic process X (or X) and all its future extensions (and on which collections they are defined).

at the beginning of Chapter 2. Even though the details of these figures may appear a tad cryptic for now, we put them here for future reference. Whenever the reader has a doubt about which concept comes before which, they may hopefully find an answer here.

Three additional structures are added to the basic axioms on \mathcal{T} . They are described in Chapter 1 and are put to use in the subsequent chapters.

- ◊ The first one is a measure m on $\mathcal{B} = \sigma(\mathcal{A})$ with a particular midpoint (or bisection) property that enables us to construct the key notion of *geodesic flow*. The initial motivation is that for a process $X = \{X_A : A \in \mathcal{A}\}$ and a geodesic flow $\phi : [0, 1] \rightarrow \mathcal{A}$ (or on the bigger class $\mathcal{A}(u)$ made of finite unions of elements of \mathcal{A}), the *projection* $X^\phi = \{X_{\phi(t)} : t \in [0, 1]\}$ becomes a one-dimensional process for which intuition may be keener. In particular, the measure m is used in Chapter 2 to define several notions of stationarity and geodesic flows are used to prove corresponding representation theorems (Theorems 2.5.30 and 2.5.31).

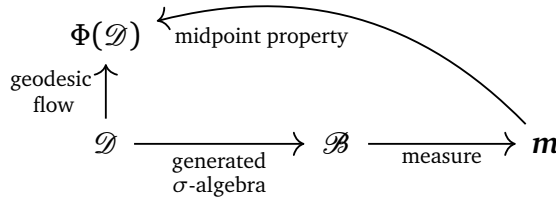


Figure 10: Geodesic \mathcal{D} -flows $\Phi(\mathcal{D})$ (usually for $\mathcal{D} = \mathcal{A}$ or $\mathcal{A}(u)$), a crucial bridge with the one-dimensional theory.

Flows are also used in the first half of Chapter 3 to define the space $\mathcal{D}^\Phi(\mathcal{A})$ of ‘càdlàg generalized maps’ in which quite a few processes of interest live.

- ◊ Finally tackling the word ‘Regularity’ of this thesis’s title, a metric $d_{\mathcal{T}}$ is introduced on \mathcal{T} , or equivalently a metric $d_{\mathcal{A}}$ on \mathcal{A} (see Figure 11). This allows us to discuss more quantitative regularity considerations like Hölder exponents in the second half of Chapter 3. Bounds on the pointwise Hölder regularity are given for the process $Y = \{Y_A : A \in \mathcal{A}\}$ defined by

$$\forall A \in \mathcal{A}, \quad Y_A = \mathcal{X}(f \mathbb{1}_A) = \int_A f dX$$

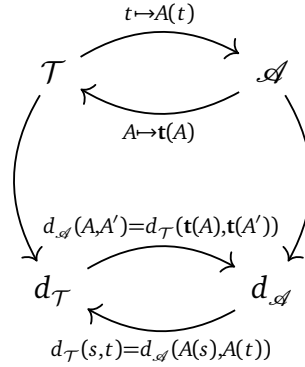


Figure 11: Correspondence between the metric spaces $(\mathcal{T}, d_{\mathcal{T}})$ and $(\mathcal{A}, d_{\mathcal{A}})$.

where $f : \mathcal{T} \rightarrow \mathbf{R}$ is a deterministic function and X is a generalized Lévy process (Theorems 3.5.2 and 3.5.12). What is observed especially is how the regularities of f and X interact and blend together to give that of Y .

- ◊ Among the ‘couple of simple axioms’ that \mathcal{T} is endowed with lies the existence of a countable subset, dense for some specific topology, that enables us to approximate the elements of \mathcal{T} . Equivalently, the elements of \mathcal{A} are approximated by a sequence of finite ‘discretizations’ $(\mathcal{A}_n)_{n \in \mathbf{N}}$ and the number of parameters required to describe such an approximation yields a notion of dimension for \mathcal{A} . Having a finite dimension has a lot of convenient consequences that will be used throughout the thesis, even if infinite dimensional examples will be considered as well.

$$\mathcal{A} \xrightarrow{\text{discretization}} \mathcal{A}_n \xrightarrow{\text{number of parameters}} \dim \mathcal{A}$$

Figure 12: The dimension of \mathcal{A} translates how well it is approximated by its discretization \mathcal{A}_n .

Last but not least, even though I will start employing the personal pronoun ‘we’, all the work that is not explicitly stated as others’ is my own²¹ or well-known general knowledge. We wish you a nice reading!

²¹As far as I know!

Notations and conventions

The notations being consistent throughout the thesis, we factorized them here. First are those known to every mathematician, but for which we felt their corresponding notations are not truly universal. Then are given the notations specific to our field together with the first time they appear in the thesis (together with the definition).

\mathbf{N}, \mathbf{N}^*	Set of non-negative (resp. positive) integers
$\mathbf{Z}, \mathbf{Q}, \mathbf{R}, \mathbf{C}$	Usual number sets
\mathbf{R}_+	Set of non-negative real numbers
$(\cdot)_+, (\cdot)_-$	Respectively positive and negative part of a real number
$\#$	Cardinal of a set
\subseteq, \subset	Respectively large and strict inclusions of sets
\sqcup	Disjoint union of sets
\simeq	Bijection between two sets
$\rightarrow, \hookrightarrow$	Surjective (resp. injective) mapping
\otimes	Product of σ -algebras or measures
$\varphi_*\mu$	Pushforward of a measure μ by a map φ
$\hat{\mu}(\xi)$	Fourier transform of a probability measure μ
$\mathcal{B}(\mathbf{R}^k)$	Borel σ -algebra of \mathbf{R}^k
$\mathcal{C}^0(E; F)$	Set of all continuous maps from E to F
o, O	Landau's 'small o' and 'big O'
\max, \vee (resp. \min, \wedge)	Maximum (resp. minimum) with respect to a partial order
$\mathfrak{S}(E)$	Group of all permutations of E
$(\Omega, \mathcal{F}, \mathbf{P})$	Complete probability space
$X \sim \mu$	Random variable X has distribution μ
$X \stackrel{\text{law}}{=} Y$	Random variables X and Y have the same distribution
$X \stackrel{\text{fdd}}{=} Y$	Processes X and Y have the same finite-dimensional distributions
\perp, \perp_Z	Independence between random variables (conditionally on Z)
$\mathbf{N}(m, \sigma^2)$	Gaussian distribution of mean m and variance σ^2
$\text{Poi}(\lambda)$	Poisson distribution of intensity λ
$\text{Uni}(0, 1)$	Uniform distribution on $[0, 1]$

$\mathcal{A}, \mathcal{A}^*, \mathcal{A}_n, A_n(t)$	Definition 1.2.1	ν_j	Equation (3.5.9)
$\mathcal{A}(u), \mathcal{A}_n(u), \mathcal{C}(u), \mathcal{D}(u), \dots$	Definition 1.2.34	$\Pi(h)$	Equation (3.3.12)
$\text{Aut}(\mathcal{T})$	Definition 2.5.24	$\Pi_{j L}$	Equation (3.5.8)
$\alpha_h(A)$	Equation (3.4.1)	$\bar{R}_f(A), \underline{R}_f(A)$	Equation (3.5.7)
$\alpha_{h,\mathcal{C}}(A)$	Equation (3.4.3)	$\bar{R}'_f(A), \underline{R}'_f(A)$	Equation (3.5.30)
$\alpha_{h,d_{\mathcal{T}}}(A)$	Equation (3.4.7)	ρ_X	Equation (2.2.3)
\mathcal{B}	Section 1.2.1	\mathbf{t}	Definition 1.2.1
\mathcal{B}_{loc}	Section 2.2.2	$(\mathcal{T}, \preccurlyeq, \mathcal{D})$	Definition 1.2.3
$\mathcal{B}_m, \mathcal{B}_\mu$	Section 1.3.1	\mathcal{T}_n	Lemma 1.3.16
$\mathcal{B}(X)$	Remark 2.2.11	\mathcal{T}_n	Equation (3.2.3)
$B_{\mathcal{A}}(A, \rho), B_{\mathcal{T}}(t, \rho)$	Section 1.4.1	$\mathcal{V}(A, \rho)$	Definition 1.4.6
$B_{\mathcal{C}}(C, \rho)$	Definition 1.4.4	$\mathcal{V}_n(A, \rho)$	Equation (1.4.12)
β	Equation (3.5.5)	$\bar{V}_n(A, \rho), \underline{V}_n(A, \rho)$	Equation (1.4.12)
$\mathcal{C}, \mathcal{C}^{(k)}$	Definition 1.2.35	$VC(\mathcal{D})$	Definition 1.5.8
$\mathcal{C}^\ell, \mathcal{C}^\ell(\mathcal{A}'), \mathcal{C}^\ell(\mathcal{A}_n), C_n(t)$	Definition 1.2.38	$v(\phi)$	Definition 1.3.5
$C^\Phi(\mathcal{A})$	Equation (3.2.4)	$\emptyset', 0_{\mathcal{T}}$	Equation (1.2.1)
Γ_j	Equation (3.4.21)	\mathcal{X}	Theorem 2.2.10
$D^\Phi(\mathcal{A})$	Equation (3.3.9)	$\chi, \llbracket \cdot \rrbracket$	Definition 2.3.4
$d_{\mathcal{A}}$	Definition 1.4.1	$\stackrel{(c)}{=}$	Definition 2.5.4
$d_{\mathcal{C}}$	Definition 1.4.4	$\ \cdot\ _\phi$	Equation (2.2.2)
$d_{\mathcal{D}(0,1)}$	Equation (3.3.1)		
d_Φ	Equation (3.3.4)		
$d_{\mathcal{H}}$	Equation (1.4.2)		
$d_{\mathcal{T}}$	Equation (1.4.1)		
d_m	Equation (1.3.2)		
$[D, D']$	Equation (1.3.1)		
$\dim \mathcal{A}$	Definition 1.5.3		
$\Delta h, \Delta X$	Proposition 1.2.44		
ΔX	Theorem 2.2.8		
\mathfrak{s}	Definition 1.4.6		
\mathcal{E}	Definition 1.2.40		
$E_{j L}^\delta$	Equation (3.4.20)		
$E_{ L}^\delta$	Equation (3.4.22)		
$\Phi(\mathcal{D}), \Phi_{\max}(\mathcal{D})$	Definition 1.3.5		
$g_n(A)$	Definition 1.2.1		
$\mathfrak{h}, \mathcal{X}$	Proposition 1.2.44		
$\mathcal{H}(\varepsilon)$	Definition 1.5.11		
$L^\gamma(b, \sigma^2, \nu)$	Section 2.4.2		
L^Φ	Definition 2.2.4		
$L(X), L_{\text{loc}}(X)$	Definition 2.2.9		
$L^p(\mathbf{m})$	Section 1.3.3		
$L^p(\Omega)$	Section 2.2.1		
$L_{f,\alpha}(A), L_{f,\alpha}(A, \rho), L_{f,\alpha}^{\mathbb{C}}(A, \rho)$	Equation (3.5.6)		
$L'_{f,\alpha}(A), L'_{f,\alpha}(A, \rho), L'^{\mathbb{C}}_{f,\alpha}(A, \rho)$	Equation (3.5.29)		
\mathbf{m}	Definition 1.3.1		
$\text{Mon}(\mathcal{T})$	Definition 2.5.24		
$\mathcal{M}(\mathcal{T})$	Section 1.3.1		

Contents

1	Set-indexed framework for generalized processes	1
1.1	Introduction	1
1.2	Indexing collections as partially ordered sets	2
1.3	Indexing collections as measured spaces	22
1.4	Indexing collections as metric spaces	32
1.5	Indexing collections of finite dimension	40
1.6	'Revamped' vs 'classical' theory	43
1.7	Perspectives	51
1.8	Perspectives (français)	52
2	Distributional properties of generalized processes	53
2.1	Introduction	54
2.2	Extensions of set-indexed processes	54
2.3	Processes with independent increments	59
2.4	Interlude: set-indexed Lévy processes	77
2.5	Processes with stationary increments	81
2.6	Perspectives	103
2.7	Perspectives (français)	105
3	Sample path properties of generalized processes	107
3.1	Introduction	108
3.2	The continuous space $\mathcal{C}^\Phi(\mathcal{A})$	108
3.3	The càdlàg space $\mathcal{D}^\Phi(\mathcal{A})$	116
3.4	Hölder exponents for generalized processes	124
3.5	Hölder regularity of the primitive with respect to a siLévy	132
3.6	Perspectives	147
3.7	Perspectives (français)	149
	Bibliography	151

Set-indexed framework for generalized processes

1

Contents

1.1	Introduction	1
1.2	Indexing collections as partially ordered sets	2
1.2.1	The class \mathcal{A} of indices	2
1.2.2	A categorical point of view: to fundamental examples and beyond	4
1.2.3	\mathcal{A} 's friends	18
1.2.4	Increment map and linear functional	20
1.3	Indexing collections as measured spaces	22
1.3.1	Measure m on $\mathcal{B} = \sigma(\mathcal{A})$	22
1.3.2	Geodesic flows	23
1.3.3	Topological consequences	29
1.4	Indexing collections as metric spaces	32
1.4.1	Metric $d_{\mathcal{A}}$ on \mathcal{A}	32
1.4.2	Metric $d_{\mathcal{C}}$ on \mathcal{C}	34
1.4.3	Divergence δ and vicinities	35
1.4.4	Topological consequence	39
1.5	Indexing collections of finite dimension	40
1.5.1	A poset approach	40
1.5.2	Entropy bound	42
1.6	'Revamped' vs 'classical' theory	43
1.6.1	A taste of the classical setting	44
1.6.2	SHAPE condition	45
1.6.3	TIP assumption	47
1.6.4	'Revamped' vs 'classical' flows	50
1.7	Perspectives	51
1.8	Perspectives (français)	52

1.1 Introduction

We develop here a theory to study stochastic processes $X = \{X_t : t \in \mathcal{T}\}$ indexed by a very general 'time' space \mathcal{T} . By 'time', we only mean that some elements may happen 'before' or 'after' others,

i.e. there will be a partial order \preceq as the basic structure on \mathcal{T} . Keeping Birkhoff's theorem [19, Chapter III, Corollary 2] in the back of our mind, we know that it is often possible to represent the elements of \mathcal{T} as sets and the partial order \preceq as the inclusion relation \subseteq . Hence we 'put the cart before the horse' and first expose in Section 1.2 this theory of generalized processes as the study of set-indexed processes $X = \{X_A : A \in \mathcal{A}\}$ where \mathcal{A} is a collection of sets instead of points. The reason we do so is twofold:

1. A successful theory for set-indexed processes has already been developed by Ivanoff and Merzbach in [46, 47, 71] together with many contributors. We turn to the previous references for an exhaustive bibliography on the matter. Even though we will not entirely follow their setting, we still borrow their notations in the hope that the already accustomed reader will find it more convenient this way.
2. Even though we will make a case in Section 1.2.2 for the equivalence between studying generalized and set-indexed processes (especially with Theorem 1.2.5), being indexed by sets has advantages of its own. Some concepts are indeed more naturally formulated in the set-indexed setting and we intend to take full advantage of it.

After exposing the relation between generalized and set-indexed processes in Section 1.2, we progressively introduce additional structures on the indexing set: a measure m in Section 1.3, a metric $d_{\mathcal{A}}$ in Section 1.4 and some finite-dimensional assumptions in Section 1.5. This presentation of the theory allows as much 'modularity' as possible. We invite the reader to see each new structure as some 'package' that they may download — and why not tweek at will — when necessary. We conclude in Section 1.6 by comparing this theory with Ivanoff and Merzbach's.

In the sequel, \mathcal{T} will always denote a 'generic' non-empty set. As mentioned above, it will progressively be endowed with additional structures. Specific choices for \mathcal{T} will be considered in examples.

1.2 Indexing collections as partially ordered sets

1.2.1 The class \mathcal{A} of indices

The following definition is inspired from [42, Definition 2.1], which is itself a careful selection of the required properties of [47, Definition 1.1.1]. We emphasize the fact that even though we borrow notations from Ivanoff and Merzbach's theory, they will not always designate the exact same objects. The differences will be highlighted and discussed in Section 1.6.

DEFINITION 1.2.1 (Indexing collection). *A class $\mathcal{A} \subseteq \mathcal{P}(\mathcal{T})$ of subsets of \mathcal{T} is an indexing collection on \mathcal{T} if the following properties hold:*

1. (Countably complete meet-semilattice). $\emptyset \in \mathcal{A}$ and the collection $\mathcal{A}^* = \mathcal{A} \setminus \{\emptyset\}$ is closed under countable intersections.
2. (Separability from above). There exists a non-decreasing sequence of finite subcollections $\mathcal{A}_n = \{A_1^n, \dots, A_{k_n}^n\} \subseteq \mathcal{A}$ ($n \in \mathbf{N}$) closed under intersections such that given the functions $g_n : \mathcal{A} \rightarrow \mathcal{A}_n \cup \{\mathcal{T}\}$ defined by

$$\forall A \in \mathcal{A}, \quad g_n(A) = \bigcap_{\substack{A' \in \mathcal{A}_n \cup \{\mathcal{T}\}: \\ A \subseteq A'}} A',$$

the elements of \mathcal{A} may be approximated as follows: for all $A \in \mathcal{A}$, $A = \bigcap_{n \in \mathbf{N}} g_n(A)$.

3. (TIP assumption). The map

$$\begin{aligned} \mathcal{T} &\longrightarrow \mathcal{A}^* \\ t &\longmapsto A(t) = \bigcap_{n \in \mathbb{N}} A_n(t), \end{aligned}$$

where $A_n(t) = \bigcap_{\substack{A \in \mathcal{A}_n \cup \{\mathcal{T}\}: \\ t \in A}} A$, is one-to-one.

Its inverse map is denoted by $\mathbf{t} : \mathcal{A}^* \rightarrow \mathcal{T}$ and called the TIP bijection. Moreover, for all $A \in \mathcal{A}^*$, the point $\mathbf{t}(A) \in \mathcal{T}$ is called the tip of A .

4. (SHAPE condition). For any $k \in \mathbb{N}^*$ and $A, A_1, \dots, A_k \in \mathcal{A}$, if $A \subseteq \bigcup_{1 \leq i \leq k} A_i$, then $A \subseteq A_j$ for some $j \in \llbracket 1, k \rrbracket$.

In the sequel, unless otherwise specified, \mathcal{A} will stand for such an indexing collection and $\mathcal{B} = \sigma(\mathcal{A})$ for the σ -algebra it generates. Even if Definition 1.2.1 might appear daunting at first, we ask the reader to bear with us until Section 1.2.2 where another — simpler — point of view is given alongside examples.

REMARK 1.2.2. What we mean by ‘indexing collection’ differs from parts of the literature. However, each time we quote a result from the literature, we will argue on why the conclusions still hold in our case. In order to make a difference, we refer to the setting presented by Ivanoff and Merzbach in [47] as being the classical setting.

For instance, a classical indexing collection [47, Definition 1.1.1] is usually supposed to be closed under arbitrary intersections, but this property also holds in our framework. Indeed, consider a subcollection $\mathcal{A}' \subseteq \mathcal{A}$. Then, separability from above tells that $\bigcap_{A \in \mathcal{A}'} A = \bigcap_{n \in \mathbb{N}} \bigcap_{A \in \mathcal{A}' \cap \mathcal{A}_n} A$ so it still belongs to \mathcal{A} by stability under countable intersections. Likewise, for all $t \in \mathcal{T}$, $\bigcap_{A \in \mathcal{A}': t \in A} A = \bigcap_{n \in \mathbb{N}} \bigcap_{A \in \mathcal{A}_n: t \in A} A = A(t)$, which links back to the usual meaning of $A(t)$ in the literature.

In particular, \mathcal{A} has a global minimum:

$$\emptyset' = A(0_{\mathcal{T}}) = \bigcap_{A \in \mathcal{A}: A \neq \emptyset} A. \quad (1.2.1)$$

where the point $0_{\mathcal{T}} \in \mathcal{T}$ will also be understood as the global minimum for \mathcal{T} as soon as it is endowed with a partial order \preceq in Definition 1.2.3. The notation will be consistent with the usual 0 whenever \mathcal{T} has one. Without loss of generality, we suppose that both \emptyset and \emptyset' belong to \mathcal{A}_n for all $n \in \mathbb{N}$.

Before proceeding any further, let us share some preliminary thoughts about each of the conditions appearing in Definition 1.2.1.

1. Having $\mathcal{A}^* = \mathcal{A} \setminus \{\emptyset\}$ closed under countable intersections implies that \mathcal{A} as well, which is a condition required in the classical setting. Remark that it is a nice setting for measure-theoretic constructions to apply.

But perhaps more importantly, it is a necessary condition if one wants to look at set-indexed processes $X = \{X_A : A \in \mathcal{A}\}$ which are ‘continuous from above’ in some sense, *i.e.* for any non-increasing sequence $(A_k)_{k \in \mathbb{N}}$ in \mathcal{A} , $X_{A_k} \rightarrow X_{\bigcap_{i \in \mathbb{N}} A_i}$ as $k \rightarrow \infty$. This property is a natural generalization of being right continuous for \mathbf{R}_+ -indexed processes, which is a condition required more often than not. In general, one often imposes the stronger condition of being *càdlàg*, that is being right continuous with left limits. This aspect will be addressed in a different way from the classical setting (see Sections 1.3.2 and 3.3).

Another reason to require stability under intersections would be that filtrations play a crucial role while studying processes such as martingales or Markov processes. In order to impose intuitive 'time consistent' relations between the σ -algebras in a set-indexed filtration, stability under intersections is necessary. For more details, we refer to [47].

2. Another key element arising in the study of \mathbf{R}_+ -indexed càdlàg processes is the use of dyadics. They are useful to get results in the continuous case from their discrete alter egos.

In the second assumption, the class \mathcal{A}_n indeed plays a role similar to the dyadics of order n in that endeavour. In the classical setting, one usually imposes some topological structure on \mathcal{T} so that A lies in the interior of $g_n(A)$ for all $n \in \mathbf{N}$ instead. This implies a 'separability *strictly from above*' property, but we chose against it here since \mathcal{A} will be endowed with a metric in Section 1.4 so that there is no competition with another topology.

3. The TIP assumption has also been introduced in [47, Assumption 2.4.2] and draws a clear correspondence between general processes $\{X_t : t \in \mathcal{T}\}$ and set-indexed processes $\{X_A : A \in \mathcal{A}\}$ such that $X_\emptyset = 0$ through the relation $X_t = X_{A(t)}$ for all $t \in \mathcal{T}$.

This bijection is the key element that allows the correspondence given by Theorem 1.2.5 to hold.

As a quick sidenote, one may remark that $g_n(A(t)) = A_n(t)$ for all $n \in \mathbf{N}$ and $t \in \mathcal{T}$.

4. The SHAPE condition has been first introduced in [47, Assumption 1.1.5] as a sufficient condition to ensure the existence of *increment maps*. In lattice-theoretic vocabulary, this condition is known as *join irreducibility*.

1.2.2 A categorical point of view: to fundamental examples and beyond

In the classical setting so far, few examples of indexing collections have been given explicitly, which might lead the reader new to this field to think that this theory might be not so rich after all. In this section, we strive to make this opinion sink into oblivion by giving a wealth of examples and indicating how to build new ones from old ones. Those examples will be continuously quoted throughout this thesis in order to illustrate the concepts at play.

In order to provide a wider view and a better grasp on what may or may not be considered as part of the theory, we chose a 'categorical' point of view where we derive general constructions, from which examples become mere applications. However, we neither expect the reader to know anything about category theory nor particularly strive to write in a 'category-friendly' manner. This approach will just show that indexing collections are nice structures by themselves.

A slight disclaimer beforehand: we do not directly study indexing collections, but another structure — called *indexing semilattice* — which turns out to be equivalent. We feel that some concepts we develop are more natural and better understood that way. This has the downside of introducing vocabulary specific to indexing semilattices. But outside of that section, we will stick as much as possible to the usual vocabulary from set-indexed theory since it has perks of its own and the reader might already be more familiar with the set-indexed theory.

Definitions and fundamental correspondence

Let us start by giving a few definitions and simple examples. First is the *indexing semilattice* (Definition 1.2.3) whose couple of simple axioms will turn out to be equivalent to the apparently more complex indexing collection (Theorem 1.2.5).

DEFINITION 1.2.3 (Indexing semilattice and sub-structures). An indexing semilattice is a triplet $(\mathcal{T}, \preceq, \mathcal{D})$ such that:

1. (Countably complete meet-semilattice). (\mathcal{T}, \preceq) is a non-empty poset (i.e. partially ordered set) such that any countable subset $\{t_i : i \in \mathbf{N}\} \subseteq \mathcal{T}$ admits a minimum $\bigwedge_{i \in \mathbf{N}} t_i \in \mathcal{T}$.
2. (Separability from above). The set \mathcal{D} is included in \mathcal{T} , at most countable, and such that for any $t \in \mathcal{T}$, there exists a non-increasing sequence $(t_n)_{n \in \mathbf{N}}$ in \mathcal{D} such that $t = \bigwedge_{n \in \mathbf{N}} t_n$.

An indexing sub-semilattice of \mathcal{T} is a subset $\mathcal{T}' \subseteq \mathcal{T}$ such that $(\mathcal{T}', \preceq, \mathcal{D} \cap \mathcal{T}')$ is an indexing semilattice.

An ideal of \mathcal{T} is an indexing sub-semilattice \mathcal{T}' of \mathcal{T} such that for all $(s, t) \in \mathcal{T} \times \mathcal{T}'$, $s \wedge t \in \mathcal{T}'$.

DEFINITION 1.2.4 (Morphism of indexing semilattices). Let $(\mathcal{T}, \preceq, \mathcal{D})$ and $(\mathcal{T}', \preceq', \mathcal{D}')$ be two indexing semilattices. A map $\varphi : \mathcal{T} \rightarrow \mathcal{T}'$ is a morphism (of indexing semilattices) if the following properties hold:

1. (Compatibility of semilattices). For all countable subset $\{t_i : i \in \mathbf{N}\}$ of \mathcal{T} ,

$$\varphi\left(\bigwedge_{i \in \mathbf{N}} t_i\right) = \bigwedge'_{i \in \mathbf{N}} \varphi(t_i)$$

where \bigwedge (resp. \bigwedge') is the minimum in \mathcal{T} (resp. \mathcal{T}').

2. (Compatibility of dense subsets). $\varphi(\mathcal{D}) \subseteq \mathcal{D}'$.

Moreover, if φ is bijective and φ^{-1} is also such a morphism, then φ is called an isomorphism (of indexing semilattices).

One may directly realize that $(\mathcal{A}, \subseteq, \bigcup_{n \in \mathbf{N}} \mathcal{A}_n)$ itself is an indexing semilattice. But one may actually say a bit more than that and establish a correspondence.

THEOREM 1.2.5 (Correspondence $(\mathcal{T}, \mathcal{A}) \leftrightarrow (\mathcal{T}, \preceq, \mathcal{D})$). Any indexing collection \mathcal{A} induces an indexing semilattice in the following way: the order relation \preceq is given by

$$\forall s, t \in \mathcal{T}, \quad s \preceq t \iff A(s) \subseteq A(t) \tag{1.2.2}$$

and $\mathcal{D} = \{\mathbf{t}(A) : A \in \bigcup_{n \in \mathbf{N}} \mathcal{A}_n\}$.

Conversely, any indexing semilattice $(\mathcal{T}, \preceq, \mathcal{D})$ induces an indexing collection $\mathcal{A} = \{A(t) : t \in \mathcal{T}\} \cup \{\emptyset\}$ on \mathcal{T} where $A(t) = \{s \in \mathcal{T} : s \preceq t\}$ for all $t \in \mathcal{T}$.

Proof. The following is plain abstract nonsense.

Suppose that \mathcal{A} is an indexing collection on \mathcal{T} and define \preceq by (1.2.2). This is an order relation: it is trivially reflexive, antisymmetric by the TIP bijection and transitive since \subseteq also is. Still thanks to the TIP bijection \mathbf{t}, \preceq verifies the first property of Definition 1.2.3. Indeed, if $\{t_i : i \in \mathbf{N}\}$ is a countable subset of \mathcal{T} , we directly have $\mathbf{t}\left(\bigcap_{i \in \mathbf{N}} A(t_i)\right) = \mathbf{t}\left(A\left(\bigwedge_{i \in \mathbf{N}} t_i\right)\right) = \bigwedge_{i \in \mathbf{N}} t_i$. As for the separability from above, the subset $\mathcal{D} = \mathbf{t}\left(\bigcup_{n \in \mathbf{N}} \mathcal{A}_n\right)$ does the job.

Conversely, consider an indexing semilattice $(\mathcal{T}, \preceq, \mathcal{D})$ and define \mathcal{A} as described in Theorem 1.2.5. Then $\emptyset \in \mathcal{A}$ by definition and the stability by intersections of $\mathcal{A}^* = \mathcal{A} \setminus \{\emptyset\}$ is trivial. For the separability from above and the TIP assumption, since \mathcal{T} is closed under minimum, without loss of generality, we may write $\mathcal{D} = \bigcup_{n \in \mathbf{N}} \mathcal{D}_n$ where $(\mathcal{D}_n)_{n \in \mathbf{N}}$ is a non-decreasing sequence of finite subsets of \mathcal{D} closed under minimum. The subcollections defined for all $n \in \mathbf{N}$ by $\mathcal{A}_n = \{A(t) : t \in \mathcal{D}_n\} \cup \{\emptyset\}$ then do the job. As for the SHAPE condition, consider $k \in \mathbf{N}$ and $A(t), A(t_1), \dots, A(t_k) \in \mathcal{A}$ such that $A(t) \subseteq \bigcup_{1 \leq i \leq k} A(t_i)$ (the case with the emptyset is trivial). Since $t \in A(t)$, there exists $j \in \llbracket 1, k \rrbracket$ such that $t \in A(t_j)$. Hence $A(t) \subseteq A(t_j)$. \square

In the sequel, unless otherwise specified, $(\mathcal{T}, \preccurlyeq, \mathcal{D})$ will stand for the indexing semilattice induced by the indexing collection \mathcal{A} considered throughout this thesis.

First, a trivial example.

EXAMPLE 1.2.6 (Indexing collection on $\mathcal{T} = \{0_{\mathcal{T}}\}$). $\mathcal{A} = \{\emptyset, \emptyset'\}$ is an indexing collection on $\mathcal{T} = \{0_{\mathcal{T}}\}$ ($= \emptyset'$) called the trivial indexing collection.

Now, let us move on to a first interesting — and fundamental! — example, which is the link between the usual theory of one-dimensional processes and the theory of generalized processes.

EXAMPLE 1.2.7 (Indexing collection on $\mathcal{T} = \mathbf{R}_+$). Since it is the first one, we present the indexing collection \mathcal{A} on \mathbf{R}_+ from both perspectives, i.e. the one of Definition 1.2.1 and the one of Theorem 1.2.5.

Let us begin with the latter: \mathbf{R}_+ is endowed with a natural total order \leq , directly making it into a complete meet-semilattice. Moreover, if we denote for all $n \in \mathbf{N}$ the set $\mathcal{D}_n = \{k2^{-n} : 0 \leq k \leq n2^n\}$ of dyadics of order n in $[0, n]$, any real number $t \in \mathbf{R}_+$ may be approximated by a non-increasing sequence in $\mathcal{D} = \bigcup_{n \in \mathbf{N}} \mathcal{D}_n$, which is countable.

According to Theorem 1.2.5, this yields an indexing collection \mathcal{A} that could directly be described as follows:

$$\mathcal{A} = \{[0, t] : t \in \mathbf{R}_+\} \cup \{\emptyset\} \quad \text{and} \quad \forall n \in \mathbf{N}, \quad \mathcal{A}_n = \{[0, t] : t \in \mathcal{D}_n\} \cup \{\emptyset\}.$$

EXAMPLE 1.2.8 (Indexing collection on $\mathcal{T} = \mathbf{R}$). In that case, an order relation \preccurlyeq on \mathbf{R} may be defined as follows:

$$\forall s, t \in \mathbf{R}, \quad s \preccurlyeq t \iff |s| \leq |t| \quad \text{and} \quad s \times t \geq 0$$

and $\mathcal{D} = \bigcup_{n \in \mathbf{N}} \mathcal{D}_n$ where for all $n \in \mathbf{N}$, \mathcal{D}_n is the set of dyadic of order n in $[-n, n]$. The resulting triplet $(\mathbf{R}, \preccurlyeq, \mathcal{D})$ easily turns out to be an indexing semilattice and the corresponding indexing collection is $\mathcal{A} = \{[0, t] : t \in \mathbf{R}_+\} \cup \{[t, 0] : t \in \mathbf{R}_-\} \cup \{\emptyset\}$. The reason why one does not choose the usual total order on \mathbf{R} is that it would not yield an indexing semilattice structure. Indeed, \mathbf{R} has no global minimum for its usual order, which would contradict $\emptyset \neq \emptyset'$ (see (1.2.1)).

This construction may seem a bit arbitrary at first, but we will show that it is part of a more general construction scheme. Namely, \mathcal{A} may be seen as the ‘gluing’ of two copies of the indexing collection on \mathbf{R}_+ (Proposition 1.2.26).

In the next few sections, we provide several methods to build new indexing collections out of old ones. The exposition goes from simple constructions to more involved ones. At the end, the table page 17 summarizes and compares all of them.

Product of indexing collections

Since the one-dimensional case has been shown to be a particular case of indexing collection in Example 1.2.7, let us rightly follow along with a construction yielding the multiparameter case. By virtue of Theorem 1.2.5, we will freely switch from indexing collections to indexing semilattices depending on which feels easier to explain.

PROPOSITION 1.2.9 (Product of indexing collections). Let $\{(\mathcal{T}_i, \mathcal{A}_i) : 1 \leq i \leq k\}$ be a finite sequence of spaces each endowed with an indexing collection. Then $\prod_{1 \leq i \leq k} \mathcal{A}_i = \{A_1 \times \dots \times A_k : \forall i \in \llbracket 1, k \rrbracket, A_i \in \mathcal{A}_i\}$ is an indexing collection on $\prod_{1 \leq i \leq k} \mathcal{T}_i$.

Let $\{(\mathcal{T}_i, \mathcal{A}_i) : i \in \mathbf{N}\}$ be an infinite sequence of spaces each endowed with an indexing collection. With the additional assumption that $\mathcal{T}_i \in \mathcal{A}_i$ for all $i > k$ for some $k \in \mathbf{N}$, the previous result still holds for the countable product, i.e. $\prod_{i \in \mathbf{N}} \mathcal{A}_i$ is an indexing collection on $\prod_{i \in \mathbf{N}} \mathcal{T}_i$.

Proof. Consider the finite case first and denote by $(\mathcal{T}_i, \preceq_i, \mathcal{D}_i)$ the indexing semilattice induced by \mathcal{A}_i thanks to Theorem 1.2.5. Define the component-wise order \preceq on $\prod_{1 \leq i \leq k} \mathcal{T}_i$ by

$$\forall s, t \in \prod_{1 \leq i \leq k} \mathcal{T}_i, \quad s \preceq t \iff \forall i \in \llbracket 1, k \rrbracket, s_i \preceq_i t_i$$

and denote $\mathcal{D} = \prod_{1 \leq i \leq k} \mathcal{D}_i$. It is then straightforward to check that $(\prod_{1 \leq i \leq k} \mathcal{T}_i, \preceq, \mathcal{D})$ is an indexing semilattice whose indexing collection corresponds to $\prod_{1 \leq i \leq k} \mathcal{A}_i$.

For the infinite case, the order relation still is the component-wise order, but the dense subset is defined a bit differently: $\mathcal{D} = \bigcup_{n \geq k} \tilde{\mathcal{D}}_n$ where k is the one from the statement of Proposition 1.2.9 and for all $n \geq k$, $\tilde{\mathcal{D}}_n = \mathcal{D}_1 \times \dots \times \mathcal{D}_n \times \{\mathbf{t}_{n+1}(\mathcal{T}_{n+1})\} \times \{\mathbf{t}_{n+2}(\mathcal{T}_{n+2})\} \times \dots$ where $\mathbf{t}_i : (\mathcal{A}_i)^* \rightarrow \mathcal{T}_i$ is the TIP bijection given by the indexing collection \mathcal{A}_i . The result follows once more by Theorem 1.2.5. \square

EXAMPLE 1.2.10 (Indexing collection on $\mathcal{T} = \mathbf{R}_+^p$). For $p \in \mathbf{N}^*$, a direct application of Proposition 1.2.9 enables us to endow $\mathcal{T} = \mathbf{R}_+^p$ with an indexing collection on its own.

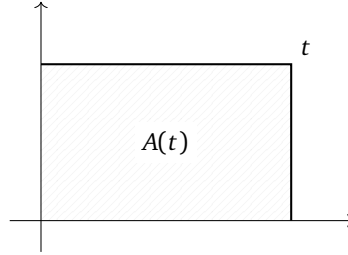


Figure 1.1: An element $A(t)$ (hatched) of the indexing collection on \mathbf{R}_+^2 .

This indexing collection is the same as the one given through Theorem 1.2.5 if one were to endow \mathbf{R}_+^p with the natural component-wise partial order and take the set of all p -tuples with dyadic coordinates as a dense subset. This is one of the core examples and whenever new hypotheses are considered, we made sure not to exclude this case.

Pushforward of an indexing collection

Here, we give some ways to transport an indexing collection structure from one space to another.

PROPOSITION 1.2.11 (Pushforward of an indexing semilattice). Let \mathcal{T}' be a set and $\varphi : \mathcal{T} \rightarrow \mathcal{T}'$ be a surjective map such that for all $s, t, u, v \in \mathcal{T}$ where $\varphi(s) = \varphi(u) \neq \varphi(t) = \varphi(v)$, we have $s \preceq t$ if and only if $u \preceq v$. Then, the binary relation \preceq' on \mathcal{T}' given for all $s', t' \in \mathcal{T}'$ by

$$\begin{aligned} s' \preceq' t' &\iff \forall (s, t) \in \varphi^{-1}(\{s'\}) \times \varphi^{-1}(\{t'\}), s \preceq t \\ &\iff \exists (s, t) \in \varphi^{-1}(\{s'\}) \times \varphi^{-1}(\{t'\}) : s \preceq t. \end{aligned}$$

is a well-defined partial order, $(\mathcal{T}', \preceq', \varphi(\mathcal{D}))$ is an indexing semilattice and φ is a morphism.

Proof. The condition on φ ensures that the equivalence defining \preceq' is correct. It readily follows that \preceq' is a partial order and that any element of \mathcal{T}' is approximated from above by elements of $\varphi(\mathcal{D})$. Only the stability under countable minimum is left. Let $\{t'_i : i \in \mathbf{N}\}$ be a countable subset of \mathcal{T}' and fix for all $i \in \mathbf{N}$ an element $t_i \in \mathcal{T}$ such that $\varphi(t_i) = t'_i$. Then, denote $t = \bigwedge_{i \in \mathbf{N}} t_i$ and $t' = \varphi(t)$. We know that for all $i \in \mathbf{N}$, $t \preceq t_i$ and thus $t' \preceq' t'_i$. If $s' = \varphi(s) \in \mathcal{T}$ is such that for all $i \in \mathbf{N}$, $s' \preceq' t'_i$, then we know that for all $i \in \mathbf{N}$, $s \preceq t_i$, thus $s \preceq t$ and $s' \preceq' t'$. Hence $t' = \bigwedge_{i \in \mathbf{N}} t'_i$. The result follows. \square

However, this general principle might not be so practical after all, so we specialized it to simpler cases. The proofs are direct applications of Proposition 1.2.11.

COROLLARY 1.2.12 (Projection of an indexing semilattice). *Suppose that \mathcal{T}' is a subset of \mathcal{T} and that $\varphi : \mathcal{T} \rightarrow \mathcal{T}'$ is a map which verifies the hypotheses of Proposition 1.2.11 and such that $\pi|_{\mathcal{T}'} = \text{id}_{\mathcal{T}'}$. Then $(\mathcal{T}', \preceq, \varphi(\mathcal{D}))$ is an indexing semilattice and φ is a morphism.*

Sketch of proof. By Proposition 1.2.11, we already know that $(\mathcal{T}', \preceq', \varphi(\mathcal{D}))$ is an indexing semilattice and φ is a morphism. Then, one just needs to exploit the fact that $\pi|_{\mathcal{T}'} = \text{id}_{\mathcal{T}'}$ to show that \preceq and \preceq' are equal. The result follows. \square

EXAMPLE 1.2.13 (\mathcal{A}_n is also an indexing collection). *Let $n \in \mathbf{N}$. Then \mathcal{A}_n is an indexing collection on $\mathbf{t}(\mathcal{A}_n)$ by Corollary 1.2.12 applied to the projection $\varphi = \mathbf{t} \circ A_n(\cdot)$. One aspect of our approach compared to the classical setting is that discrete indexing collections are still available at this level of generality. This may lead to some constructions expressing a given ‘continuous’ indexing collection \mathcal{A} as the ‘limit’ of the indexing collections \mathcal{A}_n as $n \rightarrow \infty$. We refer to the discussion in Section 1.7 for more ideas down this line.*

EXAMPLE 1.2.14 (Indexing collection on $\mathcal{T} = [0, 1]$). *$\mathcal{T} = [0, 1]$ may be endowed with an indexing collection from the one of \mathbf{R}_+ (Example 1.2.7) using Corollary 1.2.12 with the projection $\varphi : t \mapsto t \wedge 1$. It corresponds to the indexing collection $\mathcal{A} = \{[0, t] : t \in [0, 1]\} \cup \{\emptyset\}$.*

EXAMPLE 1.2.15 (Indexing collection for $\mathcal{T} = \mathbf{N}$). *$\mathcal{T} = \mathbf{N}$ may be endowed with an indexing collection from the one of \mathbf{R}_+ (Example 1.2.7) using Corollary 1.2.12 with the projection $\varphi : t \mapsto \lfloor t \rfloor$. It corresponds to the indexing collection $\mathcal{A} = \{\llbracket 0, t \rrbracket : t \in \mathbf{N}\} \cup \{\emptyset\}$.*

EXAMPLE 1.2.16 (More products). *Combining those previous examples with Proposition 1.2.9 enables us to create indexing collections on \mathbf{N}^p and the Hilbert cube $[0, 1]^{\mathbf{N}}$ for instance. But beware of the fact that even though \mathbf{R}_+ may be endowed with an indexing collection, $\mathbf{R}_+^{\mathbf{N}}$ cannot be, at least by a product procedure. Indeed, if there was such an indexing collection \mathcal{A} , then the sequence $(1 + \max_{A \in \mathcal{A}_n} \mathbf{t}(A)_n)_{n \in \mathbf{N}}$ cannot be approximated from above. This is the reason why there is an additional condition for the countable product of indexing collections in Proposition 1.2.9.*

COROLLARY 1.2.17 (Bijective pushforward of an indexing semilattice). *Let \mathcal{T}' be a set and $\varphi : \mathcal{T} \rightarrow \mathcal{T}'$ be a bijective map. Then, the binary relation \preceq' on \mathcal{T}' given for all $s', t' \in \mathcal{T}'$ by*

$$s' \preceq' t' \iff \varphi^{-1}(s') \preceq \varphi^{-1}(t')$$

is a well-defined partial order, $(\mathcal{T}', \preceq', \varphi(\mathcal{D}))$ is an indexing semilattice and φ is a morphism.

Sketch of proof. Since φ is bijective, the partial orders \preceq' defined in Proposition 1.2.11 and Corollary 1.2.17 are one and the same. The result follows. \square

EXAMPLE 1.2.18 (Indexing collection on $\mathcal{T} = [0, 1)$). $\mathcal{T} = [0, 1)$ may be endowed with an indexing collection from the one of \mathbf{R}_+ (Example 1.2.7) using Corollary 1.2.17 with the bijection $\varphi : t \mapsto t/(1+t)$. It corresponds to the indexing collection $\mathcal{A} = \{[0, t] : t \in [0, 1)\} \cup \{\emptyset\}$.

EXAMPLE 1.2.19 (Indexing collection on $\mathcal{T} = \mathbf{R}^* \cup \{\infty\}$). Compound Poisson processes indexed by \mathbf{R}_+ may be seen as an integrals against a $(\mathbf{R}_+ \times \mathbf{R}^*)$ -indexed Poisson process. This idea will later be exploited in the set-indexed setting (Example 2.4.6), but first requires to endow \mathbf{R}^* with an indexing collection. However, according to (1.2.1), indexing collections require a global minimum, which is why we consider $\mathbf{R}^* \cup \{\infty\}$ instead. For that, use Corollary 1.2.17 with the bijection

$$\begin{aligned} \varphi : \mathbf{R} &\longrightarrow \mathbf{R}^* \cup \{\infty\} \\ t &\longmapsto \begin{cases} 1/t & \text{if } t \neq 0, \\ \infty & \text{if } t = 0. \end{cases} \end{aligned}$$

This corresponds to the indexing collection

$$\mathcal{A} = \{[t, +\infty) \cup \{\infty\} : t \in \mathbf{R}_+^*\} \cup \{(-\infty, t] : t \in \mathbf{R}_-^*\} \cup \{\emptyset, \{\infty\}\}.$$

EXAMPLE 1.2.20 (Indexing collection on a graph). If $\varphi : \mathcal{T} \rightarrow \mathcal{T}'$ is a map on the indexing semilattice \mathcal{T} , then its graph $Gr(\varphi) = \{(t, \varphi(t)) : t \in \mathcal{T}\}$ may be endowed with an indexing collection by Corollary 1.2.17 with the bijection

$$\begin{aligned} \mathcal{T} &\longrightarrow Gr(\varphi) \\ t &\longmapsto (t, \varphi(t)). \end{aligned}$$

At this stage, this might not appear to amount to much. However, if φ is a map with an irregular behavior — e.g. the sample path of a stochastic process — this gives a way to endow rough surfaces with an indexing collection.

Gluing together indexing collections

Keeping on building new indexing collections from old ones, we introduce several ways to ‘glue together’ indexing collections. The main idea comes from the fact that if \mathcal{T} may be covered by a family $\{(\mathcal{T}_e, \preceq_e, \mathcal{D}_e) : e \in E\}$ of indexing semilattices in a consistent way — a bit like local maps of a manifold — then \mathcal{T} may be considered as an indexing semilattice itself.

In the probabilistic literature, gluing spaces is not a new idea, but is often used to define random metric spaces as in e.g. [7, 8, 26, 33, 89] and references therein. Our approach is slightly different since an indexing collection will end up being a space on which a stochastic process is studied, and not an object of study in itself.

We first recall a few fact about what a disjoint union is. If $\{T_e : e \in E\}$ is a collection of sets indexed by some set E , then the *disjoint union of the T_e 's* is defined by

$$\bigsqcup_{e \in E} T_e = \bigcup_{e \in E} \{(t, e) : t \in T_e\}. \quad (1.2.3)$$

Remark that for all $e \in E$, we may identify the set $\{(t, e) : t \in T_e\}$ with T_e itself, which is what we will usually do. However, do keep in mind that for $e \neq e'$ in E , even if $T_e \cap T_{e'}$ may be non-empty, it will be considered as such in $\bigsqcup_{e \in E} T_e$ precisely due to this definition. For instance, $\mathbf{R}_+ \sqcup \mathbf{R}_+$ is different from $\mathbf{R}_+ \cup \mathbf{R}_+$.

The simplest gluing is when \mathcal{T} can be realized as a disjoint union of indexing semilattices that can be ordered by means of an overarching discrete indexing semilattice E . This construction is in a stark contrast to the classical set-indexed theory where such a construction would not be allowed due to the ban on discrete structures.

PROPOSITION 1.2.21 (Disjoint union of indexing semilattices). *Let (E, \preceq_E, E) be an at most countable indexing semilattice and $\{(\mathcal{T}_e, \preceq_e, \mathcal{D}_e) : e \in E\}$ a family of indexing semilattices indexed by E . Define the order relation \preceq on the disjoint union $\bigsqcup_{e \in E} \mathcal{T}_e$ by*

$$\forall e, e' \in E, s \in \mathcal{T}_e, t \in \mathcal{T}_{e'}, \quad s \preceq t \iff (e \prec_E e') \text{ or } (e = e' \text{ and } s \preceq_e t). \quad (1.2.4)$$

If each \mathcal{T}_e admits an upper bound whenever e is not maximal in E , then $(\bigsqcup_{e \in E} \mathcal{T}_e, \preceq, \bigsqcup_{e \in E} \mathcal{D}_e)$ is an indexing semilattice.

Proof. Checking that (1.2.4) defines an order relation is straightforward. Now let us prove that $(\mathcal{T}, \preceq, \mathcal{D}) = (\bigsqcup_{e \in E} \mathcal{T}_e, \preceq, \bigsqcup_{e \in E} \mathcal{D}_e)$ is an indexing semilattice. For each non-maximal $e \in E$, denote by M_e the upper bound in \mathcal{T}_e .

We claim that for all $s, t \in \mathcal{T}$, $s \wedge t$ is well-defined. Indeed, for $s \in \mathcal{T}_e$ and $t \in \mathcal{T}_{e'}$, three cases may happen:

- ◊ s and t are comparable for \preceq , so $s \wedge t$ is equal to either s or t .
- ◊ s and t are not comparable for \preceq and $e = e'$, so $s \wedge t = s \wedge_e t$.
- ◊ s and t are not comparable for \preceq and $e \neq e'$, so $s \wedge t = M_{e \wedge_E e'}$.

Hence (\mathcal{T}, \preceq) is closed under finite minimum.

Now consider a countable subset $\{t_i : i \in \mathbb{N}\} \subseteq \mathcal{T}$ such that $t_i \in \mathcal{T}_{e_i}$ for all $i \in \mathbb{N}$ and let us show that $\bigwedge_{i \in \mathbb{N}} t_i$ exists. Since (\mathcal{T}, \preceq) is closed under finite minimum, we may suppose without loss of generality for all $i \in \mathbb{N}$ that $t_{i+1} \preceq t_i$. Denote by e the minimum in (E, \preceq_E) of $\{e_i : i \in \mathbb{N}\}$. From here, two mutually exclusive cases may happen:

- ◊ There exists $j \in \mathbb{N}$ such that $e_i = e$ for all $i \geq j$, so $\bigwedge_{i \in \mathbb{N}} t_i$ is the minimum in $(\mathcal{T}_e, \preceq_e)$ of $\{t_i : i \geq j\}$.
- ◊ For all $i \in \mathbb{N}$, $e \prec_E e_i$ for all $i \in \mathbb{N}$, so $\bigwedge_{i \in \mathbb{N}} t_i = M_e$.

Since it is also clear that $\bigsqcup_{e \in E} \mathcal{D}_e$ is a countable dense subset, it makes $(\bigsqcup_{e \in E} \mathcal{T}_e, \preceq, \bigsqcup_{e \in E} \mathcal{D}_e)$ into an indexing semilattice. \square

As a first application of Proposition 1.2.21, we give the *compactification* of an indexing semilattice, very similar in spirit to Alexandrov's compactification procedure. Although compactification does not have in general nice enough properties for the set-indexed setting (see the table page 17), it still has its uses for some constructions that follow (see e.g. Proposition 1.2.21) and examples like the hypersphere \mathbb{S}^p (Example 1.2.29).

COROLLARY 1.2.22 (Compactification of an indexing semilattice). *Let ∞ be a point that does not belong to \mathcal{T} and extend \preceq to $\mathcal{T} \cup \{\infty\}$ so that $t \preceq \infty$ for all $t \in \mathcal{T}$. The triplet $(\mathcal{T} \cup \{\infty\}, \preceq, \mathcal{D} \cup \{\infty\})$ is then an indexing semilattice of which \mathcal{T} is an ideal.*

The proof being straightforward, we choose to skip it. Remark however the possible following interpretation: the indexing semilattice $\mathcal{T} \cup \{\infty\}$ corresponds to the disjoint union $\bigsqcup_{e \in E} \mathcal{T}_e$ where $(E, \preceq_E) = (\{0, 1\}, \preceq)$, $\mathcal{T}_0 = \mathcal{T}$ and $\mathcal{T}_1 = \{\infty\}$ is the trivial indexing semilattice from Example 1.2.6.

Proposition 1.2.21 also constitutes the first step towards more interesting kinds of gluings. They are better described if split into two steps: *series gluing* and *parallel gluing*.

The series gluing's statement is longer than its proof — that we skip — and is just to express the fact that, in any disjoint union $\bigsqcup_{e \in E} \mathcal{T}_e$ of indexing semilattices, one may identify the maximum of \mathcal{T}_e together with the minimum of $\mathcal{T}_{e'}$ whenever e is a maximal in $\{e'' \in E : e'' \prec_E e'\}$.

PROPOSITION 1.2.23 (Series gluing of indexing semilattices). *Consider the following:*

1. (E, \preceq_E, E) an at most countable indexing semilattice,
2. $\{(\mathcal{T}_e, \preceq_e, \mathcal{D}_e) : e \in E\}$ a family of indexing semilattices where for all $e \in E$, 0_e denotes the global minimum of \mathcal{T}_e and whenever e is not maximal in E , \mathcal{T}_e has a global maximum $M_e \in \mathcal{T}_e$,
3. $(\mathcal{T} = \bigsqcup_{e \in E} \mathcal{T}_e, \preceq, \mathcal{D})$ the disjoint union indexing semilattice from Proposition 1.2.21,
4. \sim the finest equivalence relation on \mathcal{T} such that $M_e \sim 0_{e'}$ whenever e is a maximal element in $\{e'' \in E : e'' \prec_E e'\}$, and $\vec{\mathcal{T}} = \mathcal{T}/\sim$ (resp. $\pi : \mathcal{T} \rightarrow \vec{\mathcal{T}}$) the corresponding quotient space (resp. canonical projection).

Then \preceq is compatible with \sim , making $(\vec{\mathcal{T}}, \preceq, \pi(\mathcal{D}))$ into an indexing semilattice, called the series gluing of \mathcal{T} along E . Moreover, π is a morphism and for all $e \in E$, the set $\pi(\vec{\mathcal{T}}_e)$ — where $\vec{\mathcal{T}}_e = \bigsqcup_{e' \preceq_E e} \mathcal{T}_{e'}$ — is an ideal.

As an application, we use series gluing to make the natural indexing collection on a discrete tree (Example 1.2.24) into a continuous one where edges are part of the indexing space (Example 1.2.25).

EXAMPLE 1.2.24 (Indexing collection on a discrete tree). *Following Neveu's convention from [78], a (discrete) tree is a non-empty subset $\mathcal{U} \subseteq \bigcup_{k \in \mathbf{N}} \mathbf{N}^k$ with the convention $\mathbf{N}^0 = \emptyset' = \{0_{\mathcal{U}}\}$ and such that for all $k \in \mathbf{N}^*$ and $t = (t_1, \dots, t_k) \in \mathcal{U}$, (t_1, \dots, t_{k-1}) also belongs to \mathcal{U} (with the convention that $(t_1, \dots, t_0) = 0_{\mathcal{U}}$). Per tradition and for an easier notation, we will write $t_1 \dots t_k$ instead of (t_1, \dots, t_k) .*

The element $0_{\mathcal{U}}$ — which must then belong to \mathcal{U} — is called the root of \mathcal{U} and for any $t_1 \dots t_k \in \mathcal{U}$ and any $j \in \llbracket 0, k-1 \rrbracket$, $t_1 \dots t_j \in \mathcal{U}$ is called an ancestor of t .

The set \mathcal{U} may be endowed with the partial order \preceq for which $s \preceq t$ if and only if s is an ancestor of t . In particular, for any subset $\{t^i : i \in \mathbf{N}\} \subseteq \mathcal{U}$, $\bigwedge_{i \in \mathbf{N}} t^i$ is the biggest common ancestor to all t^i 's. Since \mathcal{U} is at most countable, it makes $(\mathcal{U}, \preceq, \mathcal{U})$ into an indexing semilattice.

EXAMPLE 1.2.25 (Indexing collection on a continuous version of a discrete tree). *Consider a discrete tree \mathcal{U} as defined in Example 1.2.24 and denote $\mathcal{T} = \{0_{\mathcal{T}}\} \sqcup \bigsqcup_{u \in \mathcal{U} \setminus \{0_{\mathcal{U}}\}} [0, 1]$.*

The series gluing $\vec{\mathcal{T}}$ of \mathcal{T} along \mathcal{U} may be endowed with an indexing collection by Proposition 1.2.23. This is illustrated in Figure 1.2.

This construction may be seen as a related — albeit simpler — case of Aldous's [7, 8] stick-breaking construction of the Continuum Random Tree (CRT).

The series gluing however has serious limitations since indexing semilattices may only be glued together by their extremal elements. We wish to express a different kind of gluing which allows for bigger overlaps between the indexing semilattices \mathcal{T}_e that cover \mathcal{T} . This is what the parallel gluing achieves, without however supplanting the series gluing (see Example 1.2.27 for a related discussion).

To our knowledge, most of the gluing procedures that have been considered by probabilists so far have been by gluing points. This construction provides a way to glue along more general surfaces.

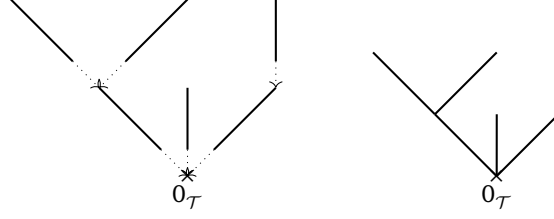


Figure 1.2: Series gluing of $\{0_{\mathcal{T}}\} \sqcup \bigsqcup_{u \in \mathcal{U} \setminus \{0_{\mathcal{U}}\}} [0, 1]$ along the discrete tree $\mathcal{U} = \{0_{\mathcal{U}}, 0, 1, 2, 00, 01, 20\}$.

PROPOSITION 1.2.26 (Parallel gluing of indexing semilattices). Consider a set \mathcal{T} and an at most countable set E with a point $0_E \in E$. Suppose that there exists a family of indexing semilattices $\{(\mathcal{T}_e, \preceq_e, \mathcal{D}_e) : e \in E \setminus \{0_E\}\}$ and denote by $\mathcal{T}_{0_E} = \{0_{\mathcal{T}}\}$ the trivial indexing semilattice from Example 1.2.6. Define the order relation \preceq_E on E as follows:

$$\forall e, e' \in E, \quad e \preceq_E e' \iff e = 0_E.$$

Denote by $\vec{\mathcal{T}}$ the resulting series gluing along (E, \preceq_E) and suppose that the following hold:

1. (Covering of \mathcal{T}). There is a surjective map $\pi : \vec{\mathcal{T}} \twoheadrightarrow \mathcal{T}$ such that for all $e \in E$, $\pi_e = \pi|_{\mathcal{T}_e}$ is injective. Denote by $\varphi_e : \pi(\mathcal{T}_e) \rightarrow \mathcal{T}_e$ its inverse.
2. (Compatibility of the covering). For all $e, e' \in E$, the set $\mathcal{T}_{ee'} = \varphi_e(\pi(\mathcal{T}_e) \cap \pi(\mathcal{T}_{e'}))$ is an ideal of \mathcal{T}_e isomorphic to $\mathcal{T}_{e'e} = \varphi_{e'}(\pi(\mathcal{T}_{e'}) \cap \pi(\mathcal{T}_e))$. Denote by $\varphi_{e'e} : \mathcal{T}_{ee'} \rightarrow \mathcal{T}_{e'e}$ the isomorphism.
3. (Retraction of the covering). For all $e, e' \in E$, there exists a projection $\pi_{ee'} : \mathcal{T}_e \twoheadrightarrow \mathcal{T}_{e'e}$ in the sense of Corollary 1.2.12.

Define a binary relation \preceq on \mathcal{T} by

$$\forall s, t \in \mathcal{T}, \quad s \preceq t \iff \exists e \in E : s, t \in \pi(\mathcal{T}_e) \text{ and } \varphi_e(s) \preceq_e \varphi_e(t). \quad (1.2.5)$$

Then $(\mathcal{T}, \preceq, \pi(\bigsqcup_{e \in E} \mathcal{D}_e))$ is an indexing semilattice and $\pi : \vec{\mathcal{T}} \twoheadrightarrow \mathcal{T}$ is a morphism.

Proof. For the diagram-lovers among us, the notations are summarized in the commutative diagram of Figure 1.3 where \twoheadrightarrow represents ‘purely’ surjective maps, \hookrightarrow canonical injections (due to the inclusions between the sets at play) and \leftrightarrow bijections.

Let us prove that (1.2.5) defines an order relation. First, we claim that

$$\forall s, t \in \mathcal{T}, \quad s \preceq t \iff (\forall e \in E, t \in \pi(\mathcal{T}_e) \implies s \in \pi(\mathcal{T}_e) \text{ and } \varphi_e(s) \preceq_e \varphi_e(t)). \quad (1.2.6)$$

The converse implication is obvious. For the direct one, consider $s, t \in \mathcal{T}$ such that $s \preceq t$ and $e \in E$ such that the right-hand side of (1.2.5) holds. Then take $e' \in E$ such that $t \in \pi(\mathcal{T}_{e'})$. Since $\varphi_e(t) \in \mathcal{T}_{ee'}$, $\varphi_e(s) \preceq_e \varphi_e(t)$ and $\mathcal{T}_{ee'}$ is an ideal of \mathcal{T}_e , we get $\varphi_e(s) \in \mathcal{T}_{ee'}$. Thus $s, t \in \pi(\mathcal{T}_e) \cap \pi(\mathcal{T}_{e'})$. In particular, it means that $\varphi_e(s), \varphi_e(t) \in \mathcal{T}_{ee'}$ and $\varphi_{e'}(s), \varphi_{e'}(t) \in \mathcal{T}_{e'e}$. Since $\mathcal{T}_{ee'}$ and $\mathcal{T}_{e'e}$ are isomorphic, $\varphi_e(s) \preceq_e \varphi_e(t)$ if and only if $\varphi_{e'}(s) \preceq_{e'} \varphi_{e'}(t)$. The claim follows.

Let us prove that \preceq is an order relation on \mathcal{T} . Reflexivity is trivial since $\mathcal{T} = \bigcup_{e \in E} \pi(\mathcal{T}_e)$ and the \preceq_e ’s are all reflexive. For antisymmetry, consider $s, t \in \mathcal{T}$ such that $s \preceq t$ and $t \preceq s$. By (1.2.5) and (1.2.6), it means that there exists $e \in E$ such that $\varphi_e(s) = \varphi_e(t)$. Hence $s = t$. As for

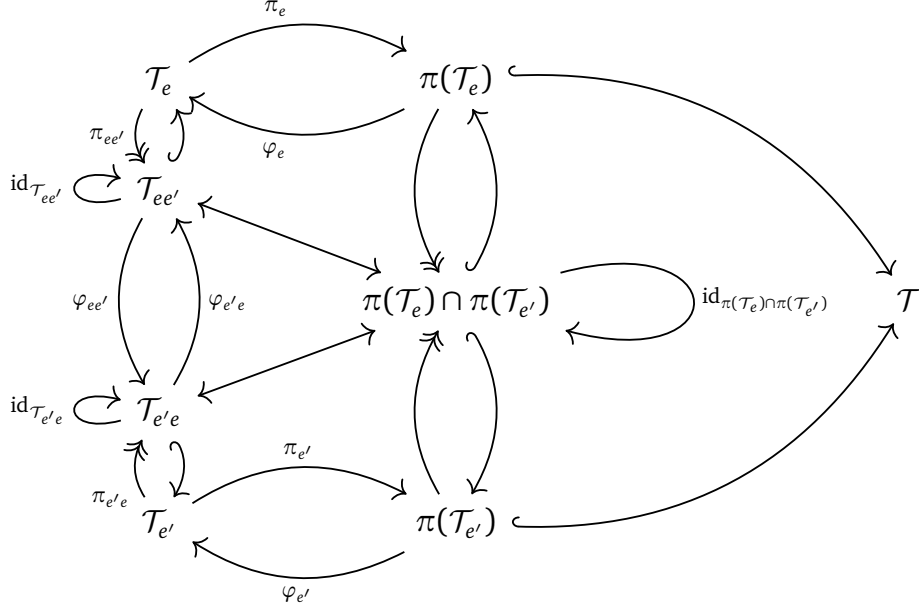


Figure 1.3: Diagram illustrating the parallel gluing from Proposition 1.2.26.

triangle inequality, consider $s, t, u \in \mathcal{T}$ such that $s \preceq t$ and $t \preceq u$. By (1.2.5), there exists $e \in E$ such that $t, u \in \pi(\mathcal{T}_e)$ and $\varphi_e(t) \preceq_e \varphi_e(u)$. By (1.2.6), $s \in \pi(\mathcal{T}_e)$ as well and $\varphi_e(s) \preceq_e \varphi_e(t)$. Thus $\varphi_e(s) \preceq_e \varphi_e(u)$, which implies that $s \preceq u$.

Hence \preceq is an order relation.

Consider $s, t \in \mathcal{T}$ and let us show that their minimum $s \wedge t$ is well-defined. For all $e, e' \in E$, denote by $\psi_{e'e} : \pi(\mathcal{T}_{e'}) \rightarrow \mathcal{T}_{ee'}$ the map $\psi_{e'e} = \varphi_{e'e} \pi_{e'e} \varphi_{e'}$ (we recommend looking at the diagram to see what this map does). Let $s, t \in \mathcal{T}$, consider $e \in E$ such that $s \in \pi(\mathcal{T}_e)$ and denote $u = \pi(\varphi_e(s) \wedge_e \psi_{e'e}(t))$ where \wedge_e is the minimum in \mathcal{T}_e (we know that \wedge_e is well-defined since \mathcal{T}_e is an ideal of $\overline{\mathcal{T}}$ by Proposition 1.2.23). By (1.2.5), $u \preceq s$. Moreover, since $\mathcal{T}_{ee'}$ and $\mathcal{T}_{e'e}$ are isomorphic, we also have $u = \pi(\psi_{e'e}(s) \wedge_{e'} \varphi_{e'}(t))$. Thus $u \preceq t$ as well. In order to show that $u = s \wedge t$, it remains to prove the following: $\forall v \in \mathcal{T}, (v \preceq s \text{ and } v \preceq t) \implies v \preceq u$. So consider $v \in \mathcal{T}$ such that $v \preceq s$ and $v \preceq t$. Fix $e \in E$ (resp. $e' \in E$) such that $v, s \in \pi(\mathcal{T}_e)$ and $\varphi_e(v) \preceq_e \varphi_e(s)$ (resp. $v, t \in \pi(\mathcal{T}_{e'})$ and $\varphi_{e'}(v) \preceq_{e'} \varphi_{e'}(t)$). Applying the morphism $\varphi_{e'e} \pi_{e'e}$ to $\varphi_{e'}(v) \preceq_{e'} \varphi_{e'}(t)$ yields $\psi_{e'e}(v) \preceq_e \psi_{e'e}(t)$. But since $v \in \pi(\mathcal{T}_e) \cap \pi(\mathcal{T}_{e'})$, we get $\psi_{e'e}(v) = \varphi_e(v)$. Thus $\varphi_e(v) \preceq_e (\varphi_e(s) \wedge_e \psi_{e'e}(t))$, which implies $v \preceq u$ by (1.2.5). Hence $u = s \wedge t$.

Let us now prove that (\mathcal{T}, \preceq) is closed under countable minimum. Since it is already closed under finite minimum, it is enough to prove that for any non-decreasing sequence $(t_i)_{i \in \mathbb{N}} \in \mathcal{T}$, $\bigwedge_{i \in \mathbb{N}} t_i$ exists. For such a sequence, consider $e \in E$ such that $t_0 \in \pi(\mathcal{T}_e)$. Since for all $i \in \mathbb{N}$, we have $t_i \preceq t_0$, we know that $t_i \in \pi(\mathcal{T}_e)$. So we may define $u = \pi(\bigwedge_{i \in \mathbb{N}} \varphi_e(t_i))$. It is obvious that for all $i \in \mathbb{N}$, $u \preceq t_i$. Conversely, if $v \in \mathcal{T}$ is such that for all $i \in \mathbb{N}$, $v \preceq t_i$, then $v \in \pi(\mathcal{T}_e)$ and for all $i \in \mathbb{N}$, $\varphi_e(v) \preceq_e \varphi_e(t_i)$. Thus $\varphi_e(v) \preceq_e \bigwedge_{i \in \mathbb{N}} \varphi_e(t_i)$, which implies that $v \preceq u$. Hence $u = \bigwedge_{i \in \mathbb{N}} t_i$.

Now, what remains is to prove that $\pi(\bigsqcup_{e \in E} \mathcal{D}_e)$ is a countable dense subset of \mathcal{T} and that π is a morphism, both of which are straightforward. \square

EXAMPLE 1.2.27 (Why such conditions for the parallel gluing?). We wish to comment on some assumptions of Proposition 1.2.26 that might appear unnecessary at first glance. Namely, we want to argue why the $\mathcal{T}_{e\epsilon'}$'s need to be ideals and why there is a need for the projections $\pi_{e\epsilon'} : \mathcal{T}_e \rightarrow \mathcal{T}_{e\epsilon'}$.

First, one could try to use parallel gluing to give an indexing collection for the circle \mathbf{S}^1 by gluing two copies of the unit segment $[0, 1]$. However, nothing in the hypotheses — apart from the ideal condition — prevents you from doing it in the silly fashion illustrated in Figure 1.4 which does not allow for a consistent indexing semilattice structure (every point is both smaller and greater than all the others).

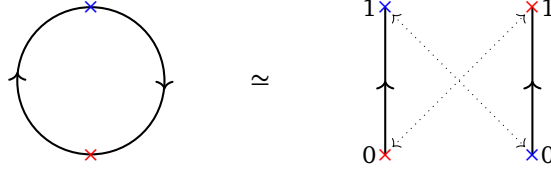


Figure 1.4: Gluing two unit segments head-to-toe fails to make an indexing collection on $\mathcal{T} = \mathbf{S}^1$.

Moving on, the ‘retraction of the covering’ condition also is necessary since gluing does not always preserve the semilattice structure. Indeed, consider once more two copies of the unit segment $[0, 1]$ and glue them so that each copy of $[0, 1/2]$ is identified with one another, i.e. $\mathcal{T} = ([0, 1] \sqcup [0, 1]) / \sim$ where

$$\forall (x, i), (y, j) \in \mathcal{T}, \quad (x, i) \sim (y, j) \iff ((x, i) = (y, j) \text{ or } x = y < 1/2).$$

In that case, the minimum between the $1/2$ belonging to the first copy of $[0, 1]$ and the $1/2$ of the second one is ill-defined. And remark that the only hypothesis that this example fails to verify is indeed the ‘retraction of the covering’.

EXAMPLE 1.2.28 (Indexing collection on $\mathcal{T} = \mathbf{R}^p$). Let $p \in \mathbf{N}^*$. Consider for all $\epsilon \in \{\pm 1\}^p$ a copy \mathcal{T}_ϵ of the indexing semilattice \mathbf{R}_+^p from Example 1.2.10. The projection $\pi_\epsilon : \mathcal{T}_\epsilon \rightarrow \mathbf{R}^p$ is then given by $\pi_\epsilon(x) = \epsilon \cdot x$ for all $x \in \mathbf{R}_+^p$, so $\pi_\epsilon(\mathcal{T}_\epsilon)$ is one of the 2^p ‘quadrants’ of \mathbf{R}^p as illustrated in Figure 1.5.

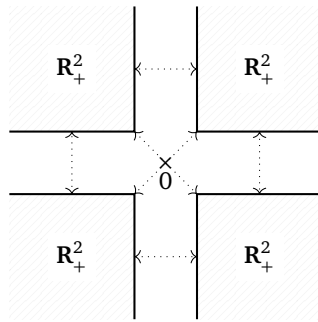


Figure 1.5: Parallel gluing of $\bigsqcup_{\epsilon \in \{\pm 1\}^2} \mathbf{R}_+^2$ onto \mathbf{R}^2 .

Then for all ϵ, ϵ' , $\mathcal{T}_{\epsilon\epsilon'}$ and $\mathcal{T}_{\epsilon'\epsilon}$ are both isomorphic to \mathbf{R}_+^k where $k = k(\epsilon, \epsilon') = \#\{i \in \llbracket 1, p \rrbracket : \epsilon_i = \epsilon'_i\}$, so the compatibility of the covering is easy to check. As for the projection $\pi_{\epsilon\epsilon'}$, it corresponds to the usual linear projection: $\forall x = (x_1, \dots, x_p) \in \mathcal{T}_\epsilon$, $\pi_{\epsilon\epsilon'}(x) = (x_{j_1}, \dots, x_{j_k})$ where $j_1 < \dots < j_k$ is such that for all i , $\epsilon_{j_i} = \epsilon'_i$. Hence Proposition 1.2.26 endows $\mathcal{T} = \mathbf{R}^p$ with an indexing collection.

Another way to go about it is to directly specify the order relation \preceq on \mathbf{R}^p , which corresponds here to

$$\forall s, t \in \mathbf{R}^p, \quad s \preceq t \iff \forall i \in \llbracket 1, p \rrbracket, s_i t_i \geq 0 \text{ and } |s_i| \leq |t_i|.$$

Remark that we cannot consider the usual component-wise order on \mathbf{R}^p for the same reason as in the case $\mathcal{T} = \mathbf{R}$ (Example 1.2.8).

EXAMPLE 1.2.29 (Indexing collection on $\mathcal{T} = \mathbf{S}^p$). Let $p \in \mathbf{N}^*$ and denote by $\varphi : \mathbf{S}^p \rightarrow \mathbf{R}^p \cup \{\infty\}$ a stereographic projection (it is a bijection). By compactification (Corollary 1.2.22), we have an indexing semilattice structure $(\mathbf{R}^p \cup \{\infty\}, \preceq, \mathcal{D})$. Since φ is bijective, by pushforward (Corollary 1.2.17), we know that $(\mathbf{S}^p, \preceq', \varphi^{-1}(\mathcal{D}))$ is an indexing semilattice where

$$\forall s, t \in \mathbf{S}^p, \quad s \preceq' t \iff \varphi(s) \preceq \varphi(t).$$

More generally, manifolds may be endowed with an indexing collection as long as one is able to order the local maps in a way consistent with a gluing procedure.

EXAMPLE 1.2.30 (Indexing collection on a continuous tree, 2nd version). In Example 1.2.25, we told how to use series gluing to endow a non-fractal continuous tree with an indexing collection. Here, we do an equivalent construction, but based on parallel gluing instead. This way is interesting as well because it respects the flows (Definition 1.3.4) in \mathcal{T} . We refer to Section 1.3.2 for more details.

Consider a discrete tree \mathcal{U} as defined in Example 1.2.24 and define the height function $h : \mathcal{U} \rightarrow \mathbf{N}$ such that for all $k \in \mathbf{N}$, $h|_{\mathcal{U} \cap \mathbf{N}^k} = k$. Define for all $u \in \mathcal{U}$, $\mathcal{T}_u = [0, h(u)]$ endowed with its usual indexing semilattice structure (see Example 1.2.14). One should think about \mathcal{T}_u as a path from the root to u . Obviously, two paths to u and v respectively should coincide until they reach their highest common ancestor $u \wedge v$ and then part ways, so the quotient space \mathcal{T} should reflect that.

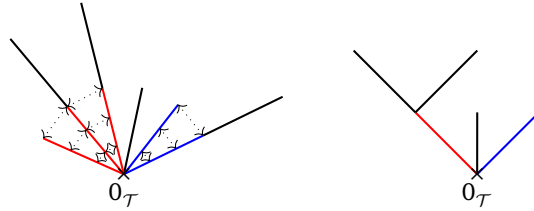


Figure 1.6: Parallel gluing based on the discrete tree $\mathcal{U} = \{0_{\mathcal{U}}, 0, 1, 2, 00, 01, 20\}$.

Just like in Proposition 1.2.26, consider the indexing semilattice $\vec{\mathcal{T}}$ obtained by identifying all the $0_{\mathcal{T}_u}$ together in $\bigsqcup_{u \in \mathcal{U}} \mathcal{T}_u$. Using the definition of the disjoint union (1.2.3), we make a slight abuse of notation quotient-wise and denote a generic element of $\vec{\mathcal{T}}$ by (s, u) for some $u \in \mathcal{U}$ and $s \in \mathcal{T}_u$. Define an equivalence relation \sim on $\vec{\mathcal{T}}$ by

$$\forall (s, u), (t, v) \in \vec{\mathcal{T}}, \quad (s, u) \preceq (t, v) \iff s = t \leq h(u \wedge v)$$

and denote $\mathcal{T} = \vec{\mathcal{T}} / \sim$ the corresponding quotient space. This is what is illustrated in Figure 1.6 where $\vec{\mathcal{T}}$ is represented on the left and \mathcal{T} on the right.

In order to endow \mathcal{T} with an indexing collection through parallel gluing, we just need to specify the elements that appear in the statement of Proposition 1.2.26. First, we have $\pi : \vec{\mathcal{T}} \rightarrow \mathcal{T}$ as

the canonical projection associated with the quotient. For all $u, v \in \mathcal{U}$, we have $\mathcal{T}_{uv} = \{(s, u) : s \in [0, h(u \wedge v)]\}$ as well as

$$\varphi_{uv} : \begin{array}{ccc} \mathcal{T}_{uv} & \longrightarrow & \mathcal{T}_{vu} \\ (s, u) & \longmapsto & (s, v) \end{array} \quad \text{and} \quad \pi_{uv} : \begin{array}{ccc} \mathcal{T}_u & \longrightarrow & \mathcal{T}_{uv} \\ (s, u) & \longmapsto & (\min\{s, h(u \wedge v)\}, u). \end{array}$$

All of those elements readily check the hypotheses of Proposition 1.2.26, so \mathcal{T} may be endowed with an indexing semilattice structure by parallel gluing.

EXAMPLE 1.2.31 (Indexing collection on a continuous tree, 3rd version). The astute reader might have already remarked that Examples 1.2.25 and 1.2.30 are but specific cases of a (rooted) \mathbf{R} -tree, i.e. a metric space $(\mathcal{T}, d_{\mathcal{T}})$ with a distinguished point $\rho \in \mathcal{T}$ called the root and such that for any $s, t \in \mathcal{T}$, there is a unique isometric embedding $\phi_{s,t} : [0, d_{\mathcal{T}}(s, t)] \rightarrow \mathcal{T}$ such that $\phi_{s,t}(0) = s$ and $\phi_{s,t}(d_{\mathcal{T}}(s, t)) = t$. Denote then by $[s, t] = \phi_{s,t}([0, d_{\mathcal{T}}(s, t)])$ the geodesic segment of extremities s and t . Using this structure, define the partial order \preceq on \mathcal{T} by:

$$\forall s, t \in \mathcal{T}, \quad s \preceq t \iff [\rho, s] \subseteq [\rho, t].$$

The ordered set (\mathcal{T}, \preceq) is closed by (finite) minimum and is such that for all $s, t \in \mathcal{T}$, $[\rho, s] \cap [\rho, t] = [\rho, s \wedge t]$ (see [31, Lemma 3.20]). So in order to show that \mathcal{T} is closed under countable minimum, it is enough to check that $\bigwedge_{i \in \mathbb{N}} t_i \in \mathcal{T}$ for all non-increasing sequence $(t_i)_{i \in \mathbb{N}}$ in \mathcal{T} . In this case, since ϕ_{ρ, t_0} is an increasing map, it is easy to see that $\bigwedge_{i \in \mathbb{N}} t_i = \phi_{\rho, t_0}(\inf_{i \in \mathbb{N}} d_{\mathcal{T}}(\rho, t_i))$. Hence \mathcal{T} may be endowed with an indexing collection provided that it is separated from above by a countable dense subset.

Remark however that might not always be the case since (\mathcal{T}, \preceq) may have uncountably many ‘leaves’ (or ends, see [31, Section 3.4.1]) which cannot be approximated from above by anything else than themselves since they are maximal elements for \preceq . We refer to [31, 65] for more general and systematic overviews on \mathbf{R} -trees and their importance in modern probability.

Summary

Later on, the indexing semilattice \mathcal{T} will be endowed with a *compatible measure* \mathbf{m} (Section 1.3) and a *compatible metric* $d_{\mathcal{T}}$ (Section 1.4). Sometimes, it may even be supposed finite-dimensional in a specific sense (Section 1.5). We collected in Table 1.1 how the several constructions we have just presented interact with those additional structures. The notations used are either the same as the corresponding propositions or sufficiently clear not to be stated more explicitly (e.g. \mathbf{m}_i corresponds to a compatible measure on the indexing semilattice \mathcal{T}_i). The cases where additional hypotheses are required for all the properties to hold are indicated in red. A ‘ \times ’ indicates a case where we did not find any obvious compatibility. The proofs are not provided since we felt they are essentially straightforward and do not bring much more to the table.

	Compatible measure \mathbf{m} (Def. 1.3.1)	Compatible metric $d_{\mathcal{T}}$ (Def. 1.4.1)	Dimension (Def. 1.5.3)
Finite product (Prop. 1.2.9)	$\mathbf{m} = \bigotimes_{i=1}^k \mathbf{m}_i$	$d_{\mathcal{T}} = \sup_{1 \leq i \leq k} d_{\mathcal{T}_i}$	$\dim \mathcal{T} = \sum_{i=1}^k \dim \mathcal{T}_i$
Countable product (Prop. 1.2.9)	$\mathbf{m} = \bigotimes_{i \in \mathbb{N}} \mathbf{m}_i$	$d_{\mathcal{T}} = \sup_{i \in \mathbb{N}} (d_{\mathcal{T}_i} \wedge 2^{-i})$	$\dim \mathcal{T} = \sum_{i \in \mathbb{N}} \dim \mathcal{T}_i$

Projection (Cor. 1.2.12)	$m = \pi_* m'$	$d_{\mathcal{T}} = d_{\mathcal{T}'} _{\mathcal{T} \times \mathcal{T}}$	$\dim \mathcal{T} \leq \dim \mathcal{T}'$
Compactification (Cor. 1.2.22)	$m(A \cup \{\infty\}) = m(A)$	\times	$\dim(\mathcal{T} \cup \{\infty\}) = \dim \mathcal{T}$
Disjoint union (Prop. 1.2.21)	$m = \sum_{e \in E} m_e(\mathcal{T}_e \cap \cdot)$	Quotient metric such that $\forall e \in E$, $d_{\mathcal{T}} _{\mathcal{T}_e \times \mathcal{T}_e} = d_{\mathcal{T}_e}$ and $d_{\mathcal{T}}(M_e, 0_{e'}) = 0$ whenever e is maximal in $[0_E, e']$.	$\dim \mathcal{T} = \max \{ \dim E, \sup_{e \in E} \dim \mathcal{T}_e \}$
Series gluing (Prop. 1.2.23)	$m = \pi_* \left(\sum_{e \in E} m_e(\mathcal{T}_e \cap \cdot) \right)$	Projection of the metric given by the disjoint union.	$\dim \mathcal{T} = \max \{ \dim E, \sup_{e \in E} \dim \mathcal{T}_e \}$
Parallel gluing (Prop. 1.2.26)	$m = \pi_* \left(\sum_{e \in E} m_e(\mathcal{T}_e \cap \cdot) \right)$	Projection of the metric given by the disjoint union.	$\dim \mathcal{T} = \max \{ \dim E, \sup_{e \in E} \dim \mathcal{T}_e \}$

Table 1.1: Compatibilities between constructions of indexing semilattices and additional structures on them.

Further examples

Even though infinite-dimensional examples might not be our focus point later on, we also wanted to give infinite-dimensional examples different from Hilbert's cube (Example 1.2.16).

EXAMPLE 1.2.32 (Indexing collection on the set of concave, non-negative functions). Suppose that \mathcal{T} is the set of all concave functions $f : [0, 1] \rightarrow \mathbf{R}_+$ endowed with the pointwise partial order \leq , i.e. $f \leq g$ if and only if $f(x) \leq g(x)$ for all $x \in [0, 1]$. Since a function is concave if and only if it may be written as the minimum of affine functions, (\mathcal{T}, \leq) is closed under any minimum.

Let $n \in \mathbf{N}$ and denote by $\mathcal{D}_n = \{k2^{-n} : 0 \leq k \leq 2^n\}$ the dyadics of order n in $[0, 1]$ and by \mathcal{D}_n the set of all functions $f \in \mathcal{T}$ such that for all $d \in \mathcal{D}_n$, $f(d) \in n\mathcal{D}_n$ and $f|_{(d, d+2^{-n})}$ is affine. Since any tangent of a concave function may be approximated from above by a function in $\mathcal{D} = \bigcup_{n \in \mathbf{N}} \mathcal{D}_n$ and that \mathcal{D} is closed un finite minimum, the set \mathcal{D} verifies the separability from above condition.

EXAMPLE 1.2.33 (Indexing collection on $\mathcal{T} = C^0(K; \mathbf{R}_+)$). Suppose that K is a compact metric space. Like Example 1.2.32, the space $\mathcal{T} = C^0(K; \mathbf{R}_+)$ is endowed with the pointwise partial order \leq . $C^0(K; \mathbf{R}_+)$ is obviously closed under finite minimum, but also under countable ones by Dini's first theorem [63, Chapter IX, Theorem 1.3]. Moreover, since it is separable by [9, Lemma 3.99], we may consider a countable subset \mathcal{D} of $C^0(K; \mathbf{R}_+)$ which is dense for the uniform convergence. Let us show that any fixed function $f \in C^0(K; \mathbf{R}_+)$ may be approximated from above by a sequence in \mathcal{D} . For all $n \in \mathbf{N}$, there exists a function $f_n \in \mathcal{D}$ such that $\|(f + 2^{-n+2}) - f_n\|_{\infty} \leq 2^{-n}$. In particular, we have

$$\forall n \in \mathbf{N}, \quad f + 3 \cdot 2^{-n} \leq f_n \leq f + 5 \cdot 2^{-n}$$

from which we deduce that the sequence $(f_n)_{n \in \mathbf{N}}$ decreases to f . Hence $(C^0(K; \mathbf{R}_+), \leq, \mathcal{D})$ is an indexing semilattice.

More generally, a lot of infinite-dimensional examples of indexing semilattices may be found among the positive cones of function spaces in the sense of [9, Definition 1.1].

1.2.3 \mathcal{A} 's friends

The indexing collection \mathcal{A} does not go alone, it naturally generates new classes of sets that we introduce here. All of them can already be found in [47], apart from the class $\mathcal{C}_{(k)}$ which is new, but has ties with the class $\mathcal{B}^k(U, \rho)$ considered in [42, Section 3.2].

DEFINITION 1.2.34 (Class $\mathcal{D}(u)$). *For any subset $\mathcal{D} \subseteq \mathcal{B}$, the class of finite unions of elements in \mathcal{D} is denoted by*

$$\mathcal{D}(u) = \left\{ \bigcup_{i=1}^k D_i : k \in \mathbf{N}, D_0, \dots, D_k \in \mathcal{D} \right\}.$$

The class $\mathcal{A}(u)$ has its importance when looking at the distribution of a set-indexed process $X = \{X_A : A \in \mathcal{A}\}$. In Section 1.3.2, we will see that the distribution of X is characterized by its distribution along increasing paths in $\mathcal{A}(u)$, but is not along paths in \mathcal{A} (see Theorem 1.3.13 for a precise statement). Remark that phenomenon is invisible in dimension one since $\mathcal{A}(u) = \mathcal{A}$ for the usual indexing collection on \mathbf{R}_+ (Example 1.2.7).

DEFINITION 1.2.35 (Increment classes). *The class of (simple) increment sets is denoted by*

$$\mathcal{C} = \{A \setminus U : A \in \mathcal{A}, U \in \mathcal{A}(u)\}.$$

For any $k \in \mathbf{N}$, the subclass $\mathcal{C}_{(k)} \subseteq \mathcal{C}$ of k -increments is given by

$$\mathcal{C}_{(k)} = \left\{ A_0 \setminus \bigcup_{i=1}^k A_i : A_0, \dots, A_k \in \mathcal{A} \right\}.$$

One obviously has $\mathcal{A} \subseteq \mathcal{C}_{(k)} \subseteq \mathcal{C}_{(k+1)} \subseteq \mathcal{C} \subseteq \mathcal{C}(u) \subseteq \mathcal{B}$ for all $k \in \mathbf{N}$ where each inclusion is strict in general.

The classes \mathcal{C} and $\mathcal{C}(u)$ are a semiring of sets and a ring of sets respectively (see [59, Definitions 1.8 and 1.9]). Thus they are well-adapted to measure-theoretic constructions, just like in Section 2.2.2 where we extend set-indexed processes to stochastic measures. The class \mathcal{C} is also a natural extension of the rectangular increments of \mathbf{R}_+^2 -indexed processes (see Example 1.2.37).

The subclasses $\mathcal{C}_{(k)}$ will play an important role to characterize a dimensional property of \mathcal{A} (see Section 1.5). They also are used to define Hölder exponents for set-indexed maps (see Section 3.4). The particular case of $\mathcal{C}_{(1)}$ — classically denoted by \mathcal{C}_0 — has been used in [39, 40] (resp. [41]) to characterize increment stationarity for set-indexed fractional Brownian motion (resp. set-indexed Lévy processes). Those processes are defined respectively in (1.6.3) and Definition 2.4.1 and a study of increment stationarity properties is carried out in Section 2.5. The reason why we chose not to stick with the former notation is that making the notation $\mathcal{C}_{(k)}$ consistent with \mathcal{C}_0 felt too confusing for us. The parentheses are here not to mix up the two. We apologize and beg the reader to bear with us on this one.

The following result appears as [47, Assumption 1.1.5]. It is useful to get rid of some redundancies when writing an element of \mathcal{C} .

PROPOSITION 1.2.36 (Extremal representation). *Any $C \in \mathcal{C}$ may be written as $C = A_0 \setminus \bigcup_{i=1}^k A_i$ where $k \in \mathbf{N}$ and $A_0, \dots, A_k \in \mathcal{A}$ are all maximal for \subseteq , i.e. for all $i \in \llbracket 1, k \rrbracket$, $A_i \subseteq A_0$ and for all $i, j \in \llbracket 1, k \rrbracket$, $A_i \subseteq A_j$ implies $i = j$. This representation, called the extremal representation of C , is unique up to a relabelling of A_1, \dots, A_k .*

Moreover, if $C \in \mathcal{C}_{(j)}$ for $j \in \mathbf{N}$, then its extremal representation is such that $k \leq j$.

Proof. Let us consider $C \in \mathcal{C}$ and write $C = A_0 \setminus \bigcup_{i=1}^j A_i$ where $j \in \mathbf{N}$ and $A_0, \dots, A_j \in \mathcal{A}$. Denote the set $\mathcal{A}' = \{A_0 \cap A_1, \dots, A_0 \cap A_j\}$. Then, the following writing

$$C = A_0 \setminus \bigcup_{\substack{A \in \mathcal{A}' \setminus \{\emptyset\}: \\ A \text{ maximal in } \mathcal{A}'}} A$$

is an extremal representation of C . Hence the existence. Moreover, since we took $C \in \mathcal{C}_{(j)}$ by definition and $\#\{A \in \mathcal{A}' \setminus \{\emptyset\} : A \text{ maximal in } \mathcal{A}'\} \leq j$, it also proves the second claim about the extremal representation.

Let us establish unicity up to relabelling. Suppose that $C = A_0 \setminus \bigcup_{i=1}^k A_i = A'_0 \setminus \bigcup_{j=1}^{k'} A'_j$ are two extremal representations of C . We have

$$A_0 = C \cup \bigcup_{i=1}^k A_i \subseteq A'_0 \cup \bigcup_{i=1}^k A_i.$$

Since A_0 cannot be included in any A_i for $i \in \llbracket 1, k \rrbracket$, the SHAPE condition (Definition 1.2.1) implies that $A_0 \subseteq A'_0$. By symmetry, we get $A_0 = A'_0$ and also $\bigcup_{i=1}^k A_i = \bigcup_{j=1}^{k'} A'_j$.

Let $i \in \llbracket 1, k \rrbracket$. Then $A_i \subseteq \bigcup_{j=1}^{k'} A'_j$. By SHAPE, there exists $\varphi(i) \in \llbracket 1, k' \rrbracket$ such that $A_i \subseteq A'_{\varphi(i)}$. By symmetry, there also exists a map ψ such that for all $j \in \llbracket 1, k' \rrbracket$, $A'_j \subseteq A_{\psi(j)}$. Since the A_i 's (resp. the A'_j 's) are maximal, $\psi \circ \varphi = \text{id}_{\llbracket 1, k \rrbracket}$ (resp. $\varphi \circ \psi = \text{id}_{\llbracket 1, k' \rrbracket}$). Hence the unicity up to relabelling. \square

EXAMPLE 1.2.37 (Increments classes for $\mathcal{T} = \mathbf{R}_+^p$). In Figure 1.7, we drew some increment sets for $\mathcal{T} = \mathbf{R}_+^p$ endowed with its indexing collection \mathcal{A} from Example 1.2.10.

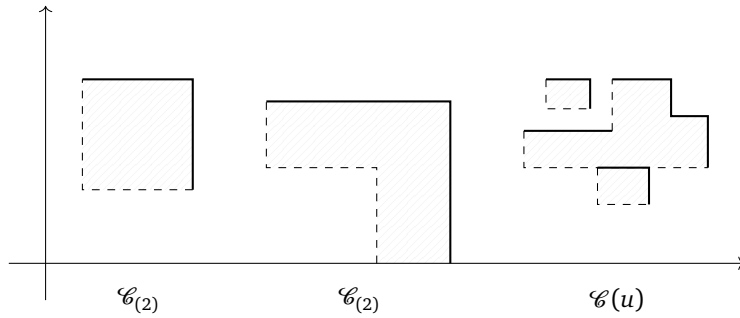


Figure 1.7: Increment sets for the indexing collection on \mathbf{R}_+^2 .

Remark that, even though \mathcal{C} is considered as the natural extension of the notion of ‘rectangular increments’ from the two-parameter setting, some elements of \mathcal{C} are far less ‘well-behaved’ (e.g. the middle one of Figure 1.7) than others (e.g. the leftmost one in Figure 1.7). Hence the need to introduce a subclass of \mathcal{C} which consists of increments sets among the best that \mathcal{C} has to offer. This class may be found in [47, Assumption 1.1.7].

DEFINITION 1.2.38 (Left neighborhoods \mathcal{C}^ℓ). For all finite subset $\mathcal{A}' \subseteq \mathcal{A}$, denote

$$\mathcal{C}^\ell(\mathcal{A}') = \left\{ A \setminus \bigcup_{\substack{A' \in \mathcal{A}' \\ A \not\subseteq A'}} A' : A \in \mathcal{A}' \right\}$$

The class of left neighborhoods is then defined as

$$\mathcal{C}^\ell = \bigcup_{n \in \mathbb{N}} \mathcal{C}^\ell(\mathcal{A}_n).$$

For all $t \in \bigcup_{C \in \mathcal{C}^\ell(\mathcal{A}_n)} C$, let $C_n(t)$ denote the unique element of $\mathcal{C}^\ell(\mathcal{A}_n)$ containing t .
For $t \notin \bigcup_{C \in \mathcal{C}^\ell(\mathcal{A}_n)} C$, set $C_n(t) = \mathcal{T}$.

The class $\mathcal{C}^\ell(\mathcal{A}')$ is made of the 'indivisible' (i.e. smallest for the inclusion), pairwise disjoint elements of \mathcal{C} that one can create from \mathcal{A}' . We will give a more lengthy comment about this class — especially $\mathcal{C}^\ell(\mathcal{A}_n)$ — in Section 1.6.

EXAMPLE 1.2.39 (Left neighborhoods for $\mathcal{T} = \mathbf{R}_+^p$). When $\mathcal{T} = \mathbf{R}_+^p$ is endowed with its usual indexing collection (Example 1.2.10), the elements of $\mathcal{C}^\ell(\mathcal{A}_n)$ are the 'hypercubes'

$$\mathbf{R}_+^p \cap \prod_{i=1}^p (d_i - 2^{-n}, d_i]$$

where the d_i 's are dyadics of order n in $[0, n]$. So the leftmost increment set of Figure 1.7 is a left neighborhood as long as its corner are dyadics.

1.2.4 Increment map and linear functional

We mentioned in Section 1.2.2 that the usual set-indexed formulation has been preferred in the rest of this thesis over the one of indexing semilattices. One reason of this preference is exposed here: any set-indexed process $X = \{X_A : A \in \mathcal{A}\}$ may be considered as a kind of cumulative distribution function of an additive map $\Delta X = \{\Delta X_U : U \in \mathcal{C}(u)\}$. This has two interesting consequences:

1. The concept of *increments* for X is well-defined (see Proposition 1.2.44) and generalizes its usual meaning in \mathbf{R}_+ - and \mathbf{R}_+^2 -indexed theories.
2. An integral against X becomes a perfectly natural object to consider and study.

This section is devoted to progressing towards those two goals. The first one is further studied in Chapter 2 whereas the construction for the second one is finished in Chapter 2 and studied in Chapter 3.

DEFINITION 1.2.40 (Simple functions). *The space of simple functions is the linear subspace \mathcal{E} of $\mathbf{R}^{\mathcal{T}}$ spanned by the indicator functions $\mathbb{1}_A$ where $A \in \mathcal{A}$.*

REMARK 1.2.41. *By the usual inclusion-exclusion formula, we know that for all $C = A_0 \setminus \bigcup_{i=1}^k A_i \in \mathcal{C}$,*

$$\mathbb{1}_C = \mathbb{1}_{A_0} - \sum_{i=1}^k (-1)^i \sum_{j_1 < \dots < j_i} \mathbb{1}_{A_0 \cap A_{j_1} \cap \dots \cap A_{j_i}}, \quad (1.2.7)$$

hence $\mathbb{1}_C$ belongs to \mathcal{E} . Since any element of $U \in \mathcal{C}(u)$ may be written as a disjoint union of elements in \mathcal{C} , its corresponding indicator function will also belong to \mathcal{E} . Hence $\{\mathbb{1}_U : U \in \mathcal{C}(u)\} \subseteq \mathcal{E}$.

The following straightforward result highlights an important aspect of \mathcal{C} : it enables us to write simple functions as sums of pairwise disjoint indicators. We give a statement a bit more specific that will be helpful to identify which elements of \mathcal{C} one may choose for such a representation.

PROPOSITION 1.2.42 (\mathcal{C} -representation of simple functions). *Let \mathcal{A}' be a finite subset of \mathcal{A}^* closed under intersections. Then, for all $f \in \text{span}(\mathbb{1}_A : A \in \mathcal{A}')$, there exists a unique $(a_C)_{C \in \mathcal{C}^\ell(\mathcal{A}')} \in \mathbf{R}^{\mathcal{C}^\ell(\mathcal{A}')}$ such that*

$$f = \sum_{C \in \mathcal{C}^\ell(\mathcal{A}')} a_C \mathbb{1}_C. \quad (1.2.8)$$

Proof. Consider \mathcal{A}' and f as stated above. The existence of the representation (1.2.8) directly follows from the fact that

$$\forall A \in \mathcal{A}', \quad \mathbb{1}_A = \sum_{C \in \mathcal{C}^\ell(\mathcal{A}')} \mathbb{1}_C.$$

As for unicity, first remark that any $C \in \mathcal{C}^\ell(\mathcal{A}')$ is non-empty since it would otherwise contradict the SHAPE condition. In particular, we may fix for any $C \in \mathcal{C}^\ell(\mathcal{A}')$ an element $t_C \in C$. Since the elements of $\mathcal{C}^\ell(\mathcal{A}')$ are pairwise disjoint, we actually have for all $C \in \mathcal{C}^\ell(\mathcal{A}')$, $a_C = f(t_C)$. The unicity follows. \square

Beware of the fact that the unicity in Proposition 1.2.42 depends on the choice of \mathcal{A}' . If one wanted a unicity independent from this choice, then one would have to rely on the class $\mathcal{C}(u)$ instead.

LEMMA 1.2.43. *The family $\{\mathbb{1}_A : A \in \mathcal{A}^*\}$ is linearly independent in \mathbf{R}^T , and thus forms a basis of \mathcal{E} .*

Proof. Assume the family is linearly dependent. We can write a dependence relation $\sum_{i=1}^k \alpha_i \mathbb{1}_{A_i} = 0$ where $k \in \mathbf{N}^*$, $\alpha_1, \dots, \alpha_k \in \mathbf{R}^*$ and $A_1, \dots, A_k \in \mathcal{A}^*$ are pairwise distinct.

By writing for all $j \in \llbracket 1, k \rrbracket$,

$$\mathbb{1}_{A_j} = \sum_{\substack{i=1 \\ i \neq j}}^k \frac{\alpha_i}{\alpha_j} \mathbb{1}_{A_i},$$

we deduce that

$$A_j \subseteq \bigcup_{\substack{i=1 \\ i \neq j}}^k A_i.$$

By the SHAPE condition (see Definition 1.2.1), we get that for all $j \in \llbracket 1, k \rrbracket$, there exists $i_j \in \llbracket 1, k \rrbracket \setminus \{j\}$ such that $A_j \subset A_{i_j}$. Notice that the inclusion is strict since the A_i 's are pairwise distinct.

Let us show that this brings a contradiction. Since $i_1 \neq 1$, instead of relabelling, we might as well suppose that $i_1 = 2$ so that $A_1 \subset A_2$. Suppose that $A_1 \subset \dots \subset A_j$, then $i_j > j$ since $i_j \neq j$ by definition and $i_j < j$ would yield the contradiction $A_j = A_{i_j}$. Hence, up to relabelling, we may suppose that $i_j = j + 1$ so that $A_1 \subset \dots \subset A_{j+1}$. By iteration, we get $A_1 \subset \dots \subset A_k$, but this is a contradiction since A_k must also be included in some A_i for $i < k$. \square

From this lemma, we deduce the existence of an *additive extension* Δh to $\mathcal{C}(u)$ of any map $h : \mathcal{A} \rightarrow \mathbf{R}$. Previously, it was known that SHAPE is a sufficient condition to the existence of such extensions (see discussion [47, p.25]), but it is also necessary to ensure the existence of all such extensions (Proposition 1.6.4).

PROPOSITION 1.2.44 (Increment map and linear functional). *Consider a map $h : \mathcal{A} \rightarrow \mathbf{R}$ such that $h(\emptyset) = 0$.*

◊ There exists a unique additive extension $\Delta h : \mathcal{C}(u) \rightarrow \mathbf{R}$ of h , i.e. such that $\Delta h|_{\mathcal{A}} = h$ and for all pairwise disjoint $U_1, U_2 \in \mathcal{C}(u)$, $\Delta h(U_1 \sqcup U_2) = \Delta h(U_1) + \Delta h(U_2)$.

The map Δh is called the increment map of h .

◊ There exists a unique linear map $\mathfrak{h} : \mathcal{E} \rightarrow \mathbf{R}$ such that $\mathfrak{h}(\mathbb{1}_A) = h(A)$ for all $A \in \mathcal{A}$. Moreover, $\mathfrak{h}(\mathbb{1}_U) = \Delta h(U)$ for all $U \in \mathcal{C}(u)$.

The map \mathfrak{h} is called the linear functional associated with h .

Proof. The existence and unicity of \mathfrak{h} are but a direct consequence of Lemma 1.2.43. For the increment map, according to Remark 1.2.41, we may define $\Delta h(U) = \mathfrak{h}(\mathbb{1}_U)$ for all $U \in \mathcal{C}(u)$, which is obviously additive.

It remains to prove uniqueness by induction. First, remark that Δh is uniquely determined on \mathcal{A} since we must have $\Delta h|_{\mathcal{A}} = h$. Suppose now that for a fixed integer $k \in \mathbf{N}$, Δh is uniquely determined on the class $\mathcal{C}_{(k)}$ (given in Definition 1.2.35) where we recall that $\mathcal{C}_{(0)} = \mathcal{A}$. Let us consider an element $C_{k+1} = A_0 \setminus \bigcup_{i=1}^{k+1} A_i \in \mathcal{C}_{(k+1)}$ and show that the value $\Delta h(C_{k+1})$ is determined by $\Delta h|_{\mathcal{C}_{(k)}}$.

Denote $C_k = A_0 \setminus \bigcup_{i=1}^k A_i$. Since $C_k = C_{k+1} \sqcup (A_{k+1} \cap C_k)$, the additivity of Δh tells us that

$$\Delta h(C_{k+1}) = \Delta h(C_k) - \Delta h(A_{k+1} \cap C_k)$$

where both C_k and $A_{k+1} \cap C_k$ actually belong to $\mathcal{C}_{(k)}$. Hence $\Delta h(C_{k+1})$ is uniquely determined by the induction hypothesis.

Thus Δh is uniquely determined on $\mathcal{C} = \bigcup_{k \in \mathbf{N}} \mathcal{C}_{(k)}$, but since any element of $\mathcal{C}(u)$ may be written as a disjoint union of elements of \mathcal{C} , Δh is uniquely determined on $\mathcal{C}(u)$ by additivity. \square

1.3 Indexing collections as measured spaces

Indexing collections are generally not ‘rigid’ nor symmetric. By that, we mean in particular that there is no group acting on \mathcal{A} which could give a notion of stationarity beloved by probabilists. Instead of a rigid group action, we capitalize on the set-indexed setting and consider a ‘softer’ concept by means of a measure \mathbf{m} .

This idea, first introduced by Herbin and Merzbach, will be further explored in Section 2.5.

1.3.1 Measure \mathbf{m} on $\mathcal{B} = \sigma(\mathcal{A})$

In the following, for any $D \subseteq D' \subseteq \mathcal{T}$, we denote the ‘segment’

$$[D, D'] = \{A \in \mathcal{A} : D \subseteq A \subseteq D'\}. \quad (1.3.1)$$

Recall that $\mathcal{B} = \sigma(\mathcal{A})$. Denote $\mathcal{M}(\mathcal{T}) = \{\mu \text{ measure on } (\mathcal{T}, \mathcal{B}) : \forall A \in \mathcal{A}, \mu(A) < \infty\}$ and for all $\mu \in \mathcal{M}(\mathcal{T})$, $\mathcal{B}_\mu = \{B \in \mathcal{B} : \mu(B) < \infty\}$.

DEFINITION 1.3.1 (Set-indexed compatible measure). *A measure \mathbf{m} on $(\mathcal{T}, \mathcal{B})$ is said to be (set-indexed) compatible if the following properties hold:*

1. (Local finiteness). For all $A \in \mathcal{A}$, $\mathbf{m}(A) < \infty$ (i.e. $\mathbf{m} \in \mathcal{M}(\mathcal{T})$).
2. (Midpoint property). For all $A_0 \subseteq A_1$ in \mathcal{A} , there exists $A \in [A_0, A_1]$ such that $\mathbf{m}(A \setminus A_0) = \mathbf{m}(A_1 \setminus A)$.

3. (Shrinking mesh property). For all $A \in \mathcal{A}$,

$$\max_{\substack{A_0, A_1 \in [\emptyset, A] \cap \mathcal{A}_n: \\ A_0 \text{ maximal proper subset of } A_1}} \mathbf{m}(A_1 \setminus A_0) \longrightarrow 0 \quad \text{as} \quad n \rightarrow \infty$$

where by ‘maximal proper subset’, we mean that $A_0 \subset A_1$ and there is no $A' \in \mathcal{A}_n$ such that $A_0 \subset A' \subset A_1$.

In the sequel, unless otherwise specified, \mathbf{m} will stand for such a compatible measure. Before moving on, let us give a few comments about Definition 1.3.1.

1. The local finiteness assumption ensures that $\mathcal{A} \subseteq \mathcal{B}_m$ so that it may be endowed with the metric d_m given by:

$$\forall B, B' \in \mathcal{B}_m, \quad d_m(B, B') = \mathbf{m}(B \Delta B') \quad (1.3.2)$$

where $B \Delta B' = (B \setminus B') \cup (B' \setminus B)$ is the symmetric set difference between B and B' . Topological properties will be studied in Section 1.3.3 (for d_m) and 1.4.4 (for more general metrics).

2. The midpoint property has this name due to the geometric fact it represents: for all $A \subseteq A'$ in \mathcal{A} , there exists a point in the middle of the segment $[A, A']$ with respect to d_m . Remark that it bestows the indexing collection \mathcal{A} with a ‘continuous’ quality, so all discrete examples from Section 1.2.2 are hereby cast away. It is an hypothesis reminiscent of the standard setting of geodesic spaces and its relevance will be further discussed in Sections 1.3.2 and 1.6.4.

3. The shrinking mesh property ensures that the ‘mesh’ created by \mathcal{A}_n shrinks correctly as n goes to ∞ . A similar, quantitative — hence more restrictive — hypothesis is found in [42, Assumption $\mathcal{H}_{\mathcal{A}}$]. A closely related assumption may also be found in [48, Assumption 2.4] where a lattice-indexed Poisson process is studied. Remark that the shrinking mesh property is weaker than what is usually required for Talagrand’s chaining argument (see [102, Section 2]).

REMARK 1.3.2. By a ‘metric’ d on a set \mathcal{D} , we actually mean ‘pseudo-metric’, i.e. we only require $d(t, t) = 0$ for all $t \in \mathcal{D}$ instead of the stronger separability axiom $d(s, t) = 0$ if and only if $s = t$ for all $s, t \in \mathcal{D}$. Such generality is welcome since any reasonable metric $d_{\mathcal{A}}$ on \mathcal{A} will verify $d_{\mathcal{A}}(\emptyset, \emptyset') = 0$ (see Definition 1.4.1) and important examples may only be pseudo-metrics (see Examples 1.4.2 and 1.4.3). In particular, this abuse explains why we called d_m a metric on \mathcal{B}_m .

EXAMPLE 1.3.3. Consider \mathcal{A} as any ‘continuous’ indexing collection taken from Examples 1.2.10, 1.2.25, 1.2.28 or 1.2.31. Then \mathbf{m} may be taken as any measure which admits a density with respect to the relevant ‘Lebesgue measure’ in each context.

Notable exceptions though are indexing collections obtained through compactification (see Corollary 1.2.22 and Example 1.2.29). In that case, the midpoint property and the shrinking mesh property do not generally hold.

1.3.2 Geodesic flows

As promised, let us now delve into the consequences of the midpoint property (Definition 1.3.1) for \mathbf{m} .

DEFINITION 1.3.4 (Flow, adapted from [47, Definition 5.1.1]). Let \mathcal{D} be a subset of $\mathcal{P}(\mathcal{T})$. A \mathcal{D} -flow is a non-decreasing map $\phi : [0, 1] \rightarrow \mathcal{D}$.

The notion of flow is well-known in classical set-indexed theory. One often requires more properties on the flow than monotonicity and this is where we split from the classical theory since our topological setting is a bit different. A more thorough comparison between the two is provided in Section 1.6.

DEFINITION 1.3.5 (Geodesic flow). Let \mathcal{D} be a subset of \mathcal{B}_m . A \mathcal{D} -flow $\phi : [0, 1] \rightarrow \mathcal{D}$ is a geodesic \mathcal{D} -flow if the following properties hold:

1. (Outer continuity). ϕ is outer continuous, i.e.

$$\forall t \in [0, 1), \quad \phi(t) = \bigcap_{s>t} \phi(s).$$

2. (Geodesic). There exists a constant $v(\phi) \geq 0$ called the speed of the geodesic such that

$$\forall t \in [0, 1], \quad \mathbf{m}(\phi(t) \setminus \phi(0)) = v(\phi)t. \quad (1.3.3)$$

The set of all geodesic \mathcal{D} -flows is denoted by $\Phi(\mathcal{D})$. A geodesic \mathcal{D} -flow ϕ is maximal if $\mathbf{m}(\phi(0)) = 0$ and $\phi(1)$ is maximal for \subseteq in \mathcal{D} . The set of all maximal geodesic \mathcal{D} -flows is denoted by $\Phi_{\max}(\mathcal{D})$.

The main use of flows is that they constitute a bridge between the set-indexed theory and the $[0, 1]$ -indexed one by means of projections.

DEFINITION 1.3.6 (Projection along a flow, adapted from [47, Section 5.1]). Let $X = \{X_A : A \in \mathcal{A}\}$ be a set-indexed process and ϕ be an \mathcal{A} -flow. We define the projection of X along ϕ to be the $[0, 1]$ -indexed process $X^\phi = \{X_{\phi(t)} : t \in [0, 1]\}$.

The projection along a \mathcal{D} -flow for $\mathcal{D} \in \{\mathcal{A}(u), \mathcal{C}, \mathcal{C}(u)\}$ is similarly defined and denoted by ΔX^ϕ .

As usual, let us give a few comments about Definition 1.3.5.

1. When studying a one-dimensional process, one often supposes — or already knows — that it has a version with right continuous sample paths. For the same reasons, when studying a set-indexed process, a property of ‘continuity from above’ is welcome. Outer continuity plays a role when this property holds by making sure that for any ‘continuous from above’ $X = \{X_A : A \in \mathcal{A}\}$ and outer continuous \mathcal{A} -flow ϕ , the projection X^ϕ is a right continuous process.
2. The reason why the second property is called as such is because any $\phi \in \Phi(\mathcal{D})$ is an actual geodesic with respect to d_m . Indeed, (1.3.3) is equivalent to

$$\forall s, t \in [0, 1], \quad d_m(\phi(s), \phi(t)) = v(\phi)|s - t|. \quad (1.3.4)$$

An idea related to projecting of along a geodesic flow may already be found in [40, Definition 2.3] under the terminology *m-standard flow*.

Before going on, we review well-known properties about d_m and set operators that will be useful later on.

LEMMA 1.3.7. For any $(B_k)_{k \in \mathbb{N}}, (B'_k)_{k \in \mathbb{N}}$ sequences in \mathcal{B}_m ,

$$d_m \left(\bigcap_{k \in \mathbb{N}} B_k, \bigcap_{k \in \mathbb{N}} B'_k \right) \leq \sum_{k=0}^{\infty} d_m(B_k, B'_k) \quad \text{and} \quad d_m \left(\bigcup_{k \in \mathbb{N}} B_k, \bigcup_{k \in \mathbb{N}} B'_k \right) \leq \sum_{k=0}^{\infty} d_m(B_k, B'_k).$$

Proof. Let $(B_k)_k$ and $(B'_k)_k$ be such sequences. Then,

$$\begin{aligned} d_m \left(\bigcap_k B_k, \bigcap_k B'_k \right) &= \mathbf{m} \left(\left(\bigcup_j B_j^c \right) \cap \left(\bigcap_k B'_k \right) \right) + \mathbf{m} \left(\left(\bigcap_k B_k \right) \cap \left(\bigcup_j B_j^c \right) \right) \\ &\leq \sum_j \mathbf{m} \left(B_j^c \cap \left(\bigcap_k B'_k \right) \right) + \sum_j \mathbf{m} \left(\left(\bigcap_k B_k \right) \cap B_j^c \right) \\ &\leq \sum_j \mathbf{m}(B'_j \setminus B_j) + \sum_j \mathbf{m}(B_j \setminus B'_j) \\ &= \sum_j d_m(B_j, B'_j). \end{aligned}$$

The inequality for \bigcup is may be immediately deduced from the previous one and the relation

$$d_m \left(\bigcup_{k \in \mathbb{N}} B_k, \bigcup_{k \in \mathbb{N}} B'_k \right) = d_m \left(\bigcap_{k \in \mathbb{N}} B_k^c, \bigcap_{k \in \mathbb{N}} B'_k{}^c \right).$$

□

REMARK 1.3.8. Lemma 1.3.7 implies the continuity of \cap and \cup with respect to d_m . Actually, we have

$$\forall B_1, B_2, B'_1, B'_2 \in \mathcal{B}_m, \quad d_m(B_1 \diamond B_2, B'_1 \diamond B'_2) \leq d_m(B_1, B'_1) + d_m(B_2, B'_2)$$

where \diamond is any of \cap, \cup or \setminus . In particular, \setminus is also continuous.

The following result has a classical alter ego in [47, Lemma 5.1.6].

PROPOSITION 1.3.9 (Existence of \mathcal{A} -flows). Let $A_0, A_1 \in \mathcal{A}$ such that $A_0 \subseteq A_1$. There exists a continuous (for d_m) \mathcal{A} -flow ϕ such that $\phi(0) = A_0$ and $\phi(1) = A_1$. Moreover, there exists $\phi \in \Phi(\mathcal{A})$ such that $A_0 \subseteq \phi(0)$, $\mathbf{m}(\phi(0) \setminus A_0) = 0$ and $\phi(1) = A_1$.

Proof. Let $A_0, A_1 \in \mathcal{A}$ such that $A_0 \subseteq A_1$. The first statement follows from the second one. Indeed, if $\phi \in \Phi(\mathcal{A})$ verifies the second statement, then the map

$$\begin{aligned} \phi' : [0, 1] &\longrightarrow \mathcal{A} \\ t &\longmapsto \begin{cases} A_0 & \text{if } t = 0, \\ \phi(t) & \text{if } t > 0 \end{cases} \end{aligned}$$

is a continuous \mathcal{A} -flow such that $\phi'(0) = A_0$ and $\phi'(1) = A_1$.

So only the second statement is left. For all $n \in \mathbb{N}$, denote by \mathcal{D}_n the set of dyadics of order n in $[0, 1]$ and $\mathcal{D} = \bigcup_{n \in \mathbb{N}} \mathcal{D}_n$. Starting from $\{A_0, A_1\}$ and repeatedly applying the midpoint property gives a family $\{A_d\}_{d \in \mathcal{D}}$ such that for all $d < d'$ in \mathcal{D} , $A_d \subset A_{d'}$ and for all $n \in \mathbb{N}^*$ and $d \in \mathcal{D}_n \setminus \mathcal{D}_{n-1}$, $\mathbf{m}(A_d \setminus A_{d-2^{-n}}) = \mathbf{m}(A_{d+2^{-n}} \setminus A_d)$. Define the map ϕ by

$$\forall t \in [0, 1], \quad \phi(t) = \begin{cases} \bigcap_{d \in \mathcal{D}: d > t} A_d & \text{if } t < 1, \\ A_1 & \text{if } t = 1. \end{cases}$$

It is obviously an outer continuous \mathcal{A} -flow. Moreover, a proof by induction readily shows that (1.3.3) holds for all $t \in \mathcal{D}$ where $\nu(\phi) = \mathbf{m}(A_1 \setminus A_0)$. By outer continuity of ϕ and \mathbf{m}^1 , (1.3.3) easily extends to all $t \in [0, 1]$. Lastly, $\phi(0) = \bigcap_{d \in \mathcal{D}: d > 0} A_d$ does contain A_0 and since $\mathbf{m}(A_1) < \infty$,

$$\mathbf{m}(\phi(0) \setminus A_0) = \lim_{\substack{d \in \mathcal{D}: \\ d \rightarrow 0^+}} \mathbf{m}(A_d \setminus A_0) = 0.$$

The result follows. \square

EXAMPLE 1.3.10. *Some might wonder whether one could improve on Proposition 1.3.9 and find $\phi \in \Phi(\mathcal{A})$ such that $A_0 = \phi(0)$ instead, but that is not true. Indeed, if $A_0 = \emptyset$ and $\mathbf{m}(A_1) > 0$, then for all $t > 0$, $\phi(t) \neq \emptyset$. Since \mathcal{A}^* is closed under intersections and ϕ is outer continuous, $\phi(0) = \bigcap_{t > 0} \phi(t) \neq \emptyset$.*

Even taking \emptyset out the picture does not solve everything. To see that, consider the set $\mathcal{T} = [0, 1] \cup \{0+\}$ where $0+$ is a point such that $0 < 0+$ and for all $t \in (0, 1]$, $0+ < t$. Suppose that \mathbf{m} is the measure on \mathcal{T} such that $\mathbf{m}([0, 1] \cap \cdot)$ is the usual Lebesgue measure on $[0, 1]$ and $\mathbf{m}(\{0+\}) = 0$. Then the same problem occurs for $A_0 = \{0\}$ and $A_1 = \mathcal{T}$ even though $A_0 \neq \emptyset$ since $\phi(0)$ will necessarily be equal to $\{0, 0+\}$.

PROPOSITION 1.3.11 (Existence of \mathcal{C} -flows). *Let $C_0, C_1 \in \mathcal{C}$ such that $C_0 \subseteq C_1$. There exists a continuous (for d_m) \mathcal{C} -flow ϕ such that $\phi(0) = C_0$ and $\phi(1) = C_1$. Moreover, if $C_0 = \emptyset$, we may choose ϕ so that for all $t \in [0, 1]$, $C_1 \setminus \phi(t) \in \mathcal{C}$.*

In general, geodesic \mathcal{C} -flows do not exist since outer continuity fails most of the time as the proof may suggest. The additional property at the end will prove to be useful much later to establish Lemma 2.5.14 when studying increment stationarity for set-indexed processes.

Such a result also holds for $\mathcal{C}(u)$ and the proof is a straightforward adaptation.

Proof. Let $C_0 = A_0 \setminus \bigcup_{i=1}^k A_i$ and $C_1 = A'_0 \setminus \bigcup_{j=1}^l A'_j$ in \mathcal{C} and such that $C_0 \subseteq C_1$. Without loss of generality, we may suppose that $k, l \in \mathbb{N}^*$ since if it is not the case, one may add the empty set. By the SHAPE condition, we know that $A_0 \subseteq A'_0$. Using Proposition 1.3.9, we denote ϕ_0 a continuous \mathcal{A} -flow from A_0 to A'_0 and for all $(i, j) \in \llbracket 1, k \rrbracket \times \llbracket 1, l \rrbracket$, ϕ_{ij} a continuous \mathcal{A} -flow from $A_i \cap A'_j$ to A_i . Then define

$$\phi : [0, 1] \rightarrow \mathcal{C} \\ t \mapsto \begin{cases} A_0 \setminus \bigcup_{\substack{1 \leq i \leq k \\ 1 \leq j \leq l}} \phi_{ij}(1-2t) & \text{if } t \leq 1/2, \\ \phi_0(2t-1) \setminus \bigcup_{1 \leq j \leq l} A'_j & \text{if } t > 1/2. \end{cases}$$

The map ϕ is obviously a \mathcal{C} -flow such that $\phi(0) = C_0$ and $\phi(1) = C_1$. It is also continuous on $[0, 1] \setminus \{1/2\}$ by continuity of the set operations with respect to d_m (see Remark 1.3.8). As for the continuity at $1/2$, we just need to show that

$$\lim_{t \rightarrow 1/2^+} \phi(t) = \phi(1/2) = A_0 \setminus \bigcup_{\substack{1 \leq i \leq k \\ 1 \leq j \leq l}} (A_i \cap A'_j) = A_0 \setminus \bigcup_{1 \leq j \leq l} A'_j = \lim_{t \rightarrow 1/2^-} \phi(t)$$

which is equivalent to show that

$$\bigcup_{\substack{1 \leq i \leq k \\ 1 \leq j \leq l}} (A_0 \cap A_i \cap A'_j) = \bigcup_{1 \leq j \leq l} (A_0 \cap A'_j). \quad (1.3.5)$$

¹By outer continuity for \mathbf{m} , we mean that for all sequence $(B_k)_{k \in \mathbb{N}}$ in \mathcal{B}_m decreasing to B , $\mathbf{m}(B_k) \downarrow \mathbf{m}(B)$ as $k \rightarrow \infty$.

The direct inclusion in (1.3.5) is trivial. Let $t \in A_0 \cap A'_j$ for some $j \in \llbracket 1, l \rrbracket$. Since $t \in A_0$, we know that either $t \in \bigcup_{1 \leq i \leq k} A_i$ or $t \in C_0$. But the latter is impossible since it would imply that $t \in C_1$ and thus $t \notin A'_j$. Hence $t \in \bigcup_{1 \leq i \leq k} A_i$, which proves (1.3.5). Hence ϕ is continuous.

Now, suppose that $C_0 = \emptyset$. Instead of the previous flow, we may consider a continuous \mathcal{A} -flow ψ from \emptyset to A'_0 and define for all $t \in [0, 1]$, $\psi(t) = A'_0 \setminus (\psi'(1-t) \cup \bigcup_{1 \leq j \leq l} A'_j)$. Since for all $t \in [0, 1]$, we have $C_1 \setminus \psi(t) = C_1 \cap \psi'(1-t) \in \mathcal{C}$: this flow answers our additional requirement. \square

PROPOSITION 1.3.12 (Interpolation of a chain in $\mathcal{A}(u)$). *For any $k \in \mathbf{N}$ and finite chain $U_0 \subseteq \dots \subseteq U_k$ in $\mathcal{A}(u)$, there exists a continuous (for d_m) $\mathcal{A}(u)$ -flow ϕ such that $\{U_0, \dots, U_k\} \subseteq \phi([0, 1])$. Moreover, there also exists $\phi \in \Phi(\mathcal{A}(u))$ and $t_0 \leq \dots \leq t_k$ in $[0, 1]$ such that for all $i \in \llbracket 0, k \rrbracket$, $U_i \subseteq \phi(t_i)$ and $\mathbf{m}(\phi(t_i) \setminus U_i) = 0$.*

Proof. For the first statement, we only prove the case $k = 1$ since the general one is deduced by ‘piecing together’ several flows. Write $U_0 = \bigcup_{i=1}^j A_{0i}$ and $U_1 = \bigcup_{i=0}^{j'} A_{1i}$. Using the SHAPE condition and without loss of generality, we may suppose that $j = j'$ and that for all $i \in \llbracket 1, j \rrbracket$, $A_{0i} \subseteq A_{1i}$. According to Proposition 1.3.9, for all $i \in \llbracket 1, j \rrbracket$, there exists a continuous \mathcal{A} -flow ϕ_i such that $\phi_i(0) = A_{0i}$ and $\phi_i(1) = A_{1i}$. Then the map $\phi = \bigcup_{i=1}^j \phi_i$ is obviously an $\mathcal{A}(u)$ -flow. According to Remark 1.3.8, it is also continuous. The first statement follows.

As a direct consequence, this result tells in particular that the midpoint property of Definition 1.3.1 also holds for elements in $\mathcal{A}(u)$. More precisely, if $\mathcal{A}_{(j)}(u)$ denotes the class of finite unions of at most j elements of \mathcal{A} , then for all $U_0 \subseteq U_1$ in $\mathcal{A}_{(j)}(u)$, there exists $U \in \mathcal{A}_{(j)}(u)$ such that $U_0 \subseteq U \subseteq U_1$ and $\mathbf{m}(U \setminus U_0) = \mathbf{m}(U_1 \setminus U)$. Using this property and the fact that $\mathcal{A}_{(j)}(u)$ is closed under monotone intersections, one may redo the proof of the second statement of Proposition 1.3.9. The second statement follows. \square

The next result says that knowing the finite-dimensional distributions of a set-indexed process is equivalent to knowing those of each projection along a geodesic flow in $\mathcal{A}(u)$. A similar result in the classical setting can already be found in [46, Lemma 6].

THEOREM 1.3.13 (Characterization by projection along $\mathcal{A}(u)$ -flows). *Let X, Y be two set-indexed processes. Then,*

$$X \stackrel{\text{fdd}}{=} Y \iff \text{For all continuous (for } d_m) \mathcal{A}(u)\text{-flow } \phi, \Delta X^\phi \stackrel{\text{fdd}}{=} \Delta Y^\phi.$$

Suppose that for all $C \in \mathcal{C}$, $\mathbf{m}(C) = 0$ implies $\Delta X_C = 0 = \Delta Y_C$ a.s. Then,

$$X \stackrel{\text{fdd}}{=} Y \iff \forall \phi \in \Phi(\mathcal{A}(u)), \Delta X^\phi \stackrel{\text{fdd}}{=} \Delta Y^\phi.$$

REMARK 1.3.14. *In Theorem 1.3.13, $\Phi(\mathcal{A}(u))$ may be replaced by $\bigcup_{n \in \mathbf{N}} \Phi_{\max}(\{U \in \mathcal{A}(u) : U \subseteq \mathcal{T}_n\})$. Indeed, any $\phi \in \Phi(\mathcal{A}(u))$ is such that $\phi(1) \subseteq \mathcal{T}_n$ for some $n \in \mathbf{N}$, so those maximal flows are enough to cover all elements in $\mathcal{A}(u)$. This will be used to establish representation theorems for certain classes processes (see Theorems 2.5.30, 2.5.31 and 3.3.8).*

Proof. The first statement being simpler to prove, we only prove the second one. More precisely, we solely focus on the ‘if’ part since the ‘only if’ part is trivial. Suppose that for all $\phi \in \Phi(\mathcal{A}(u))$, $\Delta X^\phi \stackrel{\text{fdd}}{=} \Delta Y^\phi$.

Consider $k \in \mathbf{N}$, $A_0, \dots, A_k \in \mathcal{A}$ and let us show that $(X_{A_0}, \dots, X_{A_k}) \stackrel{\text{law}}{=} (Y_{A_0}, \dots, Y_{A_k})$. Without loss of generality, we may also suppose that the family $\{A_0, \dots, A_k\}$ is made of pairwise distinct

elements, closed under intersection and *consistently* ordered, i.e. $A_i \subset A_j$ implies $i < j$ for all $i, j \in \llbracket 0, k \rrbracket$. From there, define for all $j \in \llbracket 0, k \rrbracket$, $U_j = \bigcup_{0 \leq i \leq j} A_i$. They constitute a chain of elements in $\mathcal{A}(u)$.

Then we may consider $\phi \in \Phi(\mathcal{A}(u))$ and $t_0 \leq \dots \leq t_k$ in $[0, 1]$ just like in Proposition 1.3.12.

Let $i \in \llbracket 0, k \rrbracket$. The set $V_i = \phi(t_i) \setminus U_i$ belongs to $\mathcal{C}(u)$ and is such that $\mathbf{m}(V_i) = 0$. Since any element in $\mathcal{C}(u)$ may be written as a union of elements in \mathcal{C} , we get that $\Delta X_{V_i} = 0$ a.s. In particular,

$$(\Delta X_{U_0}, \dots, \Delta X_{U_k}) = (\Delta X_{\phi(t_0)}, \dots, \Delta X_{\phi(t_k)}) \quad \text{a.s.} \quad (1.3.6)$$

and likewise for Y .

Moreover, since $\Delta X^\phi \stackrel{\text{fdd}}{=} \Delta Y^\phi$, we obtain that

$$(\Delta X_{\phi(t_0)}, \dots, \Delta X_{\phi(t_k)}) \stackrel{\text{law}}{=} (\Delta Y_{\phi(t_0)}, \dots, \Delta Y_{\phi(t_k)}). \quad (1.3.7)$$

Combining (1.3.6) and (1.3.7) yields

$$(\Delta X_{U_0}, \dots, \Delta X_{U_k}) \stackrel{\text{law}}{=} (\Delta Y_{U_0}, \dots, \Delta Y_{U_k}) \quad (1.3.8)$$

Let us show that one can reconstitute the A_i 's from the U_i 's by proving that $\text{span}(\mathbb{1}_{A_i} : i \leq j) = \text{span}(\mathbb{1}_{U_i} : i \leq j)$ by induction on j . The converse inclusion being a straightforward consequence of the inclusion-exclusion formula (and not the one needed anyway), we only prove the direct inclusion.

The case $j = 0$ is true since $A_0 = U_0$. Suppose the result holds up to $j-1$ for $j \in \llbracket 1, k \rrbracket$. Then

$$\begin{aligned} A_j &= [A_j \cap (\bigcup_{i < j} A_i)] \cup [A_j \setminus (\bigcup_{i < j} A_i)] \\ &= [\bigcup_{i < j} (A_i \cap A_j)] \cup [U_j \setminus U_{j-1}]. \end{aligned}$$

Since $\{A_0, \dots, A_k\}$ is closed under intersections and has been ordered consistently, by the induction hypothesis, we get that for all $i < j$, $\mathbb{1}_{A_i \cap A_j} \in \text{span}(\mathbb{1}_{U_k} : k \leq j)$. Hence, due to the expression of A_j above and the inclusion-exclusion formula, we have $\mathbb{1}_{A_j} \in \text{span}(\mathbb{1}_{U_i} : i \leq j)$. In particular, there exists an endomorphism $\varphi : \mathbf{R}^{k+1} \rightarrow \mathbf{R}^{k+1}$ such that

$$(\mathbb{1}_{A_0}, \dots, \mathbb{1}_{A_k}) = \varphi(\mathbb{1}_{U_0}, \dots, \mathbb{1}_{U_k}). \quad (1.3.9)$$

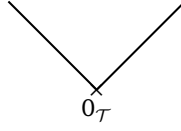
Combining (1.3.8) and (1.3.9), we get

$$(X_{A_0}, \dots, X_{A_k}) = \varphi(\Delta X_{U_0}, \dots, \Delta X_{U_k}) \stackrel{\text{law}}{=} \varphi(\Delta Y_{U_0}, \dots, \Delta Y_{U_k}) = (Y_{A_0}, \dots, Y_{A_k}).$$

Hence $X \stackrel{\text{fdd}}{=} Y$. □

EXAMPLE 1.3.15 (Projections along \mathcal{A} -flows are not enough). *To our opinion, Theorem 1.3.13 is the main reason why the class $\mathcal{A}(u)$ is important in itself and not just as an intermediate step to define other classes. Indeed, the statement of the theorem becomes false once one replaces $\mathcal{A}(u)$ by \mathcal{A} . To see that, suppose that \mathcal{T} and \mathcal{A} are the ones from Example 1.2.25 with discrete tree $\mathcal{U} = \{0_{\mathcal{U}}, 0, 1\}$ as illustrated in Figure 1.8.*

For notation's sake, we simply write $\mathcal{T} = \{a + ib \in \mathbf{C} : a, b \in [0, 1] \text{ and } a \times b = 0\}$, $\mathcal{A} = \{\emptyset\} \cup \{[0, a] : a \in [0, 1]\} \cup \{[0, ib] : b \in [0, 1]\}$ and \mathbf{m} the restriction to \mathcal{T} of the Hausdorff

Figure 1.8: Continuous tree \mathcal{T} obtained from $\mathcal{U} = \{0_{\mathcal{U}}, 0, 1\}$.

measure \mathcal{H}^1 on \mathbf{C} . Then consider two independent Brownian motions W and W' on $[0, 1]$ and define the set-indexed processes X and Y by

$$\forall A \in \mathcal{A}, \quad X_A = \begin{cases} 0 & \text{if } A \in \{\emptyset, \{0\}\}, \\ W_a & \text{if } A = [0, a] \text{ where } a \in [0, 1], \\ W'_b & \text{if } A = [0, ib] \text{ where } b \in [0, 1], \end{cases}$$

and

$$\forall A \in \mathcal{A}, \quad Y_A = \begin{cases} 0 & \text{if } A \in \{\emptyset, \{0\}\}, \\ W_a & \text{if } A = [0, a] \text{ where } a \in (0, 1], \\ W_b & \text{if } A = [0, ib] \text{ where } b \in (0, 1], \end{cases}$$

For all \mathcal{A} -flow ϕ , $X^\phi \stackrel{\text{fdd}}{=} Y^\phi$ while X and Y cannot have the same finite-dimensional distributions. Indeed, we have $\Delta X_{[0,1] \cup [0,i]} = W_1 + W'_1 \sim N(0, 2)$ while $\Delta Y_{[0,1] \cup [0,i]} = 2W_1 \sim N(0, 4)$.

Since this particular indexing collection can also be turned into a classical one (see Example 1.6.3), this constitutes a counter-example to [40, Theorem 3.4] which claims that a set-indexed Brownian motion (siBm) is the only set-indexed process such that its projections along \mathcal{A} -flows are (time-changed) Brownian motions. However, [73, Theorem 1] proves that the result holds if \mathcal{A} -flows are replaced by $\mathcal{A}(u)$ -flows.

Here, the siBm happens to be the process X . It is a set-indexed generalization of the one-dimensional Brownian motion and the \mathbf{R}_+^p -indexed Brownian sheet [4, Equation (1.4.15)]. It has been introduced together with its fractional counterpart in [39] and has been studied in [39, 41, 42]. Like in the one-dimensional case, the siBm plays an important role in set-indexed theory and a central limit theorem [47, Theorem 9.1.2] still holds.

1.3.3 Topological consequences

Section 1.2.3 introduced a whole lot of other classes than \mathcal{A} . The hypotheses on \mathbf{m} (Definition 1.3.1) do have some topological consequences on those classes once they are endowed with the metric d_m given in (1.3.2). Here, we derive those specific consequences while the others — valid for more general metrics than d_m — will be treated in Section 1.4.4.

Approximation of Borel sets and functions

In the case when $\mathcal{T} = \mathbf{R}$ and $\mathbf{m} = \mathbf{Leb}$ is the Lebesgue measure, Littlewood's principles [69] acts as guides to intuition with regards to measure theory and how can one apprehend some of its hardest concepts. In particular, the first principle tells that any Borel set $B \in \mathcal{B}_{\mathbf{Leb}}$ can be approximated as a finite union of intervals with respect to the metric $d_{\mathbf{Leb}} = \mathbf{Leb}(\Delta \cdot)$. This principle is much more general and still holds in our setting when finite unions of segments are replaced by the elements of $\mathcal{C}(u)$ and $d_{\mathbf{Leb}}$ by d_m .

Another interesting aspect of d_m is its link with the metric of $L^p(\mathbf{m}) = L^p(\mathcal{T}, \mathcal{B}, \mathbf{m})$ for $p \geq 1$, namely

$$\forall B, B' \in \mathcal{B}_m, \quad \|\mathbb{1}_B - \mathbb{1}_{B'}\|_{L^p(\mathbf{m})} = d_m(B, B')^{1/p}. \quad (1.3.10)$$

Our goal in this section is to prove a result similar to Littlewood's and apply it to approximate functions in $L^p(\mathbf{m})$.

LEMMA 1.3.16 (σ -finiteness). *The increasing sequence $(\mathcal{T}_n)_{n \in \mathbf{N}} \in \mathcal{A}(u)^{\mathbf{N}}$ given by $\mathcal{T}_n = \bigcup_{A \in \mathcal{A}_n} A$ is such that $\mathcal{T} = \bigcup_{n \in \mathbf{N}} \mathcal{T}_n$ and $\mathbf{m}(\mathcal{T}_n) < \infty$ for all $n \in \mathbf{N}$. In particular, \mathbf{m} is σ -finite.*

Proof. $\mathcal{T} = \bigcup_{n \in \mathbf{N}} \mathcal{T}_n$ directly follows from the TIP bijection. The σ -finiteness directly follows from the local finiteness of \mathbf{m} and the fact that \mathcal{A}_n is finite. \square

THEOREM 1.3.17 (Set-indexed Littlewood's first principle). *The metric space (\mathcal{B}_m, d_m) is complete and $\mathcal{C}(u)$ is a dense subset.*

Proof. Due to Lemma 1.3.16, we only prove the results when $\mathcal{T} \in \mathcal{A}(u)$. The adaptation to the σ -finite case is straightforward.

Let us show that \mathcal{B}_m is complete. Let $(B_k)_{k \in \mathbf{N}}$ be a Cauchy sequence in \mathcal{B}_m . According to (1.3.10), $(\mathbb{1}_{B_k})_{k \in \mathbf{N}}$ is a Cauchy sequence in $L^1(\mathbf{m})$. Since $L^1(\mathbf{m})$ is complete [9, Theorem 13.5], this sequence has a limit. We claim that this limit is necessarily of the form $\mathbb{1}_B$ where $B \in \mathcal{B}_m$. Indeed, since $\mathbf{m}(\mathcal{T}) < \infty$, we may apply Borel-Cantelli's lemma and obtain that $(\mathbb{1}_{B_k})_{k \in \mathbf{N}}$ has a subsequence that converges \mathbf{m} -a.e. Its limit must thus take its values in $\{0, 1\}$ \mathbf{m} -a.e., i.e. be equal to $\mathbb{1}_B$ for some $B \in \mathcal{B}_m$. Applying (1.3.10) once more yields that $d_m(B_k, B) \rightarrow 0$ as $k \rightarrow \infty$. The completeness follows.

Let us prove that $\mathcal{C}(u)$ is a dense subset of \mathcal{B}_m . Denote by $\overline{\mathcal{C}(u)}$ its closure in \mathcal{B}_m . Since $\mathcal{A} \subseteq \overline{\mathcal{C}(u)}$ and $\sigma(\mathcal{A}) = \mathcal{B}_m$ — recall that $\mathbf{m}(\mathcal{T}) < \infty$ — it is sufficient to prove that $\overline{\mathcal{C}(u)}$ is a σ -algebra. Since $\emptyset \in \mathcal{A}$, then $\emptyset \in \overline{\mathcal{C}(u)}$. Stability by complement is but a consequence of the relation $d_m(B_1 \Delta B_2) = d_m(B_1^c \Delta B_2^c)$ for all $B_1, B_2 \in \mathcal{B}_m$. As for stability under countable intersections, consider a sequence $(B_k)_{k \in \mathbf{N}}$ in $\overline{\mathcal{C}(u)}$ and $\varepsilon > 0$. For all $k \in \mathbf{N}$, consider $U_k \in \mathcal{C}(u)$ such that $d_m(B_k, U_k) \leq \varepsilon 2^{-(k+1)}$. By Lemma 1.3.7,

$$d_m \left(\bigcap_{k \in \mathbf{N}} B_k, \bigcap_{k \in \mathbf{N}} U_k \right) \leq \varepsilon/2. \quad (1.3.11)$$

Moreover, since $\mathbf{m}(\mathcal{T}) < \infty$, the monotone continuity of \mathbf{m} implies that

$$d_m \left(\bigcap_{k \in \mathbf{N}} U_k, \bigcap_{i \leq j} U_i \right) \rightarrow 0 \quad \text{as } j \rightarrow \infty. \quad (1.3.12)$$

Hence (1.3.11), (1.3.12) and the triangle inequality imply that for some j big enough,

$$d_m \left(\bigcap_{k \in \mathbf{N}} B_k, \bigcap_{i \leq j} U_i \right) \leq \varepsilon.$$

Since $\mathcal{C}(u)$ is closed under finite intersections, $\bigcap_{i \leq j} U_i \in \mathcal{C}(u)$. Hence $\bigcap_{k \in \mathbf{N}} B_k \in \overline{\mathcal{C}(u)}$ and the result follows. \square

REMARK 1.3.18. *Actually, \mathcal{B}_m is also separable — and hence Polish — since the density of the countable class $\mathcal{C}^\ell(u) = \left\{ \bigcup_{i=1}^k C_i : k \in \mathbf{N}, C_1, \dots, C_k \in \mathcal{C}^\ell \right\}$ is easily deduced from that of $\mathcal{C}(u)$.*

COROLLARY 1.3.19 (Density of simple functions). *For any $p \geq 1$, $\mathcal{E} = \text{span}(\mathbb{1}_A : A \in \mathcal{A})$ is dense in $L^p(\mathbf{m})$.*

Proof. Since $\text{span}(\mathbb{1}_B : B \in \mathcal{B}_m)$ is dense in $L^p(\mathbf{m})$ [9, Theorem 13.8], we just need to show that for all $B \in \mathcal{B}_m$, $\mathbb{1}_B$ belongs to the closure of \mathcal{E} in $L^p(\mathbf{m})$, but that is exactly a consequence of Theorem 1.3.17 and (1.3.10). \square

Path-connectedness

As a question of independent interest, we use the existence of flows to prove path-connectedness results for several classes. Even though those results will not be used elsewhere, we hope that it will strengthen one's intuition about indexing collections.

PROPOSITION 1.3.20. *Any of the following classes \mathcal{A} , $\mathcal{A}(u)$, $\mathcal{C}_{(k)}$ ($k \in \mathbf{N}$), \mathcal{C} , $\mathcal{C}(u)$ and \mathcal{B}_m is path-connected for d_m .*

Beginning of the proof of Proposition 1.3.20. Path-connectedness for \mathcal{A} and $\mathcal{A}(u)$ is a consequence of Propositions 1.3.9 and 1.3.12 respectively.

Let us fix $k \in \mathbf{N}$ and prove that $\mathcal{C}_{(k)}$ is path-connected. Consider $C = A_0 \setminus \bigcup_{i=1}^k A_i$ and $C' = A'_0 \setminus \bigcup_{i=1}^k A'_i$ in $\mathcal{C}_{(k)}$. For all $i \in \llbracket 0, k \rrbracket$, consider a continuous path $\phi_i : [0, 1] \rightarrow \mathcal{A}$ such that $\phi_i(0) = A_i$ and $\phi_i(1) = A'_i$. Then the map $\psi = \phi_0 \setminus \bigcup_{i=1}^k \phi_i$ is such that $\psi(0) = C$ and $\psi(1) = C'$. Due to Remark 1.3.8, ψ is also continuous. Hence $\mathcal{C}_{(k)}$ is path-connected.

Then $\mathcal{C} = \bigcup_{k \in \mathbf{N}} \mathcal{C}_{(k)}$ is path-connected as a non-decreasing union of path-connected sets.

The class $\mathcal{C}(u)$ is path-connected since \mathcal{C} is and \cup is continuous (Remark 1.3.8). \square

For \mathcal{B}_m , we need a bit more effort concerning the kind of paths connecting sets in \mathcal{C} and $\mathcal{C}(u)$. Indeed, Theorem 1.3.17 tells that \mathcal{B}_m is the closure of $\mathcal{C}(u)$. However the result does not follow directly since the closure of a path-connected space is not necessarily path-connected itself. A well-known counter-example is the closure in \mathbf{R}^2 of $\{(t, \sin(1/t)) : t \in (0, 1]\}$ which is not path-connected. It teaches us that one has to first control the oscillations of the connecting paths.

LEMMA 1.3.21. *For any $U, U' \in \mathcal{C}(u)$, there exists a continuous (for d_m) map $\phi : [0, 1] \rightarrow \mathcal{C}(u)$ such that the following holds:*

1. $[0, 1] \ni t \mapsto \phi(1/2 - t/2)$ is a $\mathcal{C}(u)$ -flow from $U \cap U'$ to U ,
2. $[0, 1] \ni t \mapsto \phi(1/2 + t/2)$ is a $\mathcal{C}(u)$ -flow from $U \cap U'$ to U' .

Proof. This lemma is just a convoluted way of putting together two continuous $\mathcal{C}(u)$ -flows. Use Proposition 1.3.11 to create a $\mathcal{C}(u)$ -flow ψ (resp. ψ') from $U \cap U'$ to U (resp. U'). Then, the map defined by

$$\begin{aligned} \phi : [0, 1] &\longrightarrow \mathcal{C}(u) \\ t &\longmapsto \begin{cases} \psi(1-2t) & \text{if } t \leq 1/2, \\ \psi'(2t-1) & \text{if } t > 1/2 \end{cases} \end{aligned}$$

answers our needs. \square

End of the proof of Proposition 1.3.20. Let us prove that \mathcal{B}_m is path-connected. Let $B \in \mathcal{B}_m$ and let us show that B is connected to \emptyset . Due to Theorem 1.3.17, there is a sequence $(U_k)_{k \in \mathbf{N}^*}$ in $\mathcal{C}(u)$ such that $U_1 = \emptyset$ and $d_m(U_k, B) \rightarrow 0$ as $k \rightarrow \infty$. For all $k \in \mathbf{N}^*$, let $x_k = 1 - 1/k$ and

$\phi_k : [0, 1] \rightarrow \mathcal{C}(u)$ be the map given by Lemma 1.3.21 for $U = U_k$ and $U' = U_{k+1}$. Then define for all $k \in \mathbf{N}^*$ the map

$$\begin{aligned} \psi_k : [x_k, x_{k+1}) &\longrightarrow \mathcal{C}(u) \\ t &\longmapsto \phi_k((t - x_k)/(x_{k+1} - x_k)) \end{aligned}$$

as well as

$$\begin{aligned} \phi : [0, 1) &\longrightarrow \mathcal{C}(u) \\ t &\longmapsto \psi_k(t) \text{ if } t \in [x_k, x_{k+1}) \text{ where } k \in \mathbf{N}^*. \end{aligned}$$

By construction, ϕ is continuous. Let us show that it may be extended by continuity by posing $\phi(1) = B$. Let $t \in [0, 1)$ and denote by $k = k(t)$ the only positive integer such that $t \in [x_k, x_{k+1})$. Then, since for all $s \in [0, 1]$, $(U_k \cap U_{k+1}) \subseteq \phi_k(s) \subseteq (U_k \cup U_{k+1})$, we have

$$\begin{aligned} d_m(\phi(t), U_k \cap U_{k+1}) &\leq \max_{s \in [0, 1]} d_m(\phi_k(s), U_k \cap U_{k+1}) \\ &= \max_{s \in [0, 1]} m(\phi_k(s) \setminus (U_k \cap U_{k+1})) \\ &\leq m((U_k \cup U_{k+1}) \setminus (U_k \cap U_{k+1})) \\ &= d_m(U_k, U_{k+1}). \end{aligned}$$

Thus, since $U_k \rightarrow B$ as $k \rightarrow \infty$ and \cap is continuous (Remark 1.3.8),

$$d_m(\phi(t), B) \leq d_m(\phi(t), U_k \cap U_{k+1}) + d_m(U_k \cap U_{k+1}, B) \longrightarrow 0 \quad \text{as } t \rightarrow 1^-.$$

So setting $\phi(1) = B$ defines a continuous path from \emptyset to B . Hence \mathcal{B}_m is path-connected. \square

1.4 Indexing collections as metric spaces

1.4.1 Metric $d_{\mathcal{A}}$ on \mathcal{A}

Since we want to have a look at the regularity of set-indexed processes, we require a metric $d_{\mathcal{A}}$ on \mathcal{A} . Moreover, we want $d_{\mathcal{A}}$ to interact well with the already existing order structure of the indexing collection \mathcal{A} . A similar approach has been undertaken in [42] to obtain a set-indexed version of Kolmogorov-Čentsov's regularity theorem, but here we do not require any quantitative hypothesis.

Some might say that the metric d_m introduced in (1.3.2) should be enough to our purposes and they would be mostly right. However, d_m is not always an intuitive metric when trying to come up with examples. For instance, even in the simple case when $\mathcal{T} = \mathbf{R}_+^2$ (Example 1.2.10) and m is the Lebesgue measure, the balls for d_m are not exactly well-behaved per say. Indeed, when the center is not on the axes, a ball of small radius is delimited by four branches of hyperbolae (see Figure 1.9).

Remark that due to the TIP assumption (Definition 1.2.1), it is equivalent to either give a metric $d_{\mathcal{A}}$ on \mathcal{A} or a metric $d_{\mathcal{T}}$ on \mathcal{T} and to define the other through the relations:

$$\begin{aligned} \forall s, t \in \mathcal{T}, \quad d_{\mathcal{T}}(s, t) &= d_{\mathcal{A}}(A(s), A(t)), \\ \forall A, A' \in \mathcal{A}, \quad d_{\mathcal{A}}(A, A') &= d_{\mathcal{T}}(\mathbf{t}(A), \mathbf{t}(A')). \end{aligned} \tag{1.4.1}$$

with the convention that $\mathbf{t}(\emptyset) = \mathbf{t}(\emptyset') (= 0_{\mathcal{T}})$.

This pushes further the correspondence established by Theorem 1.2.5. We emphasize that even though \mathcal{T} might sometimes be endowed with a natural metric, if $d_{\mathcal{A}}$ is given, then $d_{\mathcal{T}}$ will always be defined by (1.4.1) and vice versa.

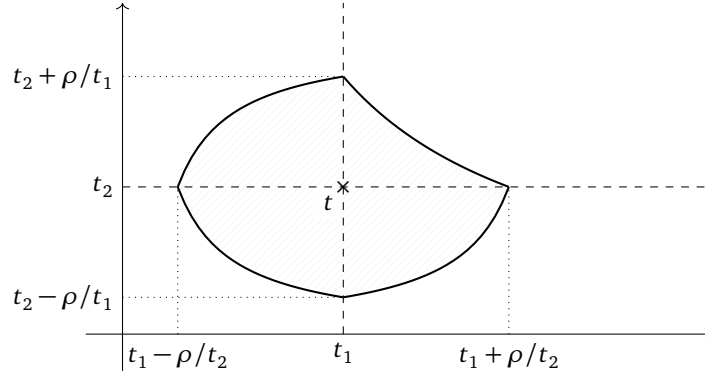


Figure 1.9: Ball of center $t \in (\mathbb{R}_+^*)^2$ and radius $\rho \in (0, t_1 t_2]$ for d_m .

DEFINITION 1.4.1 (Set-indexed compatible metric). A metric $d_{\mathcal{A}}$ on \mathcal{A} is said to be (set-indexed) compatible if the following properties hold:

1. (Contractivity). For any $A, A', A'' \in \mathcal{A}$, $d_{\mathcal{A}}(A \cap A'', A' \cap A'') \leq d_{\mathcal{A}}(A, A')$.
2. (Outer continuity). For any non-increasing sequence $(A_k)_{k \in \mathbb{N}}$ in \mathcal{A} , $d_{\mathcal{A}}(A_k, A) \rightarrow 0$ as $k \rightarrow \infty$ where $A = \bigcap_{n \in \mathbb{N}} A_n$.
3. (Shrinking mesh property). For all $A \in \mathcal{A}$,

$$\max_{\substack{A_0, A_1 \in [\emptyset, A] \cap \mathcal{A}_n: \\ A_0 \text{ maximal proper subset of } A_1}} d_{\mathcal{A}}(A_0, A_1) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

In the sequel, unless otherwise specified, $d_{\mathcal{A}}$ will stand for such a compatible metric. In particular, the metric $d_{\mathcal{T}}$ given by (1.4.1) endows \mathcal{T} with a topology, relating back to the usual definition of indexing collection given in [47] where \mathcal{T} is supposed to be a topological space from the start (see Section 1.6.1 for more details).

Open balls for $d_{\mathcal{A}}$ and $d_{\mathcal{T}}$ will be denoted by $B_{\mathcal{A}}(A, \rho)$ and $B_{\mathcal{T}}(t, \rho)$ respectively.

The first two properties of Definition 1.4.1 may already be found in [42, Definition 2.2] and ensure that $d_{\mathcal{A}}$ is compatible with the order structure on \mathcal{T} . As for the shrinking mesh property, it is just a metric version of the one given in Definition 1.3.1 and serves the same purpose.

EXAMPLE 1.4.2. The metric $d_m = \mathbf{m}(\Delta)$ is compatible due to the compatibility of \mathbf{m} itself. It plays an important role in the context of the regularity of stochastic processes. In that endeavor, the first milestone certainly is the Kolmogorov-Čentsov's regularity theorem and its variants [80]. They teach us that for a process $X = \{X_A : A \in \mathcal{A}\}$ in $L^2(\Omega)$, a natural metric to consider is $(A, A') \mapsto \|X_A - X_{A'}\|_{L^2(\Omega)} = \mathbf{E}((X_A - X_{A'})^2)^{1/2}$.

Moreover, when X is a Gaussian process [4, Section 1.3] or more generally, a Lévy process with stationarity measure \mathbf{m} (Definition 2.4.1), we have $\mathbf{E}[(X_A - X_{A'})^2] \propto d_m(A, A')$ for all $A, A' \in \mathcal{A}$. Hence the need to study d_m . Remark that elevating d_m to a power smaller than 1 changes nothing with regards to its set-indexed compatibility.

EXAMPLE 1.4.3 (Hausdorff metric). In the case where \mathcal{T} itself is endowed with a metric d , instead of directly defining $d_{\mathcal{A}}$ through $d_{\mathcal{T}} = d$, we may also consider the induced Hausdorff metric given by

$$\forall A, A' \in \mathcal{A}^*, \quad d_{\mathcal{H}}(A, A') = \inf \{ \varepsilon > 0 : A' \subseteq A^\varepsilon \text{ and } A \subseteq (A')^\varepsilon \} \quad (1.4.2)$$

where $A^\varepsilon = \{ t \in \mathcal{T} : d(t, A) \leq \varepsilon \}$ and $d_{\mathcal{H}}(\emptyset, A) = d_{\mathcal{H}}(\emptyset', A)$ for all $A \in \mathcal{A}$. The metric $d_{\mathcal{H}}$ is always contractive, outer continuous if d is with respect to the semilattice (\mathcal{T}, \leq) . Similarly, the shrinking mesh property may be formulated in terms of d .

The Hausdorff metric has been considered in the classical set-indexed setting from the beginning and many developments have it as a centerpiece (see [47, Chapters 1 and 7] for more details).

1.4.2 Metric $d_{\mathcal{C}}$ on \mathcal{C}

The metric $d_{\mathcal{A}}$ on \mathcal{A} may be naturally extended to a metric $d_{\mathcal{C}}$ on \mathcal{C} with corresponding interesting properties.

DEFINITION 1.4.4 (Metric $d_{\mathcal{C}}$ on \mathcal{C}). For any $C = A_0 \setminus \bigcup_{i=1}^k A_i$ and $C' = A'_0 \setminus \bigcup_{j=1}^{k'} A'_j$ in \mathcal{C} written with their extremal representations (Proposition 1.2.36), denote by $d_{\mathcal{C}}(C, C')$ the Hausdorff distance between the sets $\{A_0, \dots, A_k\}$ and $\{A'_0, \dots, A'_{k'}\}$, i.e.

$$d_{\mathcal{C}}(C, C') = \max \left\{ \max_{0 \leq i \leq k} \min_{0 \leq j \leq k'} d_{\mathcal{A}}(A_i, A'_j), \max_{0 \leq j \leq k'} \min_{0 \leq i \leq k} d_{\mathcal{A}}(A_i, A'_j) \right\}.$$

$(\mathcal{C}, d_{\mathcal{C}})$ is a metric space for which the canonical injection $(\mathcal{A}, d_{\mathcal{A}}) \hookrightarrow (\mathcal{C}, d_{\mathcal{C}})$ is an isometry, i.e. $d_{\mathcal{C}}(A, A') = d_{\mathcal{A}}(A, A')$ for all $A, A' \in \mathcal{A}$. Open balls for $d_{\mathcal{C}}$ will be denoted by $B_{\mathcal{C}}(C, \rho)$.

In order to comprehend what is going on for $d_{\mathcal{C}}$, the minimum is here in order to 'match' A_i with the closest A'_j and vice versa while the maximum takes the total error into account for the best matching. $d_{\mathcal{C}}$ is well-defined since the extremal representation of an element of \mathcal{C} is unique due to Proposition 1.2.36.

The next lemma basically tells that all constitutive elements of $C_n(t)$ (Definition 1.2.38) converge to $A(t)$ as n tends to infinity and sheds a new light on the metric $d_{\mathcal{T}}$ given by (1.4.1).

LEMMA 1.4.5. For all $t \in \mathcal{T}$, $d_{\mathcal{C}}(C_n(t), A(t)) \rightarrow 0$ as $n \rightarrow \infty$. Moreover,

$$\forall s, t \in \mathcal{T}, \quad d_{\mathcal{T}}(s, t) = \lim_{n \rightarrow \infty} d_{\mathcal{C}}(C_n(s), C_n(t)). \quad (1.4.3)$$

Proof. Let $t \in \mathcal{T}$ and $\varepsilon > 0$. By outer continuity of $d_{\mathcal{A}}$, there exists an integer n_0 such that

$$\forall n \geq n_0, \quad A_n(t) \in \mathcal{A}_n \quad \text{and} \quad d_{\mathcal{A}}(A_n(t), A(t)) \leq \varepsilon/2. \quad (1.4.4)$$

By the shrinking mesh property, there exists $n_1 \geq n_0$ such that

$$\forall n \geq n_1, \quad \max_{\substack{A, A' \in [\emptyset, A_{n_0}(t)] \cap \mathcal{A}_n: \\ A \text{ maximal proper subset of } A'}} d_{\mathcal{A}}(A, A') \leq \varepsilon/2. \quad (1.4.5)$$

By (1.4.4), for all $n \geq n_0$, $C_n(t) \in \mathcal{C}^\ell$. In particular, we may write its extremal representation $C_n(t) = A_n^0 \setminus \bigcup_{i=1}^{k_n} A_n^i$. By (1.4.5), we get

$$\forall n \geq n_1, \quad \forall i \leq k_n, \quad d_{\mathcal{A}}(A_n(t), A_n^i) \leq \varepsilon/2. \quad (1.4.6)$$

Hence it follows that for all $n \geq n_1$,

$$\begin{aligned} d_{\mathcal{C}}(C_n(t), A(t)) &\leq d_{\mathcal{C}}(C_n(t), A_n(t)) + d_{\mathcal{A}}(A_n(t), A(t)) \\ &\leq \varepsilon/2 + \varepsilon/2 \end{aligned} \quad \text{by (1.4.4) and (1.4.6).}$$

Hence $d_{\mathcal{C}}(C_n(t), A(t)) \rightarrow 0$ as $n \rightarrow \infty$.

Only (1.4.3) remains to prove. Let $s, t \in \mathcal{T}$. Remark that $C_n(s)$ and $C_n(t)$ both belong to \mathcal{C} for all n big enough, say $n \geq n_2$. Then we can write for all $n \geq n_2$,

$$d_{\mathcal{T}}(s, t) = d_{\mathcal{C}}(A(s), A(t)) \leq d_{\mathcal{C}}(A(s), C_n(s)) + d_{\mathcal{C}}(C_n(s), C_n(t)) + d_{\mathcal{C}}(C_n(t), A(t)).$$

Taking lower limits yields $d_{\mathcal{T}}(s, t) \leq \liminf_{n \rightarrow \infty} d_{\mathcal{C}}(C_n(s), C_n(t))$. Conversely, for all $n \geq n_2$,

$$d_{\mathcal{C}}(C_n(s), C_n(t)) \leq d_{\mathcal{C}}(C_n(s), A(s)) + d_{\mathcal{T}}(s, t) + d_{\mathcal{C}}(C_n(t), A(t)).$$

Taking upper limits yields $\limsup_{n \rightarrow \infty} d_{\mathcal{C}}(C_n(s), C_n(t)) \leq d_{\mathcal{T}}(s, t)$. The limit (1.4.3) follows. \square

1.4.3 Divergence δ and vicinities

We introduce here a notion of interaction between points $t \in \mathcal{T}$ and $A \in \mathcal{A}$ that the metrics $d_{\mathcal{A}}$ and $d_{\mathcal{T}}$ fail to capture. It will prove to be useful when looking at which points in \mathcal{T} have an influence over the regularity of some set-indexed process around a given $A \in \mathcal{A}$.

At the most basic level, looking at the regularity of a set-indexed process $X = \{X_{A'} : A' \in \mathcal{A}\}$ at some fixed $A \in \mathcal{A}$ means studying the behavior of increments of the form $X_A - X_{A'}$ for all $A' \in \mathcal{A}$ such that $d_{\mathcal{A}}(A, A')$ is small.

Using the additive extension ΔX of X provided by Proposition 1.2.44, we get for all A' ,

$$X_A - X_{A'} = (\Delta X_{A \setminus A'} + X_{A \cap A'}) - (\Delta X_{A' \setminus A} + X_{A \cap A'}) = \Delta X_{A \setminus A'} - \Delta X_{A' \setminus A}. \quad (1.4.7)$$

So the points in \mathcal{T} that are concerned by the increment $X_A - X_{A'}$ are those in $A \Delta A' = (A \setminus A') \sqcup (A' \setminus A)$. However for some point $t \in A \Delta A'$, having $d_{\mathcal{A}}(A, A')$ small does not necessarily mean that $d_{\mathcal{A}}(A, A(t))$ is small as well. This is illustrated Figure 1.10 when $\mathcal{T} = \mathbf{R}_+^2$ and $d_{\mathcal{T}}$ is the euclidean distance. Hence a need to characterize ‘how close’ such t is to A .

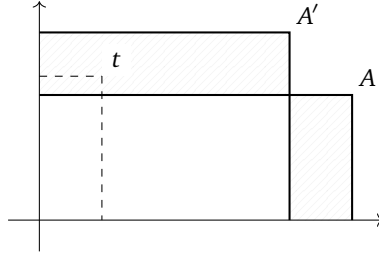


Figure 1.10: Point $t \in A \Delta A'$ where $d_{\mathcal{A}}(A, A')$ is much smaller than $d_{\mathcal{A}}(A, A(t))$.

DEFINITION 1.4.6 (Vicinity \mathcal{V} and divergence δ). For all $A \in \mathcal{A}$, $t \in \mathcal{T}$ and $\rho > 0$, define

$$\mathcal{V}(A, \rho) = \bigcup_{A' \in \mathcal{B}_{\mathcal{A}}(A, \rho)} (A \Delta A') \quad \text{and} \quad \delta(t, A) = \inf \{ \rho > 0 : t \in \mathcal{V}(A, \rho) \}$$

with the convention that $\mathcal{V}(A, \rho) = \emptyset$ for $\rho \leq 0$.

$\mathcal{V}(A, \rho)$ is called the vicinity of A of size ρ and $\delta(t, A)$ the divergence between t and A .

This definition naturally yields two notions of ‘open balls’ for $t \in \mathcal{T}$, $A \in \mathcal{A}$ and $\rho > 0$:

$$\mathcal{V}(A, \rho) = \{s \in \mathcal{T} : \mathfrak{s}(s, A) < \rho\}, \quad (1.4.8)$$

$$\mathcal{V}'(t, \rho) = \{A' \in \mathcal{A} : \mathfrak{s}(t, A') < \rho\}. \quad (1.4.9)$$

where we check that (1.4.8) is consistent with Definition 1.4.6. $\mathcal{V}'(t, \rho)$ will be called *dual vicinity of t of size ρ* .

EXAMPLE 1.4.7. Although $\mathfrak{s}(t, A) = d_{\mathcal{A}}(A(t), A)$ in the case where $\mathcal{T} = \mathbf{R}_+$ (or a tree more generally, see Example 1.2.25), other behaviors start to appear in higher-dimensional examples.

When $\mathcal{T} = \mathbf{R}_+^2$ and $d_{\mathcal{T}} = d_2$ is the usual euclidean distance, the vicinities are illustrated in Figure 1.11.

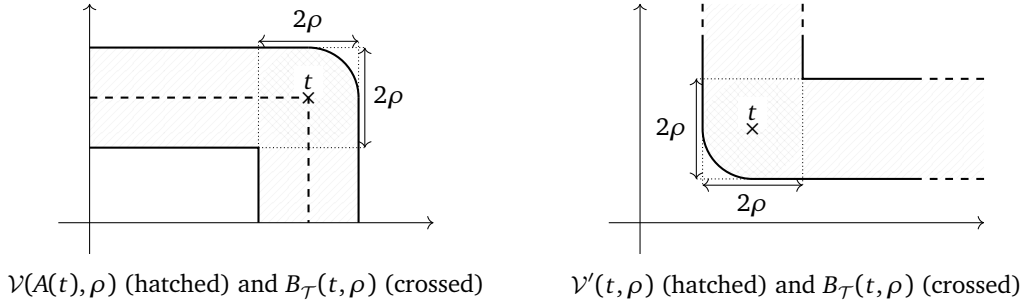


Figure 1.11: Vicinity and dual vicinity for $(\mathcal{T}, d_{\mathcal{T}}) = (\mathbf{R}_+^2, d_2)$.

Some might wonder whether through the TIP bijection, one could obtain a metric on \mathcal{T} with the formula $(s, t) \mapsto \mathfrak{s}(s, A(t))$. However that is not the case since both symmetry and triangle inequality fail in general. Indeed, if symmetry was true, we would have $\mathcal{V}(A(t), \rho) = \mathbf{t}(\mathcal{V}'(t, \rho))$ but Figure 1.11 strongly suggests that is generally not the case. As for triangle inequality, Figure 1.12 illustrates a case where while both $\mathfrak{s}(s, A(t))$ and $\mathfrak{s}(t, A(u))$ are small — s and t are contained in corresponding small vicinities — $\mathfrak{s}(s, A(u))$ is big, so $\mathfrak{s}(s, A(u)) \leq \mathfrak{s}(s, A(t)) + \mathfrak{s}(t, A(u))$ cannot hold.

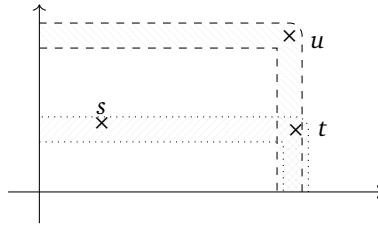


Figure 1.12: The hatched regions represent vicinities of $A(t)$ and $A(u)$.

Before moving on, let us briefly study \mathcal{V} and \mathfrak{s} in order to see that, even though there is no metric structure in general, those objects still retains some related nice geometric properties.

PROPOSITION 1.4.8 (Properties of \mathcal{V}). *The following properties hold:*

1. (Vicinity behaves like open balls). For all $A, A' \in \mathcal{A}$ and $\rho > 0$,

$$\bigcup_{\rho' < \rho} \mathcal{V}(A, \rho') = \mathcal{V}(A, \rho) \subseteq \bigcap_{\rho' > \rho} \mathcal{V}(A, \rho') \quad (1.4.10)$$

and

$$\mathcal{V}(A', \rho - d_{\mathcal{A}}(A, A')) \subseteq \mathcal{V}(A, \rho) \subseteq \mathcal{V}(A', \rho + d_{\mathcal{A}}(A, A')). \quad (1.4.11)$$

2. (Discretization of the vicinity). For all $A \in \mathcal{A}$, $\rho > 0$ and $n \in \mathbf{N}$, denote

$$\mathcal{V}_n(A, \rho) = \bigcup_{\substack{\bar{A}, \underline{A} \in \mathcal{A}_n \cap B_{\mathcal{A}}(A, \rho) \\ \bar{A} \subseteq A}} (\bar{A} \setminus \underline{A}) = \bar{\mathcal{V}}_n(A, \rho) \setminus \underline{\mathcal{V}}_n(A, \rho) \quad (1.4.12)$$

where $\bar{\mathcal{V}}_n(A, \rho)$ (resp. $\underline{\mathcal{V}}_n(A, \rho)$) is the union (resp. intersection) of all maximal (resp. minimal) elements for \subseteq in $\mathcal{A}_n \cap B_{\mathcal{A}}(A, \rho)$. Then

$$\mathcal{V}(A, \rho) = \bigcup_{n \in \mathbf{N}} \mathcal{V}_n(A, \rho). \quad (1.4.13)$$

The vicinity $\mathcal{V}(A, \rho)$ will actually play an important role to understand the Hölder regularity of some set-indexed processes (see (3.5.6) and Theorem 3.5.2). As an intermediary step, the $\mathcal{V}_n(A, \rho)$'s should be seen as a discretized version of $\mathcal{V}(A, \rho)$ and will be used to get an upper bound on the Hölder regularity based on the jumps of the process (Lemma 3.4.8). The vicinity also intervenes in a set-indexed versions of Doob's maximal inequality (Corollary 2.3.31) as well as a 0-1 law similar to Blumenthal's (Theorem 2.3.33).

Proof. Let us fix $A, A' \in \mathcal{A}$, $n \in \mathbf{N}$ and $\rho > 0$.

1. The relation (1.4.10) is a straightforward consequence of the definition.

Let us prove the second inclusion of (1.4.11). Consider $A'' \in B_{\mathcal{A}}(A, \rho)$. Then

$$\begin{aligned} A \Delta A'' &= (A \setminus A'') \cup (A'' \setminus A) \\ &\subseteq [(A \setminus A') \cup (A' \setminus A'')] \cup [(A'' \setminus A') \cup (A' \setminus A)] \\ &= (A \Delta A') \cup (A' \Delta A''). \end{aligned}$$

Since both $d_{\mathcal{A}}(A', A)$ and $d_{\mathcal{A}}(A', A'')$ are smaller than $\rho + d_{\mathcal{A}}(A, A')$, we get

$$A \Delta A'' \subseteq \mathcal{V}(A', \rho + d_{\mathcal{A}}(A, A')).$$

Hence $\mathcal{V}(A, \rho) \subseteq \mathcal{V}(A', \rho + d_{\mathcal{A}}(A, A'))$. The first inclusion follows from this one by permuting A and A' as well as replacing ρ by $\rho - d_{\mathcal{A}}(A, A')$.

2. Let us prove that the definition (1.4.12) of $\mathcal{V}_n(A, \rho)$ is consistent.

Denote

$$\begin{aligned} \bar{\mathcal{V}}_n(A, \rho) &= \bar{A}_1 \cup \dots \cup \bar{A}_k \\ \underline{\mathcal{V}}_n(A, \rho) &= \underline{A}_1 \cap \dots \cap \underline{A}_\ell \end{aligned}$$

where the \bar{A}_i 's (resp. \underline{A}_j 's) are the maximal (resp. minimal) elements in $\mathcal{A}_n \cap B_{\mathcal{A}}(A, \rho)$. Then

$$\bar{\mathcal{V}}_n(A, \rho) \setminus \underline{\mathcal{V}}_n(A, \rho) = \bigcup_{\substack{1 \leq i \leq k \\ 1 \leq j \leq \ell}} (\bar{A}_i \setminus \underline{A}_j).$$

From this expression, the converse inclusion in (1.4.12) is straightforward whereas the direct inclusion comes from the fact that any $\bar{A} \setminus \underline{A}$ is included in some $\bar{A}_i \setminus \underline{A}_j$.

Let us prove the direct inclusion in (1.4.13). For $A' \in B_{\mathcal{A}}(A, \rho)$, we have $A\Delta A' = [A \setminus (A \cap A')] \cup [A' \setminus (A \cap A')]$ by definition. Hence, by separability from above (Definition 1.2.1), we get

$$\forall n_0 \in \mathbf{N}, \quad A\Delta A' \subseteq \bigcup_{n \geq n_0} [g_n(A) \setminus g_n(A \cap A')] \cup [g_n(A') \setminus g_n(A \cap A')] \quad (1.4.14)$$

which would then be included in $\bigcup_{n \in \mathbf{N}} \mathcal{V}_n(A, \rho)$ as long as there exists $n_0 \in \mathbf{N}$ such that

$$\forall n \geq n_0, \quad \max \{d_{\mathcal{A}}(A, g_n(A)), d_{\mathcal{A}}(A, g_n(A')), d_{\mathcal{A}}(A, g_n(A \cap A'))\} < \rho. \quad (1.4.15)$$

So let us find such n_0 . Using outer continuity and contractivity for the last inequality (Definition 1.4.1), we get the following:

$$\begin{array}{llll} d_{\mathcal{A}}(A, g_n(A)) & \xrightarrow{n \rightarrow \infty} & d_{\mathcal{A}}(A, A) & = 0 < \rho \\ d_{\mathcal{A}}(A, g_n(A')) & \xrightarrow{n \rightarrow \infty} & d_{\mathcal{A}}(A, A') & < \rho \\ d_{\mathcal{A}}(A, g_n(A \cap A')) & \xrightarrow{n \rightarrow \infty} & d_{\mathcal{A}}(A, A \cap A') & \leq d_{\mathcal{A}}(A, A') < \rho. \end{array}$$

Thus (1.4.15) is true for a big enough $n_0 \in \mathbf{N}$. Hence

$$\mathcal{V}(A, \rho) \subseteq \bigcup_{n \in \mathbf{N}} \mathcal{V}_n(A, \rho).$$

For the converse inclusion in (1.4.13), fix $\bar{A}, \underline{A} \in B_{\mathcal{A}}(A, \rho)$ such that $\underline{A} \subset \bar{A}$. Then,

$$\begin{aligned} \bar{A} \setminus \underline{A} &= [(A \cap \bar{A}) \setminus \underline{A}] \cup [\bar{A} \setminus (A \cap \underline{A})] \\ &\subseteq [(A \cap \bar{A}) \Delta \underline{A}] \cup [\bar{A} \Delta (A \cap \underline{A})]. \end{aligned}$$

Contractivity then shows that both $A \cap \bar{A}$ and $A \cap \underline{A}$ belong to $B_{\mathcal{A}}(A, \rho)$. The result follows. \square

PROPOSITION 1.4.9 (Ersatz of triangle inequality). *For all $t \in \mathcal{T}$, the map $\mathfrak{s}(t, \cdot)$ is 1-Lipschitz, i.e.*

$$\forall A, A' \in \mathcal{A}, \quad |\mathfrak{s}(t, A) - \mathfrak{s}(t, A')| \leq d_{\mathcal{A}}(A, A').$$

Proof. Let us fix $t \in \mathcal{T}$, $A, A' \in \mathcal{A}$ and $\rho > d_{\mathcal{A}}(A, A')$. Denoting $\varepsilon = \rho - d_{\mathcal{A}}(A, A')$ and using (1.4.11), we obtain

$$\begin{aligned} t &\in \mathcal{V}(A, \mathfrak{s}(t, A) + \varepsilon) \subseteq \mathcal{V}(A', \mathfrak{s}(t, A) + \rho), \\ \mathcal{V}(A', \mathfrak{s}(t, A) - \rho) &\subseteq \mathcal{V}(A, \mathfrak{s}(t, A) - \varepsilon) \not\ni t. \end{aligned}$$

Hence, by definition of $\mathfrak{s}(t, A')$, for all $\rho > d_{\mathcal{A}}(A, A')$,

$$\mathfrak{s}(t, A) - \rho \leq \mathfrak{s}(t, A') \leq \mathfrak{s}(t, A) + \rho.$$

The result follows from taking $\rho \rightarrow d_{\mathcal{A}}(A, A')^+$ in the previous inequality. \square

1.4.4 Topological consequence

The following property will be useful when studying weak convergence in some functional spaces in Chapter 3.

PROPOSITION 1.4.10 (σ -compactity). *For all $U \in \mathcal{A}(u)$, the set $[\emptyset, U] = \{A \in \mathcal{A} : A \subseteq U\}$ is compact. In particular, $(\mathcal{A}, d_{\mathcal{A}})$ is σ -compact, i.e. it is a countable union of compact sets.*

Proof. Let $U \in \mathcal{A}(u)$. Without loss of generality, we might as well suppose that the constitutive elements of U all belong to \mathcal{A}_0 so that for all $n \in \mathbf{N}$ and $A \in [\emptyset, U]$, $g_n(A) \in \mathcal{A}$.

Let us consider a sequence $(A_k)_{k \in \mathbf{N}}$ in $[\emptyset, U]$ and prove that it has a convergent subsequence in $[\emptyset, U]$. By the pigeonhole principle, there exists $A'_0 \in \mathcal{A}_0$ and an extraction φ_0 such that $g_0(A_{\varphi_0(k)}) = A'_0$ for all $k \in \mathbf{N}$. Iterating in this manner yields a non-increasing sequence $(A'_n)_{n \in \mathbf{N}}$ in $[\emptyset, U]$ and a sequence of extractions $(\varphi_n)_{n \in \mathbf{N}}$ such that

$$\forall n, k \in \mathbf{N}, \quad g_n(A_{\varphi_0 \circ \dots \circ \varphi_n(k)}) = A'_n.$$

Define then the diagonal extraction ψ by $\psi(n) = \varphi_0 \circ \dots \circ \varphi_n(n)$ for all $n \in \mathbf{N}$ and $A' = \bigcap_{n \in \mathbf{N}} A'_n$. By outer continuity, we have

$$d_{\mathcal{A}}(g_n(A_{\psi(n)}), A') = d_{\mathcal{A}}(A'_n, A') \longrightarrow 0 \quad \text{as} \quad n \rightarrow \infty. \quad (1.4.16)$$

By the shrinking mesh property, $d_{\mathcal{A}}(g_n(A), A) \rightarrow 0$ as $n \rightarrow \infty$ uniformly in $A \in [\emptyset, U]$. In particular,

$$d_{\mathcal{A}}(g_n(A_{\psi(n)}), A_{\psi(n)}) \longrightarrow 0 \quad \text{as} \quad n \rightarrow \infty. \quad (1.4.17)$$

Combining (1.4.16) and (1.4.17) yields $d_{\mathcal{A}}(A_{\psi(n)}, A') \rightarrow 0$ as $n \rightarrow \infty$. Hence $(A_k)_{k \in \mathbf{N}}$ has a convergent subsequence in $[\emptyset, U]$, so $[\emptyset, U]$ is compact.

As for σ -compactness, it is a consequence of the previous result and writing $\mathcal{A} = \bigcup_{n \in \mathbf{N}} [\emptyset, \mathcal{T}_n]$ where $\mathcal{T}_n = \bigcup_{A \in \mathcal{A}_n} A$. \square

EXAMPLE 1.4.11 (\mathcal{A} is generally not locally compact). *Although \mathcal{A} is σ -compact, it does not imply local compactity. To see that, consider the continuous tree $\mathcal{T} \simeq \{0_{\mathcal{T}}\} \sqcup ((0, 1] \times \mathbf{N})$ from Example 1.2.25 with the discrete tree $\mathcal{U} = \{0_{\mathcal{U}}\} \cup \mathbf{N}$ and illustrated in Figure 1.13.*

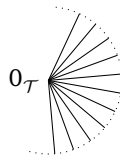


Figure 1.13: Continuous tree obtained from $\mathcal{U} = \{0_{\mathcal{U}}\} \cup \mathbf{N}$.

The space \mathcal{T} is endowed with the ‘SNCF metric’ $d_{\mathcal{T}}$ given for all $s = (x, u)$ and $t = (y, v)$ in $(0, 1] \times \mathbf{N}$ by

$$d_{\mathcal{T}}(s, 0_{\mathcal{T}}) = x, \quad d_{\mathcal{T}}(0_{\mathcal{T}}, t) = y \quad \text{and} \quad d_{\mathcal{T}}(s, t) = \begin{cases} |x - y| & \text{if } u = v, \\ x + y & \text{if } u \neq v. \end{cases}$$

where we made the non-consequential abuse of considering that the \simeq above is in fact an equality.

Let $\rho > 0$. The open ball $B_{\mathcal{T}}(0_{\mathcal{T}}, \rho)$ cannot be compact since it contains the sequence $((\rho/2, u))_{u \in \mathbf{N}}$ which has no convergent subsequence. By the relation (1.4.1) between $d_{\mathcal{T}}$ and $d_{\mathcal{A}}$, \mathcal{A} is not locally compact around \emptyset .

1.5 Indexing collections of finite dimension

1.5.1 A poset approach

In this part, we propose a set of hypotheses which embodies the notion for an indexing collection \mathcal{A} to have finite dimension. They will be used in particular to establish a strong Lévy-Itô decomposition (Theorem 3.3.8) and combined with the concepts of the previous section to prove some martingale related results (Section 2.3.3).

DEFINITION 1.5.1 (Order embedding). *If (E, \preceq_E) and (F, \preceq_F) are two partially ordered sets, an order embedding $\varphi : E \hookrightarrow F$ is a map such that*

$$\forall x, y \in E, \quad x \preceq_E y \iff \varphi(x) \preceq_F \varphi(y).$$

DEFINITION 1.5.2 (Order dimension). *A partially ordered set (E, \preceq) has poset (or order) dimension $\leq p$ where $p \in \mathbf{N}$ if there exists an order embedding $\varphi : E \hookrightarrow \mathbf{N}^p$ where \mathbf{N}^p is endowed with the usual componentwise partial order and $\mathbf{N}^0 = \{0\}$ by convention.*

Some authors prefer an alternative definition to Definition 1.5.2 based on the intersection of linear orders. We refer to [79, Theorem 10.4.2] to see they are equivalent.

DEFINITION 1.5.3 (Indexing collection of finite dimension). *The indexing collection \mathcal{A} is said to have finite dimension if there exists $p \in \mathbf{N}$ such that the following properties hold:*

1. *The \mathcal{A}_n 's all have poset dimension $\leq p$.*
2. *$\mathcal{C}^\ell \subseteq \mathcal{C}_{(p)}$, i.e. any left-neighborhood $C \in \mathcal{C}^\ell$ may be written $C = A_0 \setminus \bigcup_{i=1}^p A_i$ where $A_0, \dots, A_p \in \mathcal{A}$.*

The smallest of such integers will be called the dimension of \mathcal{A} and denoted $\dim \mathcal{A}$.

In the sequel, the dimension of \mathcal{A} shall not always be finite. The hypothesis will appear explicitly each time we suppose so.

As usual, we comment on the definition before moving on. The two properties bound the local of elements in the ‘mesh’ \mathcal{A}_n as n goes to ∞ but each in a different way. They follow a philosophy similar to [42, Assumption $\mathcal{H}_{\mathcal{A}}$] albeit without being equivalent.

EXAMPLE 1.5.4. *The usual indexing collection of $\mathcal{T} = \mathbf{R}_+^p, \mathbf{R}^p$ or $[0, 1]^p$ (see Examples 1.2.10 and 1.2.28) has dimension p .*

EXAMPLE 1.5.5. *The indexing collection for a continuous tree from Example 1.2.25 or 1.2.30 has dimension 2 as long as it not a chain (in this cas, it has dimension 1).*

Remark that being of finite dimension actually concerns the sequence $(\mathcal{A}_n)_{n \in \mathbf{N}}$ more than \mathcal{A} itself. If that were not the case, most of those continuous trees would have a dimension much bigger than 2.

EXAMPLE 1.5.6 (1 $\not\leq$ 2 in Definition 1.5.3). *One might think that if the \mathcal{A}_n 's have poset dimension $\leq p$, then $\mathcal{C}^\ell \subseteq \mathcal{C}_{(q)}$ for some $q \in \mathbf{N}$ possibly greater than p . Unfortunately, that is generally not the case even for $\mathcal{T} = [0, 1]^2$ endowed with the usual component-wise partial order (restrict Example 1.2.10 to $[0, 1]^2$) if $(\mathcal{A}_n)_{n \in \mathbf{N}}$ happens to be badly chosen. Figure 1.14 illustrates such a case where we drew the first steps to a construction where for all $n \in \mathbf{N}$, the poset dimension of \mathcal{A}_n is ≤ 2 while $C_n(1, 1) \in \mathcal{C}_{(n+1)} \setminus \mathcal{C}_{(n)}$.*

EXAMPLE 1.5.7 ($2 \not\approx 1$ in Definition 1.5.3). Just as well, having $\mathcal{C}^\ell \subseteq \mathcal{C}_{(p)}$ does not impose a uniform bound on the poset dimension of the \mathcal{A}_n 's. To present a counter-example, we will make full use of Theorem 1.2.5 and specify the indexing semilattice (\mathcal{T}, \preceq) together with an increasing sequence $(\mathcal{D}_n)_{n \in \mathbf{N}}$ of finite subsets of \mathcal{T} closed under minima such that $\mathcal{D} = \bigcup_{n \in \mathbf{N}} \mathcal{D}_n$ satisfies the separability from above condition (Definition 1.2.3). For all $n \in \mathbf{N}$, the class \mathcal{A}_n is then recovered from \mathcal{D}_n through the TIP bijection \mathbf{t} : $\mathcal{A}_n = \{\mathbf{t}^{-1}(D) : D \in \mathcal{D}_n\} \cup \{\emptyset\}$.

Denote $\mathcal{T} = \{E \subseteq \mathbf{N} : \#E \leq 2\}$ ordered by inclusion and for all $n \in \mathbf{N}$, $\mathcal{D}_n = \{E \subseteq \llbracket 0, n \rrbracket : \#E \leq 2\}$. The poset (\mathcal{T}, \subseteq) is closed under any intersections and is countable, so $\mathcal{D} = \bigcup_{n \in \mathbf{N}} \mathcal{D}_n = \mathcal{T}$ directly satisfies separability from above. Hence (\mathcal{T}, \subseteq) is an indexing semilattice.

In this context, condition 2 of Definition 1.5.3 translates into: for any $E \in \mathcal{T}$, E has at most p maximal proper subsets in \mathcal{T} . It is then easily seen that $p = 3$ works fine here. As for condition 1, [94] tells us that the poset dimension of \mathcal{D}_n is of order $\log_2 \log_2 n$ (an asymptotic expansion is worked out in [32], see also [98] for an introductory survey on the matter). In particular, the poset dimension of \mathcal{D}_n cannot be uniformly bounded with respect to n .

Remark that even though this example is discrete — and thus excluded once a compatible measure \mathbf{m} is considered — it can be made into a continuous one by a series gluing of unit segments similar to Example 1.2.25.

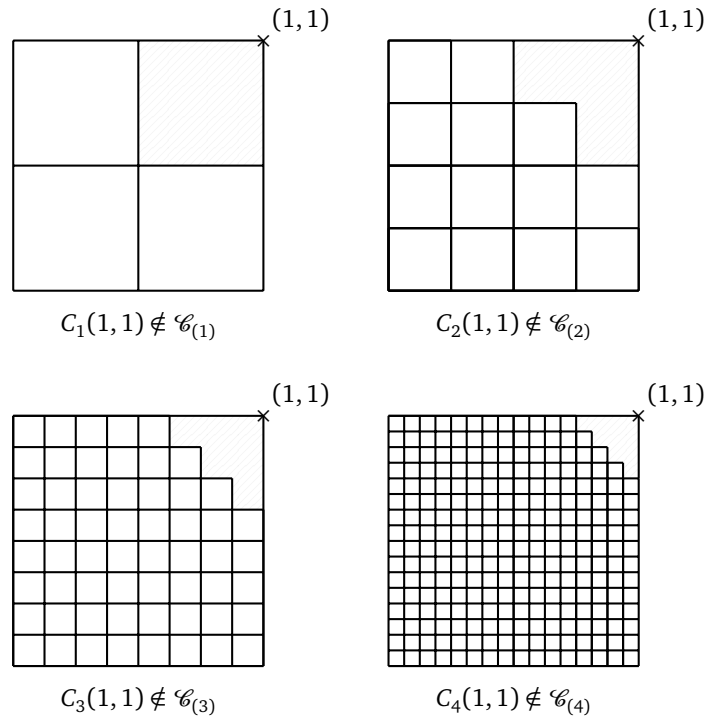


Figure 1.14: $\mathcal{T} = [0, 1]^2$ and a bad choice of $(\mathcal{A}_n)_{n \in \mathbf{N}}$.

1.5.2 Entropy bound

An important consequence of being finite-dimensional is that it gives a bound on the entropy of \mathcal{A} . This property has been successfully applied by numerous authors to investigate the regularity of stochastic processes as well as Donsker classes (see [4, 100] and references therein for more details).

DEFINITION 1.5.8 (VC-class, [100, Section 2.6.1]). *Let \mathcal{D} be a collection of subsets of \mathcal{T} . For any finite subset $E \subseteq \mathcal{T}$, denote*

$$\Delta_{\mathcal{D}}(E) = \#\{E \cap D : D \in \mathcal{D}\}.$$

The class \mathcal{D} shatters E if $\Delta_{\mathcal{D}}(E) = \#\mathcal{D}(E) = 2^{\#E}$. The VC-index $VC(\mathcal{D})$ of \mathcal{D} is defined as

$$VC(\mathcal{D}) = \inf \left\{ k \in \mathbf{N} : \max_{\substack{E \subseteq \mathcal{T}: \\ \#E=k}} \Delta_{\mathcal{D}}(E) < 2^k \right\}$$

with the convention $\inf \emptyset = \infty$. If $VC(\mathcal{D}) < \infty$, then \mathcal{D} is called a Vapnik-Červonenkis class.

EXAMPLE 1.5.9 (VC-index of the indexing collection on \mathbf{R}_+^p). *Suppose that \mathcal{A} is the usual indexing collection of rectangles on $\mathcal{T} = \mathbf{R}_+^p$ where $p \in \mathbf{N}^*$ (Example 1.2.10). Let us show that $VC(\mathcal{A}) = p + 1$.*

First, the canonical basis $\mathcal{B} = \{e_1, \dots, e_p\}$ of \mathbf{R}^p is shattered by \mathcal{A} . Indeed, for any subset $\mathcal{B}' \subseteq \mathcal{B}$, we have $\mathcal{B} \cap A(t_{\mathcal{B}'}) = \mathcal{B}'$ where $t_{\mathcal{B}'} = \sum_{e \in \mathcal{B}'} e$. Hence $\Delta_{\mathcal{A}}(\mathcal{B}) = 2^p$, so $VC(\mathcal{A}) > p$.

Let us show the converse inequality. Let $E = \{t_1, \dots, t_{p+1}\}$ be a subset of \mathbf{R}_+^p of cardinal $p + 1$. Instead of relabelling the vectors, we might as well suppose that for all $j \in \llbracket 1, p \rrbracket$, $t_{p+1,j} \leq \max_{1 \leq i \leq p} t_{i,j}$. Denote $s = (\max_{1 \leq i \leq p} t_{i1}, \dots, \max_{1 \leq i \leq p} t_{ip}) \in \mathbf{R}_+^p$. Then the previous assumption reads $t_{p+1} \preceq s$ where \preceq is the component-wise partial order on \mathbf{R}_+^p . Since the smallest set in \mathcal{A} containing $\{t_1, \dots, t_p\}$ is $A(s)$ and $t_{p+1} \in A(s)$ as well, E cannot be shattered by \mathcal{A} . Hence $VC(\mathcal{A}) \leq p + 1$. The result follows.

LEMMA 1.5.10 (Bounding the VC-index by the dimension). *The following inequality holds: $VC(\mathcal{A}) \leq \dim \mathcal{A} + 1$. In particular, if \mathcal{A} is finite-dimensional, it is also a Vapnik-Červonenkis class.*

Proof. Denote $p = \dim \mathcal{A}$. If $p = \infty$, there is nothing to prove.

Let us suppose that $p < \infty$ and show $VC(\mathcal{A}) \leq p + 1$. Let E be a finite subset of \mathcal{T} shattered by \mathcal{A} . By separability from above (Definition 1.2.1), E must also be shattered by \mathcal{A}_n for some big enough $n \in \mathbf{N}$. Consider an order embedding $\varphi : \mathcal{A}_n \hookrightarrow \mathbf{N}^p$ and denote $F = \{\varphi(A_n(t)) : t \in E\} \subseteq \mathbf{N}^p$. Since E is shattered by \mathcal{A}_n and φ is injective, we have $\#E = \#F$. Moreover, since φ is also order embedding, F is shattered by the indexing collection on \mathbf{R}_+^p . In particular, $\#E = \#F \leq p$ according to Example 1.5.9. The result follows. \square

DEFINITION 1.5.11 (Log-entropy, [4, Definition 1.3.2]). *Let $(\mathcal{D}, d_{\mathcal{D}})$ be a metric space. For all $\varepsilon > 0$, denote by $N(\varepsilon, \mathcal{D}, d_{\mathcal{D}}) = N(\varepsilon)$ the minimal number — possibly infinite — of balls of radius ε required to cover \mathcal{D} . The log-entropy of \mathcal{D} is then defined for $\varepsilon > 0$ as $\mathcal{H}(\varepsilon, \mathcal{D}, d_{\mathcal{D}}) = \mathcal{H}(\varepsilon) = \log N(\varepsilon)$.*

One could not overstate the importance of log-entropy when studying the regularity of stochastic processes. For instance, Dudley [28, Theorem 1.1] proved that for a centered Gaussian process $X = \{X_t : t \in \mathcal{T}\}$ to have continuous sample paths, it is sufficient that $\int_0^1 \mathcal{H}^{1/2}(\varepsilon) d\varepsilon < \infty$ where \mathcal{H} is the log-entropy on $(\mathcal{T}, d_{\mathcal{T}})$ for $d_{\mathcal{T}}(s, t) = \|X_s - X_t\|_{L^2(\Omega)}$. Moreover, this condition

becomes necessary if \mathcal{T} is a group and the law X is invariant by translation. We refer to [4, Theorems 1.3.5 and 1.5.4] for more details.

In order to prepare for the study of sample paths of set-indexed processes, we give an entropy bound which is mainly a consequence of [100, Theorem 2.6.7].

THEOREM 1.5.12 (Entropy bound, adapted from [100, Theorem 2.6.7]). *If $p = \dim \mathcal{A}$ is finite, then for all $U \in \mathcal{A}(u)$, there exists $\kappa_{p,m(U)} > 0$ such that*

$$\forall \varepsilon \in (0, \mathbf{m}(U) \wedge 1), \quad N(\varepsilon, [\emptyset, U], d_m) \leq \kappa_{p,m(U)} \left(\frac{1}{\varepsilon} \right)^p.$$

In particular,

$$\mathcal{H}(\varepsilon, [\emptyset, U], d_m) = O(\log(1/\varepsilon)) \quad \text{as} \quad \varepsilon \rightarrow 0^+.$$

Proof. Suppose that $p = \dim \mathcal{A}$ is finite and consider $U \in \mathcal{A}(u)$. If $\mathbf{m}(U) = 0$, then $N(\varepsilon, [\emptyset, U], d_m) = 1$ for all $\varepsilon > 0$ so the two estimates are trivial. Now, let us suppose that $\mathbf{m}(U) > 0$. Consider the class $\mathcal{F} = \{\mathbb{1}_A : A \in [\emptyset, U]\}$. By (1.3.10),

$$\forall \varepsilon > 0, \quad N(\varepsilon, [\emptyset, U], d_m) = N(\varepsilon, \mathcal{F}, \|\cdot\|_{L^1(m)}). \quad (1.5.1)$$

According to [100, Theorem 2.6.7] and since $\mathbf{m}(U \cap \cdot)$ is a finite measure, there exists a constant $\kappa_{VC(\mathcal{F}),m(U)} > 0$ (increasing in both its parameters) such that

$$\forall \varepsilon \in (0, \mathbf{m}(U)), \quad N(\varepsilon, \mathcal{F}, \|\cdot\|_{L^1(m)}) \leq \kappa'_{VC(\mathcal{F}),m(U)} \left(\frac{1}{\varepsilon} \right)^{VC(\mathcal{F})-1} \quad (1.5.2)$$

where $VC(\mathcal{F})$ is the VC-index of the class \mathcal{F}' of subgraphs of \mathcal{F} , i.e. $\mathcal{F}' = \{C_A : A \in [\emptyset, U]\}$ where

$$\forall A \in \mathcal{A}, \quad C_A = (A \times (-\infty, 1)) \cup (A^b \times (-\infty, 0)) \quad (\subseteq \mathcal{T} \times \mathbf{R}).$$

It is straightforward to see that for all $A, A' \in \mathcal{A}$, $A \subseteq A'$ if and only if $C_A \subseteq C_{A'}$. Thus $\mathcal{A} \ni A \mapsto C_A \in \mathcal{F}'$ is an order embedding. In particular, we may copy the proof of Lemma 1.5.10 to get $VC(\mathcal{F}) \leq p + 1$. Combining (1.5.1) and (1.5.2) then gives

$$\forall \varepsilon \in (0, \mathbf{m}(U) \wedge 1), \quad N(\varepsilon, [\emptyset, U], d_m) \leq \kappa'_{VC(\mathcal{F}),m(U)} \left(\frac{1}{\varepsilon} \right)^p$$

and since $\kappa'_{VC(\mathcal{F}),m(U)}$ is increasing in its parameters, we may replace it by some other constant $\kappa_{p,m(U)}$. The result follows. \square

1.6 ‘Revamped’ vs ‘classical’ theory

The setting developed in this chapter differs on several account from the classical one presented in [46, 47, 71]. To our opinion, the main difference between the two resides in the way it is presented:

- ◊ The classical setting puts all the main hypotheses in the definition of the indexing collection (see Definition 1.6.2 below). Once the definition is well-understood, it has the advantage of studying a unique object throughout the theory. However, one drawback is that it becomes hard to think about new examples since they have to check every condition in a long definition. This has led a lot of people to mistakenly think that the only example of interest is the multiparameter case $\mathcal{T} = \mathbf{R}_+^p$ (Example 1.2.10).

◊ Our ‘revamped’ setting chooses a more ‘modular’ approach in the sense that the definition of an indexing collection is axiomatically light, especially its formulation through indexing semilattices (Theorem 1.2.5), and more structure is added as we go on: a measure \mathbf{m} (Definition 1.3.1), a metric $d_{\mathcal{A}}$ (Definition 1.4.1) and sometimes some finite-dimensional assumption (Definition 1.5.3). We hope that this approach will make set-indexed theory more widespread and accepted among probabilists. Of course, there are still some differences in the choice of the core assumptions, but that will be explained as we go.

In this section, we will work our way through several concepts of the classical theory and explain how they tie (or not) with our setting. The goal is not to advocate for one theory or the other since they have been developed for different purposes. Historically, classical theory has been developed to study generalisations of multiparameter random fields, especially from the martingale point of view. Through decades of research, the hypotheses were then tailor-made to tackle concepts such as stopping and the several types of ‘pasts’ already arising in the theory of two-parameter martingales. We refer to the introduction of [47] for a more detailed account and its bibliography for an exhaustive list of references. On the other side, when devising definitions, we had in mind the regularity of sample paths, which lead to different choices that will be progressively motivated.

REMARK 1.6.1. *In this section and this section only, the notations \mathcal{A} , \mathcal{A}_n , $\mathcal{A}(u)$, $\mathcal{A}_n(u)$, \mathcal{T}_n , g_n , \mathbf{m} and $d_{\mathcal{A}}$ may differ from what we usually mean and may stand for their classical counterparts.*

1.6.1 A taste of the classical setting

In the classical setting, the space \mathcal{T} is most of the time supposed to be endowed with a (true) metric $d_{\mathcal{T}}$ making it a complete, separable, σ -compact metric space. In our setting, $(\mathcal{T}, d_{\mathcal{T}})$ is a (pseudo) metric space which is separable (due to (1.4.1) and separability from above) and σ -compact (Proposition 1.4.10). So far, our setting is a tad more general.

In the following, $\overline{(\cdot)}$ (resp. $(\cdot)^\circ$) denotes the closure (resp. the interior) with respect to the topology on $(\mathcal{T}, d_{\mathcal{T}})$.

DEFINITION 1.6.2 (Classical indexing collection). *A class \mathcal{A} of compact, connected subsets of \mathcal{T} is a classical indexing collection on \mathcal{T} if the following properties hold:*

1. $\emptyset \in \mathcal{A}$ and there exists an increasing sequence $(\mathcal{T}_n)_{n \in \mathbf{N}}$ in $\mathcal{A}(u)$ such that $\mathcal{T} = \bigcup_{n \in \mathbf{N}} \mathcal{T}_n^\circ$.
2. \mathcal{A} is closed under arbitrary intersections and $\mathcal{A}^* = \mathcal{A} \setminus \{\emptyset\}$ is closed under finite intersections. If $(A_k)_{k \in \mathbf{N}}$ is an increasing sequence in \mathcal{A} and there exists $n \in \mathbf{N}$ such that $A_k \subseteq \mathcal{T}_n$ for all $k \in \mathbf{N}$, then $\bigcup_{k \in \mathbf{N}} A_k \in \mathcal{A}$.
3. $\sigma(\mathcal{A}) = \mathcal{B}$ is the Borel σ -algebra of $(\mathcal{T}, d_{\mathcal{T}})$.
4. (Classical separability from above). There exists a non-decreasing sequence of finite subcollections $\mathcal{A}_n = \{A_1^n, \dots, A_{k_n}^n\} \subseteq \mathcal{A}$ ($n \in \mathbf{N}$) closed under intersections with \emptyset , $\mathcal{T}_n \in \mathcal{A}_n(u)$ and a sequence of functions $g_n : \mathcal{A} \rightarrow \mathcal{A}_n(u) \cup \{\mathcal{T}\}$ such that: (a) g_n preserves arbitrary intersections and finite unions, (b) for all $A \in \mathcal{A}$, $A \subseteq g_n(A)^\circ$, $A = \bigcap_{n \in \mathbf{N}} g_n(A)$ and $g_n(A) \subseteq g_m(A)$ if $n \geq m$, (c) $g_n(A) \cap A' \in \mathcal{A}$ if $A, A' \in \mathcal{A}$ and $g_n(A) \cap A' \in \mathcal{A}_n$ if $A' \in \mathcal{A}_n$, (d) $g_n(\emptyset) = \emptyset$.

A first thing that comes to mind is that a classical indexing collection is *almost* an indexing collection but for TIP and SHAPE, which are classical hypotheses. Our choice to have them as core hypotheses will be defended in the subsequent Sections 1.6.2 and 1.6.3.

A first difference is that a classical indexing collection is irremediably linked to the topology on \mathcal{T} .

For the first property of Definition 1.6.2, having $\mathcal{T} = \bigcup_{n \in \mathbb{N}} \mathcal{T}_n^\circ$ instead of $\mathcal{T} = \bigcup_{n \in \mathbb{N}} \mathcal{T}_n$ (Lemma 1.3.16) is more restrictive. For instance, Example 1.4.11 would not work at all in the classical setting.

For the second property, as far as we could tell, the closure- and interior-related conditions are solely to ensure that any element in \mathcal{A} may be approximated ‘strictly’ from above and also from below. Those lead later on to a seemingly more powerful property: the existence of (classical) flows [47, Definition 5.1.1 and Lemma 5.1.6]. We do get an ersatz though by means of \mathcal{A} -flows (Proposition 1.3.9). A comparison is made in Section 1.6.4.

The condition that $\sigma(\mathcal{A}) = \mathcal{B}$ is the Borel σ -algebra becomes a consequence of Theorem 1.3.17 in our setting.

Lastly, \mathcal{A}_n and g_n play the same role in both settings. However, a few differences are worth noting. In our setting, g_n is unequivocally defined from \mathcal{A}_n by the formula $g_n(A) = \bigcap_{A' \in \mathcal{A}_n : A \subseteq A'} A'$ and is \mathcal{A}_n -valued. In the classical one, g_n is not so easily accessible. To our opinion, this is mainly due to the condition $A \subseteq g_n(A)^\circ$. Taking the latter into account for the formula for g_n would intuitively give $g_n(A) = \bigcap_{A' \in \mathcal{A}_n : A \subseteq (A')^\circ} A'$ just like in [42, Definition 2.1]. However, it does not always hold as we illustrate in Example 1.6.3 below. To solve this problem, g_n is authorized to take values in $\mathcal{A}_n(u)$ instead.

EXAMPLE 1.6.3 (Why classical g_n must be $\mathcal{A}_n(u)$ -valued). *Consider the set $\mathcal{T} = \{a + ib \in \mathbf{C} : a, b \in [0, 1] \text{ and } a \times b = 0\}$ and its indexing collection $\mathcal{A} = \{\emptyset\} \cup \{[0, a] : a \in [0, 1]\} \cup \{[0, ib] : b \in [0, 1]\}$ from Example 1.3.15. Let us show that \mathcal{A} is also a classical indexing collection.*

Remark that \mathcal{T} is endowed with the (metrizable) topology induced by the usual one on \mathbf{C} . In particular, \mathcal{T} is compact and each element of \mathcal{A} is compact and connected. The first three properties of Definition 1.6.2 are straightforward to check for $\mathcal{T}_n = \mathcal{T}$ for all $n \in \mathbb{N}$. Let us show that the classical separability from above holds as well. Let $n \in \mathbb{N}$, denote $\mathcal{D}_n = \{k2^{-n} : 0 \leq k \leq 2^n\}$ and $\mathcal{A}_n = \{\emptyset\} \cup \{[0, d] : d \in \mathcal{D}_n\} \cup \{[0, id] : d \in \mathcal{D}_n\}$. Let us consider $A \in \mathcal{A}^$ and define the classical $g_n(A)$. The issue is that there is no element $A' \in \mathcal{A}_n$ such that $A \subseteq (A')^\circ$. Thus $g_n(A)$ cannot belong to \mathcal{A}_n . However, there is no issue once one accepts to have $g_n(A) \in \mathcal{A}_n(u)$. For all $x \in [0, 1 - 2^{-n}]$, denote by $d_n(x) = \lfloor 2^n x + 1 \rfloor 2^{-n} \in \mathcal{D}_n$ the smallest dyadic of order n strictly greater than x . Then define*

$$\begin{aligned} \forall a \in [0, 1], \quad g_n([0, a]) &= \begin{cases} [0, d_n(a)] \cup [0, i2^{-n}] & \text{if } a \leq 1 - 2^{-n}, \\ \mathcal{T} & \text{if } a > 1 - 2^{-n}, \end{cases} \\ \forall b \in [0, 1], \quad g_n([0, ib]) &= \begin{cases} [0, 2^{-n}] \cup [0, id_n(b)] & \text{if } b \leq 1 - 2^{-n}, \\ \mathcal{T} & \text{if } b > 1 - 2^{-n}. \end{cases} \end{aligned}$$

The map g_n then easily verifies all the required conditions.

This illustrates the importance to authorize g_n to be $\mathcal{A}_n(u)$ -valued in the classical setting. However, this is not without repercussions in the structures that follow. For instance, once \mathcal{A} is endowed with a metric $d_{\mathcal{A}}$, one of the first key quantities that one looks at is $d_{\mathcal{A}}(A, g_n(A))$ (see the shrinking mesh property in Definition 1.4.1 for our setting or [47, Assumption 1.3.2] for the classical one). So $d_{\mathcal{A}}$ has to be extendable to $\mathcal{A}(u)$ in a meaningful way.

1.6.2 SHAPE condition

In this section, we study the importance of the SHAPE condition and hopefully convince the reader that it is needed as a core hypothesis. In the classical literature, if one is given a set-

indexed process $X = \{X_A : A \in \mathcal{A}\}$, the existence of an additive extension ΔX for X (Proposition 1.2.44) is a supplementary hypothesis that is made most of the time, sometimes even implicitly. SHAPE has been known so far to be a sufficient condition for the existence of the additive extension of set-indexed processes (see [47, p.27]). What we prove is that SHAPE is also necessary.

Let us suppose that \mathcal{A} is only a collection of subsets of \mathcal{T} such that $\emptyset \in \mathcal{A}$ and $\mathcal{A}^* = \mathcal{A} \setminus \{\emptyset\}$ is closed under finite intersections. The classes \mathcal{C} , $\mathcal{C}_{(k)}$ and $\mathcal{C}(u)$ are then defined as before (Definition 1.2.35).

PROPOSITION 1.6.4 (Equivalent formulations to SHAPE). *The following properties are equivalent:*

- (i) \mathcal{A} verifies the SHAPE condition.
- (ii) For any $k \in \mathbf{N}^*$ and $A_1, \dots, A_k \in \mathcal{A}$, if $\bigcup_{1 \leq i \leq k} A_i \in \mathcal{A}$, then $\bigcup_{1 \leq i \leq k} A_i = A_j$ for some $j \in \llbracket 1, k \rrbracket$.
- (iii) The family $\{\mathbb{1}_A : A \in \mathcal{A}^*\}$ is linearly independent in $\mathbf{R}^{\mathcal{T}}$.

Proof. (i) \Rightarrow (iii) follows from Lemma 1.2.43. The reader may readily check that only the stability of \mathcal{A} under intersections has been used for the proof.

Let us suppose (iii) and consider $A_1, \dots, A_k \in \mathcal{A}$ such that $U = \bigcup_{1 \leq i \leq k} A_i \in \mathcal{A}$. By the inclusion-exclusion formula we know that $\mathbb{1}_U$ may be expressed as a linear combination of the $\mathbb{1}_{A_{j_1} \cap \dots \cap A_{j_i}}$ where $1 \leq j_1 < \dots < j_i \leq k$. According to (iii), it means that $\mathbb{1}_U = \mathbb{1}_{A_{j_1} \cap \dots \cap A_{j_i}}$ for some tuple $1 \leq j_1 < \dots < j_i \leq k$. In particular, $U = A_{j_1}$, which proves (iii) \Rightarrow (ii).

Let us suppose (ii) and consider $A, A_1, \dots, A_k \in \mathcal{A}$ such that $A \subseteq \bigcup_{1 \leq i \leq k} A_i$. In particular, $\bigcup_{1 \leq i \leq k} (A_i \cap A) = A$ belongs to \mathcal{A} . Since \mathcal{A} is closed under intersection, we may apply (ii) to conclude that there is $j \in \llbracket 1, k \rrbracket$ such that $A \subseteq (A_j \cap A)$. Hence $A \subseteq A_j$, which proves (ii) \Rightarrow (i). \square

PROPOSITION 1.6.5. *The following properties are equivalent:*

- (i) \mathcal{A} verifies the SHAPE condition.
- (ii) Any map $h : \mathcal{A} \rightarrow \mathbf{R}$ such that $h(\emptyset) = 0$ has an increment map $\Delta h : \mathcal{C}(u) \rightarrow \mathbf{R}$.
- (iii) Any map $h : \mathcal{A} \rightarrow \mathbf{R}$ such that $h(\emptyset) = 0$ has an associated linear functional $\flat : \mathcal{E} \rightarrow \mathbf{R}$.

Proof. (i) \Rightarrow (ii) \Leftrightarrow (iii) follows from Proposition 1.2.44 (once more, no unwanted property of \mathcal{A} has been used for the proof).

Let us suppose that (i) does not hold. According to Proposition 1.6.4, it means that we can find $\alpha_1, \dots, \alpha_k \in \mathbf{R}^*$ and $A_0, A_1, \dots, A_k \in \mathcal{A}^*$ pairwise distinct such that $\mathbb{1}_{A_0} = \sum_{1 \leq i \leq k} \alpha_i \mathbb{1}_{A_i}$. Then, the map h defined for all $A \in \mathcal{A}$ by $h(A) = \mathbb{1}_{A=A_0}$ cannot have an associated linear functional. Hence (iii) \Rightarrow (i). \square

In conclusion, as long as one has to consider an additive extension of a set-indexed process, the SHAPE condition is morally required. And additive extensions are indeed required here for several reasons:

- ◊ Constructing an integral with respect to a set-indexed process $X = \{X_A : A \in \mathcal{A}\}$ first requires to be able to do so for simple functions, which is exactly what the linear functional $\mathcal{X} = \{\mathcal{X}(f) : f \in \mathcal{E}\}$ associated with X is for.

- ◊ In the same way that rectangular increments for \mathbf{R}_+^2 -indexed processes are sometimes better than regular increments, increments ΔX_C for $C \in \mathcal{C}$ do have some advantages over the regular increments $X_A - X_{A'}$ for $A, A' \in \mathcal{A}$. In the classical setting, it is the case whenever one looks at weak and strong martingales [47, Definition 3.1.1]. Here, it will be the case when trying to understand the Hölder regularity of set-indexed processes (see Section 3.4).

We see one main objection to this statement, namely that there are classical indexing collections for which some processes still have an additive extension even though SHAPE does not hold. To our knowledge, there is in essence only one such example in the literature.

EXAMPLE 1.6.6 (Lower layers on $\mathcal{T} = \mathbf{R}_+^p$). This example is taken from [47, Example 1.2.4] and could be adapted to \mathbf{R}^p or some other manifold. A lower layer on $\mathcal{T} = \mathbf{R}_+^p$ is a compact subset $A \subseteq \mathbf{R}_+^p$ such that for all $t \in \mathbf{R}_+^p$, $t \in A$ implies $[0, t] \subseteq A$ (where $[0, t]$ is the usual ‘rectangle’ in \mathbf{R}_+^p). The collection \mathcal{A} of lower layers does not verify SHAPE (for $p \geq 2$) since all three of $A_1 = [0, 1] \times \{0\}^{p-1}$, $A_2 = \{0\}^{p-1} \times [0, 1]$ and $A_1 \cup A_2$ are lower layers, which would contradict (ii) in Proposition 1.6.4.

Hence in general, a set-indexed process does not have an associated increment map. However, we will see in Chapter 2 that the set-indexed Brownian motion $W = \{W_A : A \in \mathcal{A}\}$ (Example 2.2.12) still has an increment map. But as we see it, the deeper reason is that W may be extended all the way to \mathcal{B}_m in the first place while still being characterized by the rectangles in \mathbf{R}_+^p (i.e. the ‘usual’ indexing collection in this context). So a study of a process indexed by the lower sets can be still be attained in our case by studying an extension of one defined on the rectangles.

1.6.3 TIP assumption

For this endeavor, let us suppose that \mathcal{A} verifies all the conditions for being an indexing collection (Definition 1.2.1) but for the TIP assumption. The goal becomes to study under which conditions the map

$$\begin{aligned} A(\cdot) : \mathcal{T} &\longrightarrow \mathcal{A}^* \\ t &\longmapsto A(t) = \bigcap_{n \in \mathbf{N}} A_n(t) \end{aligned}$$

where $A_n(t) = \bigcap_{\substack{A \in \mathcal{A}_n \cup \{\mathcal{T}\} \\ t \in A}} A$ is one-to-one.

Injectivity of $t \mapsto A(t)$

The TIP map $t = A(\cdot)^{-1}$ might not always be well-defined in this context, but some approximations still are.

LEMMA 1.6.7 (TIP_n bijections). For all $n \in \mathbf{N}$, the map

$$\begin{aligned} \mathbf{t}_n : \mathcal{A}_n &\longrightarrow \mathcal{C}^\ell(\mathcal{A}_n) \\ A &\longmapsto A \setminus \bigcup_{A' \in \mathcal{A}_n : A' \subset A} A' \end{aligned}$$

is one-to-one and such that for all $t \in \mathcal{T}_n$, $\mathbf{t}_n(A_n(t)) = C_n(t)$.

Proof. Let $n \in \mathbf{N}$. By definition of $\mathcal{C}^\ell(\mathcal{A}_n)$, \mathbf{t}_n is surjective. Let $A, A' \in \mathcal{A}_n$ such that $\mathbf{t}_n(A) = \mathbf{t}_n(A')$. Writing the extremal decomposition for this element of $\mathcal{C}^\ell(\mathcal{A}_n)$, we get

$$A \setminus \bigcup_{A'' \in \mathcal{A}_n : A'' \subset A} A'' = A_0 \setminus \bigcup_{i=1}^k A_i = A' \setminus \bigcup_{A'' \in \mathcal{A}_n : A'' \subset A'} A''.$$

Since the extremal representation is obtained only by selecting elements from $\{A' \in \mathcal{A}_n : A' \subseteq A\}$ (see the proof of Proposition 1.2.36), we have $A = A_0$. Likewise, $A' = A_0$. Hence \mathbf{t}_n is injective.

Let $t \in \mathcal{T}_n$. Since $A_n(t)$ belongs to \mathcal{A}_n , it is the smallest element in \mathcal{A}_n containing t by definition. Thus for all $A' \in \mathcal{A}_n$ such that $A \subset A_n(t)$, $t \notin A'$. In particular, $\mathbf{t}_n(A_n(t))$ is the only element of $\mathcal{C}^\ell(\mathcal{A}_n)$ containing t , and so is equal to $C_n(t)$ by definition of the latter. \square

The injectivity of $t \mapsto A(t)$ is closely linked with another property that may be found in the classical setting under [47, Assumption 1.1.7].

PROPOSITION 1.6.8 (Dissecting system). *The map $t \mapsto A(t)$ is injective if and only if \mathcal{C}^ℓ is a dissecting system, i.e. for all $s, t \in \mathcal{T}$ such that $s \neq t$, there exists $C, C' \in \mathcal{C}^\ell$ such that $s \in C$, $t \in C'$ and $C \cap C' = \emptyset$.*

Proof. Suppose that $t \mapsto A(t)$ is injective and consider $s, t \in \mathcal{T}$ such that $s \neq t$. By injectivity, there exists $n \in \mathbb{N}$ such that $A_n(s) \neq A_n(t)$ while they both belong to \mathcal{A}_n . Since the extremal representation of $C_n(s)$ is of the form $C_n(s) = A_n(s) \setminus \dots$ and likewise for $C_n(t)$, we have $C_n(s) \neq C_n(t)$. Hence $C_n(s) \cap C_n(t) = \emptyset$ since the elements of $\mathcal{C}^\ell(\mathcal{A}_n)$ are pairwise disjoint. So \mathcal{C}^ℓ is a dissecting system.

Conversely, suppose that \mathcal{C}^ℓ is a dissecting system and consider $s, t \in \mathcal{T}$ such that $A(s) = A(t)$. Hence for all $n \in \mathbb{N}$, $A_n(s) = A_n(t)$ and thus $C_n(s) = C_n(t)$ by Lemma 1.6.7. Since the only elements in \mathcal{C}^ℓ that contain s (resp. t) are of the form $C_n(s)$ (resp. $C_n(t)$) for $n \in \mathbb{N}$, it means that $s = t$. The injectivity of $t \mapsto A(t)$ follows. \square

This feature tells us in particular that one may ‘pick out’ any element $t \in \mathcal{T}$ by means of the sequence $(C_n(t))_{n \in \mathbb{N}}$. It has been used in [41] to define the *point mass jumps* of a set-indexed process (Definition 3.3.10). This article shows for instance that although a set-indexed Gaussian process may have discontinuous sample paths, it remains ‘jumpless’ nonetheless [41, Theorem 7.3].

Since understanding the jump structure of set-indexed processes will be an important point of focus for us (see Section 3.4.2), the injectivity of $t \mapsto A(t)$ is a must-have.

Surjectivity of $t \mapsto A(t)$

What we actually show here is that surjectivity is a non-issue since one may always slightly modify the indexing collection \mathcal{A} into another one $[\mathcal{A}]$ for which the surjectivity of $t \mapsto A(t)$ holds and the previous existing structures (namely, \preceq , \mathbf{m} and $d_{\mathcal{A}}$) naturally extend.

EXAMPLE 1.6.9. *Suppose that $\mathcal{T} = [0, 1)$ and $\mathcal{A} = \{[0, t] : t \in [0, 1)\} \cup \{\emptyset\} \cup \{[0, 1)\}$. The class \mathcal{A} could be an indexing collection but for the fact that the element $[0, 1)$ is ‘tipless’. A natural way here is to complete the space $[0, 1)$ so that the extra-point 1 given by the completion will serve as a tip for $[0, 1)$. But since the tip has to belong to the set, we need to replace $[0, 1)$ by $[0, 1]$.*

Any measure \mathbf{m} on $[0, 1)$ could be extended to $[0, 1]$ by imposing that $\mathbf{m}([0, 1]) = \mathbf{m}([0, 1))$. Likewise, the relation $d_{[\mathcal{A}]}(A, [0, 1]) = d_{\mathcal{A}}(A, [0, 1))$ would extend the metric. Remark that those extensions modify neither the poset structure of \mathcal{A} nor the ‘compatible’ aspect of \mathbf{m} and $d_{\mathcal{A}}$ in the sense of Definitions 1.3.1 and 1.4.1.

What follows is just a fancier version of Example 1.6.9. Suppose that \mathbf{m} (resp. $d_{\mathcal{A}}$) is a measure on $\sigma(\mathcal{A})$ (resp. metric on \mathcal{A}). Let us introduce the following:

1. The collection $[\mathcal{A}] = \{[\emptyset', A] : A \in \mathcal{A}\}$.

2. The measure $[m]$ on the measurable space $(\mathcal{A}, \sigma([\mathcal{A}]))$ which is the pushforward of m by $t \mapsto A(t)$.
3. The metric $d_{[\mathcal{A}]}$ on $[\mathcal{A}]$ defined for all $A, A' \in \mathcal{A}$ by $d_{[\mathcal{A}]}([\emptyset', A], [\emptyset', A']) = d_{\mathcal{A}}(A, A')$.

THEOREM 1.6.10 (TIP embedding). *The map $\mathcal{A} \ni A \mapsto [\emptyset', A] \in [\mathcal{A}]$ is one-to-one and order embedding. Moreover, $[\mathcal{A}]$ is an indexing collection on \mathcal{A}^* for which the tip map is $(A \mapsto [\emptyset', A])^{-1}$ and the following holds:*

- ◇ For all $A \in \mathcal{A}$, $[m]([\emptyset', A]) = m(A)$, thus if m is compatible, so is $[m]$.
- ◇ The map $A \mapsto [\emptyset', A]$ is an isometry, thus if $d_{\mathcal{A}}$ is compatible, so is $d_{[\mathcal{A}]}$.

Proof. The map $\mathcal{A} \ni A \mapsto [\emptyset', A] \in [\mathcal{A}]$ is one-to-one by definition. Being order embedding and isometric are just straightforward matters. Let us show that $[\mathcal{A}]$ is an indexing collection on \mathcal{A}^* . The elements relative to $[\mathcal{A}]$ as an indexing collection will be denoted using '[']' such as $[g_n]$.

1. $\emptyset = [\emptyset', \emptyset] \in [\mathcal{A}]$ and $[\mathcal{A}] \setminus \{\emptyset\}$ is closed under intersections since \mathcal{A}^* is and for any collection $(A_i)_{i \in I}$ of sets in \mathcal{A} , we have

$$\left[\emptyset', \bigcap_{i \in I} A_i \right] = \bigcap_{i \in I} [\emptyset', A_i]. \quad (1.6.1)$$

2. For all $n \in \mathbf{N}$, define $[\mathcal{A}_n] = \{[\emptyset', A] : A \in \mathcal{A}_n\}$. The corresponding maps $[g_n]$ are given by

$$\forall A \in \mathcal{A}, \quad [g_n]([\emptyset', A]) = \bigcap_{\substack{B \in [\mathcal{A}_n] \cup \{\mathcal{A}^*\}: \\ [\emptyset', A] \subseteq B}} B = \bigcap_{\substack{A' \in \mathcal{A}_n \cup \{\mathcal{T}\}: \\ A \subseteq A'}} [\emptyset', A'] = [\emptyset', g_n(A)]$$

where the first equality is by definition, the second due to the previous order embedding and the third to (1.6.1). The approximation from above property readily follows.

3. For all $n \in \mathbf{N}$ and $A \in \mathcal{A}$, we have

$$[A_n](A) = \bigcap_{\substack{B \in [\mathcal{A}_n] \cup \{\mathcal{A}^*\}: \\ A \in B}} B = \bigcap_{\substack{A' \in \mathcal{A}_n \cup \{\mathcal{T}\}: \\ A \subseteq A'}} [\emptyset', A'] = [\emptyset', g_n(A)] \quad (1.6.2)$$

for the same reasons as above. From (1.6.1) and (1.6.2), we get for all $A \in \mathcal{A}$,

$$\bigcap_{n \in \mathbf{N}} [A_n](A) = \left[\emptyset', \bigcap_{n \in \mathbf{N}} g_n(A) \right] = [\emptyset', A].$$

Hence $A \mapsto \bigcap_{n \in \mathbf{N}} [A_n](A)$ is one-to-one and the TIP bijection is indeed the one claimed in Theorem 1.6.10.

4. The SHAPE condition for $[\mathcal{A}]$ is just a consequence of the same condition for \mathcal{A} and the fact that for all $k \in \mathbf{N}^*$ and $A_1, \dots, A_k \in \mathcal{A}$,

$$\bigcup_{i=1}^k [\emptyset', A_i] = \left[\emptyset', \bigcup_{i=1}^k A_i \right].$$

Hence $[\mathcal{A}]$ is an indexing on \mathcal{A} .

Concerning the measure $[\mathbf{m}]$, we have for all $A \in \mathcal{A}$,

$$[\mathbf{m}](\llbracket \emptyset', A \rrbracket) = \mathbf{m}\left(\underbrace{\{t \in \mathcal{T} : A(t) \subseteq A\}}_{=A}\right) = \mathbf{m}(A).$$

The rest follows immediately. \square

REMARKS 1.6.11. \diamond If we were to use the vocabulary of Section 1.2.2, we could say that $[\mathcal{A}]$ is the indexing collection associated with the indexing semilattice $(\mathcal{A}^*, \subseteq)$.

\diamond This construction also provides injectivity for free, but if injectivity did hold in the first place, then $[\mathcal{A}]$ could be seen as an indexing collection on $\mathcal{T} \cup \mathcal{T}'$ where \mathcal{T}' is the set of ‘missing tips’.

\diamond In the classical setting, the TIP assumption is found under [47, Assumption 2.4.2] and suppose that for all $A \in \mathcal{A}$, $A \setminus \bigcup_{A' \in \mathcal{A} : A' \subset A} A' \neq \emptyset$. A TIP map is then given by an appeal to the axiom of choice in order to choose an element $t_A \in A \setminus \bigcup_{A' \in \mathcal{A} : A' \subset A} A'$ for all $A \in \mathcal{A}$. Let us point out that our TIP assumption implies the classical one and does not rely on the axiom of choice.

1.6.4 ‘Revamped’ vs ‘classical’ flows

Flows have been introduced in [23, §4] while developing a stochastic integration with \mathbf{R}_+^2 -indexed martingales as integrators. They have then been studied in the set-indexed setting and are described in [47, Section 5.1].

Flows are a very interesting tool. As mentioned in Section 1.3.2, they constitute the link between set-indexed and one-dimensional processes. They may serve as a guide for intuition when defining set-indexed processes. For instance, a *set-indexed fractional Brownian motion (sifBm)* with control measure \mathbf{m} and Hurst exponent $H \in (0, 1/2]$ is a centered Gaussian process $W^H = \{W_A^H : A \in \mathcal{A}\}$ with covariance function given by

$$\forall A, A' \in \mathcal{A}, \quad \text{Cov}(W_A^H, W_{A'}^H) = \frac{1}{2} [\mathbf{m}(A)^{2H} + \mathbf{m}(A')^{2H} - \mathbf{m}(A \Delta A')^{2H}] \quad (1.6.3)$$

where the expression is non-negative definite by [39, Lemma 2.9].

This definition might appear arbitrary at first glance, but becomes less so if one remarks that for any geodesic \mathcal{A} -flow $\phi \in \Phi(\mathcal{A})$, $\nu(\phi)^{-H}(W^H)^\phi$ is the usual $[0, 1]$ -indexed fractional Brownian motion of Hurst exponent H .

Since a flow at its bare minimum is the same in both theories ([47, Definition 5.1.1] and Definition 1.3.4), it is more meaningful to compare the properties that a ‘typical’ \mathcal{A} -flow usually has. When $A_0, A_1 \in \mathcal{A}$ are such that $A_0 \subseteq A_1$, both [47, Lemma 5.1.6] and Proposition 1.3.9 ensure the existence of a *continuous* \mathcal{A} -flow ϕ from A_0 to A_1 . However, the notion of continuity depends on the setting. In the classical one, ϕ is outer continuous and *inner continuous* in the sense that

$$\forall t \in (0, 1], \quad \phi(t) = \overline{\bigcup_{s < t} \phi(s)}. \quad (1.6.4)$$

In our setting, ϕ is continuous with respect to the metric d_m . Under the additional assumption that \mathbf{m} is a Radon measure, being continuous in the classical sense implies being continuous in our sense.

Due to our choice of separating order and topological assumptions for the indexing collection, there is no particular reason for (1.6.4) to hold in our setting. However, the continuity with

respect to d_m is an ersatz that works well enough for our purposes. The tradeoff is that we have a quicker construction of geodesic flow and do not require the additional classical assumption that m is Radon, whereas classical flows obey more properties.

1.7 Perspectives

The framework developed here enables us to tackle a variety of situations ranging from the multiparameter case to some \mathbf{R} -trees. By ‘some’, we mean that unfortunately not all \mathbf{R} -trees may be endowed with an indexing collection.

EXAMPLE 1.7.1 (Infinite binary tree with ends). *Suppose that \mathcal{T} is the ends compactification [31, Section 3.4.2] of the infinite binary \mathbf{R} -tree (Figure 1.15). Since \mathcal{T} has an uncountable number of leaves and each of them is maximal for the order on \mathcal{T} , the separability from above cannot hold. Hence, even though \mathcal{T} is an \mathbf{R} -tree, it cannot be endowed with an indexing collection.*

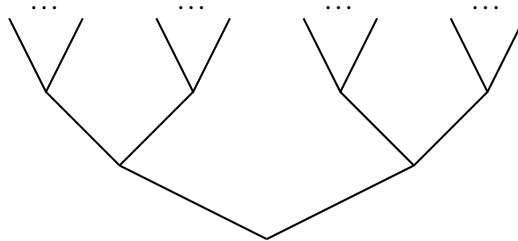


Figure 1.15: Infinite binary \mathbf{R} -tree where each leaf corresponds to an ever-increasing path in the discrete infinite binary tree.

However, any (separable) \mathbf{R} -tree is the limit — for the Gromov-Hausdorff topology — of discrete trees, which can be endowed with indexing collections of their own. From that remark, we would find interesting to develop a topology on the ‘set’ of all indexing collections in the spirit of the Gromov-Hausdorff topology for \mathbf{R} -trees. Indeed, objects arising as limits of a structure encompassing both finite dimensional vector spaces and \mathbf{R} -trees may very well serve as relevant random environments. Hopefully, this setting would cover:

- ◊ \mathbf{R} -trees, and especially celebrated objects such as Aldous’s Continuum Random Tree (CRT) [7],
- ◊ manifolds as mentioned in Example 1.2.29, hopefully extending works like [2, 61],
- ◊ simplicial complexes from homotopy theory through gluing procedures, with a view to reaching random complexes described in [103] for instance.

1.8 Perspectives (français)

Le cadre développé ici nous permet d'aborder une variété de situations allant du cas multi-paramétrique à certains \mathbf{R} -arbres. Par "certains", nous voulons bien dire que tous les \mathbf{R} -arbres ne peuvent pas être munis d'une *indexing collection*.

EXEMPLE 1.8.1 (Arbre binaire infini compactifié par les bouts.). *Supposons que \mathcal{T} est le compactifié par les bouts [31, Section 3.4.2] du \mathbf{R} -arbre binaire infini (Figure 1.16). Comme \mathcal{T} a une infinité indénombrable de feuilles et que chaque feuille est maximale pour la relation d'ordre sur \mathcal{T} , la séparabilité par au-dessus ne peut avoir lieu. Donc, même si \mathcal{T} est un \mathbf{R} -arbre, il ne peut être équipé d'une *indexing collection*.*

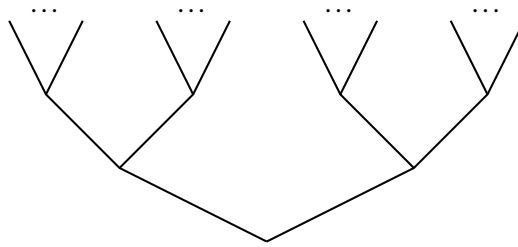


Figure 1.16: \mathbf{R} -arbre binaire infini où chaque feuille correspond à un chemin strictement croissant dans l'arbre binaire discret.

Cependant, tout \mathbf{R} -arbre (séparable) devrait être la limite — pour la topologie de Gromov-Hausdorff — d'arbres discrets, qui peuvent eux être équipés d'*indexing collections*. De ce fait, nous trouverions intéressant de pouvoir développer une topologie sur l'"ensemble" des *indexing collections* dans l'esprit de la topologie de Gromov-Hausdorff pour les \mathbf{R} -arbres. En effet, les objets émergeant comme limites de structures pouvant être des variétés différentielles ou des \mathbf{R} -arbres promet d'être très riche, et pourrait avec un peu de chance servir d'une solide assise pour des modèles d'environnements aléatoires. Avec un peu d'espoir, ce cadre pourrait englober :

- ◊ des \mathbf{R} -arbres tels que le célèbre arbre continu d'Aldous (CRT) [7],
- ◊ des variétés différentielles comme mentionné dans l'Exemple 1.2.29, faisant on l'espère avec des travaux comme [2, 61],
- ◊ des complexes simpliciaux en théorie homotopique grâce aux procédures de recollement, avec en tête d'atteindre les complexes aléatoires décrits dans [103] par exemples

Distributional properties of generalized processes **2**

Contents

2.1	Introduction	54
2.2	Extensions of set-indexed processes	54
2.2.1	Prerequisites on modular spaces	54
2.2.2	Set-indexed processes as stochastic measures	56
2.3	Processes with independent increments	59
2.3.1	Equivalent definitions	59
2.3.2	Infinitely divisible processes	61
2.3.3	Martingale inequalities	70
2.3.4	A 0-1 law	76
2.4	Interlude: set-indexed Lévy processes	77
2.4.1	Definition and examples	77
2.4.2	Weak Lévy-Itô decomposition	79
2.5	Processes with stationary increments	81
2.5.1	General philosophy for set-indexed stationarity	82
2.5.2	\mathcal{A} -increment stationarity	83
2.5.3	\mathcal{C} -exchangeability	87
2.5.4	Functional exchangeability	91
2.5.5	Examples of \mathcal{C} -exchangeable processes	92
2.5.6	Representation theorems	95
2.6	Perspectives	103
2.6.1	Characteristics for stochastic measures	103
2.6.2	Towards functional martingale inequalities	104
2.6.3	More representation theorems	104
2.7	Perspectives (français)	105
2.7.1	Caractéristiques pour les mesures stochastiques	105
2.7.2	Vers des inégalités fonctionnelles de type martingale	105
2.7.3	Davantage de théorèmes de représentation	105

2.1 Introduction

The whole chapter is devoted to the study of the distributional properties of a given set-indexed process $X = \{X_A : A \in \mathcal{A}\}$ and several of its extensions. Namely, we will turn our attention to the following ones progressively introduced in Section 2.2:

1. The increment map $\Delta X = \{\Delta X_U : U \in \mathcal{C}(u)\}$ given by Proposition 1.2.44.
2. The stochastic measure $\Delta X = \{\Delta X_B : B \in \mathcal{B}_{\text{loc}}\}$ given by Theorem 2.2.8.
3. The linear process $\mathcal{X} = \{\mathcal{X}(f) : f \in \mathcal{L}\}$ given by Theorem 2.2.10.

All of those extend X in one way or another, but may not always exist apart from ΔX . We review in this section some conditions that ensure their existence. This is naturally linked to the existence of an integral with respect to X , which has been extensively studied by Kwapien, Rajput, Rosiński, Urbanik, Woyczyński and more.

We then proceed to study the distributional properties of X . Two cases are of particular interest to us:

- ◊ When X has independent increments (Section 2.3), which is already well-known due to the work of Rajput and Rosiński [83]. As such, Sections 2.3.1 and 2.3.2 are mainly expository and are meant to link the literature with the set-indexed setting. Then, we take up on the occasion to prove two kinds of results: martingale inequalities (Section 2.3.3) and a 0-1 law (Section 2.3.4), both of which will be used to establish regularity results in Chapter 3.
- ◊ When X has stationary increments (Section 2.5), which may be formulated in several — not equivalent — ways in our general setting. The starting idea steams from the works of Herbin and Merzbach [39, 40, 41] who used m to measure the size of the increments. In Section 2.5.1, we explain the general scheme that stems from their approach and propose several definitions all base on the same notion of *equality in configuration* (Definition 2.5.4): \mathcal{A} -*increment stationarity* (Section 2.5.2) and \mathcal{C} -*exchangeability* (Section 2.5.3) among others. For the latter, we prove representation theorems (Theorems 2.5.30 and 2.5.31) which generalize Bühlmann's [55, Theorem 1.19] and Kallenberg's [55, Theorem 3.15] for one-dimensional processes with exchangeable increments.

2.2 Extensions of set-indexed processes

2.2.1 Prerequisites on modular spaces

In this section, we recall without proof the few facts about modular spaces that will be required in building the extensions mentioned in the introduction. For more details about the general theory of modular spaces, we refer to [76]. In the abundant literature on the subject, not all authors have had the same definitions. We tried to follow [62], but tailored to the scope of our needs.

DEFINITION 2.2.1 (Modular, [62, Section 0.7]). *Let V be a real vector space. A modular on V is a map $\rho : V \rightarrow \mathbf{R}_+$ which verifies the following properties:*

1. (*Separation*). For all $u \in V$, $\rho(u) = 0$ if and only if $u = 0_V$.

2. (Symmetry and scaling compatibility). For all $u \in V$, the map $\mathbf{R} \ni \lambda \mapsto \rho(\lambda u)$ is even and continuous on \mathbf{R} and non-decreasing on \mathbf{R}_+ .
3. (Quasi triangle inequality). There exists a constant $\kappa > 0$ such that

$$\forall u, v \in V, \quad \rho(u + v) \leq \kappa(\rho(u) + \rho(v)). \quad (2.2.1)$$

The pair (V, ρ) is called a modular space.

PROPOSITION 2.2.2 ([62, Section 0.7]). A modular space (V, ρ) has a natural topology where open sets are subsets $U \subseteq V$ such that for any $u \in U$, there is $r > 0$ such that $\{v \in V : \rho(u - v) < r\} \subseteq U$. This makes (V, ρ) into a topological vector space where a sequence $(u_k)_{k \in \mathbf{N}}$ converges to $u \in V$ if and only if $\rho(u_k - u) \rightarrow 0$ as $k \rightarrow \infty$.

EXAMPLE 2.2.3 (Space of random variables). Let $L^0(\Omega)$ denote the space of (equivalence classes of) real-valued random variables. The map $\rho_0 : X \mapsto \mathbf{E}[|X| \wedge 1]$ is a modular on $L^0(\Omega)$. The resulting topology is metrizable — with the metric $(X, Y) \mapsto \rho_0(X - Y)$ — complete and corresponds to the convergence in probability.

A kind of modular spaces are Musielak-Orlicz spaces or generalized Orlicz spaces. As it has been shown in [83, 99], they naturally arise when building a stochastic integral with respect to processes with independent increments. We refer to [36] for a modern and detailed exposition on those spaces and [62] for their use in stochastic integration.

Recall that $L^0(\mathbf{m}) = L^0(\mathcal{T}, \mathcal{B}, \mathbf{m})$ stands for the space of (equivalence classes of) real-valued Borel maps $f : \mathcal{T} \rightarrow \mathbf{R}$.

DEFINITION 2.2.4 (Musiela-Orlicz space, [62, Section 0.8]). Let $\phi : \mathcal{T} \times \mathbf{R}_+ \rightarrow \mathbf{R}_+$ be a map which verifies the following properties:

1. (Measurability). For all $f \in L^0(\mathbf{m})$, the map $\phi(\cdot, |f(\cdot)|)$ is measurable.
2. (Scaling compatibility). For all $t \in \mathcal{T}$, the map $\phi(t, \cdot)$ is continuous, non-decreasing and $\phi(t, 0) = 0$.
3. (Moderate growth). There exists a constant $\kappa > 0$ such that

$$\forall (t, x) \in \mathcal{T} \times \mathbf{R}_+, \quad \phi(t, \kappa^{-1}x) \leq \kappa \phi(t, x).$$

Then the map ρ_ϕ given by

$$\forall f \in L^0(\mathbf{m}), \quad \rho_\phi(f) = \int_{\mathcal{T}} \phi(t, |f(t)|) \mathbf{m}(dt)$$

induces a modular on the space

$$L^\phi = \{f \in L^0(\mathbf{m}) : \rho_\phi(f) < \infty\}$$

which is called the Musielak-Orlicz space (of modular ϕ).

PROPOSITION 2.2.5 (Luxemburg norm). *The Musielak-Orlicz space L^ϕ may be characterized equivalently by*

$$L^\phi = \{f \in L^0(\mathbf{m}) : \|f\|_\phi < \infty\}$$

where $\|\cdot\|_\phi$ is the Luxemburg (quasi)norm given by

$$\forall f \in L^0(\mathbf{m}), \quad \|f\|_\phi = \inf\{c \geq 0 : \rho_\phi(c^{-1}f) \leq 1\} \quad (2.2.2)$$

with the convention $\inf \emptyset = \infty$.

We emphasize the fact that even if $\|\cdot\|_\phi$ is called a ‘norm’ by custom, it is not one in general but only a *quasinorm*, i.e. a map that verifies the separability, symmetry and quasi triangle inequality (2.2.1). In particular, $\|\cdot\|_\phi$ is also a modular.

EXAMPLE 2.2.6. *For the map $\phi : (t, x) \mapsto |x|^p$ where $p \in \mathbf{R}_+^*$, the space L^ϕ corresponds to the usual space $L^p(\mathbf{m})$.*

2.2.2 Set-indexed processes as stochastic measures

We mainly follow the construction given in [62, Chapter 7] for stochastic integrals on $[0, 1]$ while making the necessary adjustments to the set-indexed theory. We refer to the bibliographical notes in *loc. cit.* for further references.

We also note that a set-indexed theory of stochastic integration has been developed in [72, 86]. However, the integrands may be non-deterministic, which is not our focus for now.

The starting point is the set-indexed process X . From it, Proposition 1.2.44 gives the existence of the increment map ΔX as well as the linear functional \mathcal{X} which serve as embryos of a measure and an integral respectively.

Let us start with the increment map ΔX . It is a $L^0(\Omega)$ -valued (finitely) additive map defined on the ring $\mathcal{C}(u)$. In particular, it can be seen as a vector-valued premeasure, which should be extendable under some additional continuity assumption. Denote $\mathcal{B}_{loc} = \{B \cap \mathcal{T}_n : B \in \mathcal{B}, n \in \mathbf{N}\}$.

DEFINITION 2.2.7 (Stochastic measure, [62, Section 7.1]). *] A stochastic measure is a map $M : \mathcal{B}_{loc} \rightarrow L^0(\Omega)$ which verifies the following properties:*

1. $M_\emptyset = 0$ a.s.
2. (σ -additivity). *For all sequence $(B_j)_{j \in \mathbf{N}}$ of pairwise disjoint elements of \mathcal{B}_{loc} such that $\bigsqcup_{j \in \mathbf{N}} B_j \in \mathcal{B}_{loc}$ (equivalently, $B_j \subseteq \mathcal{T}_n$ for some n and all j),*

$$\sum_{j=0}^k M_{B_j} \xrightarrow{\mathbf{P}} M_{\bigsqcup_{j \in \mathbf{N}} B_j} \quad \text{as} \quad k \rightarrow \infty.$$

THEOREM 2.2.8 (Extension of ΔX , adapted from [62, Theorem B.1.1]). *] The following statements are equivalent:*

- (i) *There exists a stochastic measure $\Delta \mathbf{X} = \{\Delta \mathbf{X}_B : B \in \mathcal{B}_{loc}\}$ such that $\Delta \mathbf{X}|_{\mathcal{C}} = X$.*
- (ii) *For any $n \in \mathbf{N}$ and any sequence $(U_k)_{k \in \mathbf{N}}$ in $\mathcal{C}(u)$ included in \mathcal{T}_n , if $\limsup_{k \rightarrow \infty} U_k = \emptyset$, then $\Delta X_{U_k} \xrightarrow{\mathbf{P}} 0$ as $k \rightarrow \infty$.*

- (iii) For any $n \in \mathbf{N}$ and any sequence $(U_k)_{k \in \mathbf{N}}$ in $\mathcal{C}(u)$ included in \mathcal{T}_n , if either $U_k \downarrow \emptyset$ or the U_k 's are pairwise disjoint, then $\Delta X_{U_k} \xrightarrow{P} 0$ as $k \rightarrow \infty$.
- (iv) a. For any $n \in \mathbf{N}$ and any sequence $(U_k)_{k \in \mathbf{N}}$ in $\mathcal{C}(u)$ included in \mathcal{T}_n , if $U_k \downarrow \emptyset$ and a subsequence of $(\Delta X_{U_k})_{k \in \mathbf{N}}$ converges in probability to some $Z \in L^0(\Omega)$, then $Z = 0$ a.s.
- b. For any $n \in \mathbf{N}$ and any sequence $(C_k)_{k \in \mathbf{N}}$ in \mathcal{C} included in \mathcal{T}_n , if the C_k 's are pairwise disjoint, then the series $\sum_k \Delta X_{C_k}$ converges in probability.

Under those conditions, ΔX is unique up to a version and verifies $\Delta X|_{\mathcal{C}(u)} = \Delta X$. The process ΔX is called stochastic measure associated with X .

What we mean here by 'unique up to a version' here is that if ΔX and $\Delta X'$ are two measures processes associated with X , then for all $B \in \mathcal{B}_{\text{loc}}$, $\Delta X_B = \Delta X'_B$ a.s.

Proof. Suppose that $\mathcal{T} \in \mathcal{A}_0(u)$ (the general case is a straightforward consequence of this one). In particular, $\mathcal{C}(u)$ becomes an algebra of sets instead of just a ring, i.e. it is stable under complement and not just set difference. Moreover, $\mathcal{B}_{\text{loc}} = \mathcal{B}$. Under this additional hypothesis, [62, Theorem B.1.1] together with the remark that follows it proves (i) \Leftrightarrow (ii) \Leftrightarrow (iii) \Rightarrow (iv).

Suppose that (iv) holds and let us prove (iii). Consider a sequence $(U_k)_{k \in \mathbf{N}}$ in $\mathcal{C}(u)$ of pairwise disjoint elements. Then, there exists a sequence $(C_j)_{j \in \mathbf{N}}$ of pairwise disjoint elements of \mathcal{C} and an increasing sequence of integers $(n_k)_{k \in \mathbf{N}}$ such that for all $k \in \mathbf{N}$, $U_k = \bigsqcup_{j=n_k}^{n_{k+1}-1} C_j$. Applying (iv)b. yields the convergence of the series $\sum_j \Delta X_{C_j}$. In particular, $\Delta X_{U_k} = \sum_{j=n_k}^{n_{k+1}-1} \Delta X_{C_j} \xrightarrow{P} 0$ as $k \rightarrow \infty$.

Now, let us suppose instead that $U_k \downarrow \emptyset$ as $k \rightarrow \infty$. According to (iv)a., it is enough to prove that $(\Delta X_{U_k})_{k \in \mathbf{N}}$ converges. Applying (iv)b. to $(U_k \setminus U_{k+1})_{k \in \mathbf{N}}$, we know that the series $\sum_k (\Delta X_{U_k} - \Delta X_{U_{k+1}})$ converges. Hence $\Delta X_{U_p} - \Delta X_{U_q} = \sum_{k=p}^{q-1} (\Delta X_{U_k} - \Delta X_{U_{k+1}}) \xrightarrow{P} 0$ as $p, q \rightarrow \infty$ ($p \leq q$). So $(\Delta X_{U_k})_{k \in \mathbf{N}}$ is a Cauchy sequence and thus converges since $L^0(\Omega)$ is complete. The equivalence follows.

Let us establish unicity. Suppose that ΔX and $\Delta X'$ are two stochastic measures associated with X and consider the set $\mathcal{B}' = \{B \in \mathcal{B} : \Delta X_B = \Delta X'_B \text{ a.s.}\}$. Due to the properties of ΔX and $\Delta X'$, \mathcal{B}' is a λ -system. By definition, it also contains the π -system \mathcal{A} . Hence, by Dynkin's π - λ theorem [59, Theorem 1.19], $\mathcal{B}' = \mathcal{B}$. The unicity follows. \square

Now that the stochastic measure is well-defined, let us move on to the integral itself. All the results whose proofs are skipped may be found in [62, Section 7.1]. In the same way we began with the increment map ΔX for the stochastic measure, we take here the linear functional \mathcal{X} as our starting point for the integral. Define

$$\forall f \in \mathcal{E}, \quad \rho_X(f) = \sup_{\substack{g \in \mathcal{E}: \\ |g| \leq 1}} \mathbf{E}[|\mathcal{X}(fg)| \wedge 1]. \quad (2.2.3)$$

Remark that ρ_X defines a modular on \mathcal{E} based on the usual modular space $(L^0(\Omega), \mathbf{E}[|\cdot| \wedge 1])$ mentioned in Example 2.2.3. In order to have a 'good' integration theory, it is desirable for the space of so-called 'integrable' functions to be complete. In order to do so, we need to define the following outer measure:

$$\forall B \subseteq \mathcal{T}, \quad \mathbf{m}_X(B) = \inf_{\substack{(C_k)_{k \in \mathbf{N}} \in \mathcal{C}^{\mathbf{N}}: \\ B \subseteq \bigcup_{k \in \mathbf{N}} C_k}} \sum_{k \in \mathbf{N}} \rho_X(\mathbb{1}_{C_k}). \quad (2.2.4)$$

Indeed, like in usual integration theory, not all integrable functions are pointwise limits of simple functions, but the statement becomes true if one requires the pointwise convergence to hold almost everywhere.

DEFINITION 2.2.9 (Integrability, [62, Definition 7.1.1]).] A map $f : \mathcal{T} \rightarrow \mathbf{R}$ is said to be X -integrable if there exists a sequence $(f_k)_{k \in \mathbf{N}}$ of simple functions in \mathcal{E} such that:

1. (Almost everywhere convergence). For \mathbf{m}_X -a.e. $t \in \mathcal{T}$, $f_k(t) \rightarrow f(t)$ as $k \rightarrow \infty$.
2. (Cauchy sequence for ρ_X). $\rho_X(f_j - f_k) \rightarrow 0$ as $j, k \rightarrow \infty$.

The function f is locally X -integrable if for all $A \in \mathcal{A}$, $f|_A$ is X -integrable. The space of all X -integrable (resp. locally X -integrable) functions is denoted by $L(X)$ (resp. $L_{loc}(X)$).

THEOREM 2.2.10 (Extension of \mathcal{X} , adapted from [62, Theorem 7.1.2]).] The following statements are equivalent:

- (i) X extends to a stochastic measure $\Delta X = \{\Delta X_B : B \in \mathcal{B}_{loc}\}$.
- (ii) For any $n \in \mathbf{N}$ and any sequence $(f_k)_{k \in \mathbf{N}}$ of simple functions converging pointwise to 0 and such that for all $k \in \mathbf{N}$, $\text{supp}(f_k) \subseteq \mathcal{T}_n$ and $|f_k| \leq 1$, $\rho_X(f_k) \rightarrow 0$ as $k \rightarrow \infty$.

Under those conditions, the modular ρ_X extends to a modular on $L(X)$, which then becomes a complete linear metric space where \mathcal{E} is a dense subset. In particular, there exists a unique continuous linear map $\mathcal{X} = \{\mathcal{X}(f) : f \in L(X)\}$ such that $\mathcal{X}|_{\mathcal{E}} = \mathcal{X}$. The process \mathcal{X} is called the linear process associated with X . We also use the notation

$$\mathcal{X}(f) = \int_{\mathcal{T}} f dX = \int_{\mathcal{T}} f(t)X(dt)$$

when we want to emphasize that $\mathcal{X}(f)$ should be seen as the stochastic integral of f with respect to X .

REMARK 2.2.11. Actually, this construction enables us to extend ΔX even further than \mathcal{B}_{loc} , that is on the class $\mathcal{B}(X) = \{B \in \mathcal{B} : \mathbb{1}_B \in L(X)\}$ by the formula $\Delta X = \mathcal{X}(\mathbb{1}_B)$. Even though the definition of stochastic measure does not usually consider this extension, this extra step is no hurdle and the extended process $\Delta X = \{\Delta X_B : B \in \mathcal{B}(X)\}$ will still be referred to as a stochastic measure.

EXAMPLE 2.2.12 (White noise and isonormal process). Remark that there is in general no hope to be able to extend a stochastic measure to the whole σ -algebra \mathcal{B} . Indeed, even deterministic signed measures cannot do so in general. But let us give a probabilistic example.

Consider a set-indexed Brownian motion (siBm) $W = \{W_A : A \in \mathcal{A}\}$ with control measure \mathbf{m} , i.e. a centered Gaussian process with covariance function given by

$$\forall A, A' \in \mathcal{A}, \quad \text{Cov}(W_A, W_{A'}) = \mathbf{m}(A \cap A') \quad (2.2.5)$$

(where we remark that it indeed corresponds to the siBm (1.6.3) with Hurst exponent $H = 1/2$).

Then, it is well-known that W extends to a stochastic measure $\Delta W = \{\Delta W_B : B \in \mathcal{B}_m\}$, called white noise, which is a centered Gaussian process with covariance function given by

$$\forall B, B' \in \mathcal{B}_m, \quad \text{Cov}(\Delta W_B, \Delta W_{B'}) = \mathbf{m}(A \cap A') \quad (2.2.6)$$

which cannot be extended to \mathcal{B} in any meaningful way. One may remark that $\mathcal{B}_{loc} \neq \mathcal{B}_m$, but the knowledge of the stochastic measure on \mathcal{B}_{loc} is enough to build the resulting integral. In general,

when it exists, the stochastic measure can be extended to \mathcal{B}_{m_X} but identifying the outer measure m_X may prove to be a challenge.

In turn, the white noise extends to the isonormal process $\mathfrak{W} = \{\mathfrak{W}(f) : f \in L^2(\mathbf{m})\}$ which is a centered Gaussian process with covariance function given by

$$\forall f, g \in L^2(\mathbf{m}), \quad \text{Cov}(\mathfrak{W}(f), \mathfrak{W}(g)) = \int_{\mathcal{T}} f g \, d\mathbf{m}. \quad (2.2.7)$$

For more details about those processes, especially in the multiparameter setting, we refer to [4, Section 1.4.3] where the siBm is called Gaussian \mathbf{m} -noise.

EXAMPLE 2.2.13 (Independently scattered random measure). Reformulating Example 2.2.12, we could say that for a set-indexed Brownian motion W with control measure \mathbf{m} , the linear process \mathfrak{W} exists and is such that $L(W) = L^2(\mathbf{m})$. More generally, when X has independent increments, [83, 99] proved that the linear process \mathfrak{X} always exists and is such that $L(X) = L^\phi$ where ϕ depends on the Lévy-Khintchine triplet of X . We refer to [62, Theorem 8.3.1] for a more precise statement and its proof.

Terminology-wise, we know that ‘Independently Scattered Random Measure’ (ISRM) is the one that stuck. However, we will refrain from using it for two reasons: first, it clashes with what ‘random measure’ means for some authors (see Definition 2.2.14 below) and second, the study we lead in Section 2.3 on processes with independent increments will require a richer vocabulary anyway (see Definition 2.3.1).

In Section 1.3.1, we denoted by $\mathcal{M}(\mathcal{T})$ the set of all Borel measures μ on $(\mathcal{T}, \mathcal{B})$ such that for all $A \in \mathcal{A}$, $\mu(A) < \infty$. It is endowed with the smallest σ -algebra which makes the maps $\mathcal{M}(\mathcal{T}) \ni \mu \mapsto \mu(B) \in \mathbf{R}_+ \cup \{+\infty\}$ measurable for all $B \in \mathcal{B}$.

DEFINITION 2.2.14 (Random measure). A random measure on \mathcal{T} is a random variable $\mathbf{M} : \Omega \rightarrow \mathcal{M}(\mathcal{T})$.

In particular, a random measure is almost surely σ -additive, which is a much stronger condition than being a stochastic measure.

Conversely, a non-negative stochastic measure is not always a random measure. Indeed, suppose that $\mathcal{T} = \mathbf{R}_+$ is endowed with its usual indexing collection from Example 1.2.7 and \mathbf{M} is a random measure on \mathbf{R}_+ . In particular, we know that the \mathbf{R}_+ -indexed process $\{\mathbf{M}_{[0,t]} : t \in \mathbf{R}_+\}$ has almost sure finite variation. However, it is known that there are \mathbf{R}_+ -indexed Lévy processes that extends to non-negative stochastic measures [62, Theorem 8.3.1] while having almost sure infinite variation [10, Theorem 2.4.25].

In our setting, Lévy processes also play a central role since they have both independent and stationary increments, which are respectively the focus of the coming Sections 2.3 and 2.5.

2.3 Processes with independent increments

2.3.1 Equivalent definitions

When one thinks about a set-indexed process as a kind of random measure, most authors suppose that it must have independent increments at the very least (just like in Example 2.2.13). Even though we will study more general cases, it remains a very rich and interesting one nonetheless. The following definition of \mathcal{C} -independent increments may be found in [47, Section 3.4] under the denomination *independent increments*. Proposition 2.3.2 will argue in favor of that less precise terminology.

DEFINITION 2.3.1 (Independent increments). Let \mathcal{D} be \mathcal{C} or $\mathcal{C}(u)$. The process X has independent \mathcal{D} -increments if for all $k \in \mathbf{N}^*$ and pairwise disjoint $D_1, \dots, D_k \in \mathcal{D}$, the random variables $\Delta X_{D_1}, \dots, \Delta X_{D_k}$ are independent.

The process X has independent $\mathcal{B}(X)$ -increments if the stochastic measure ΔX exists and for all $k \in \mathbf{N}^*$ and pairwise disjoint $B_1, \dots, B_k \in \mathcal{B}(X)$, the random variables $\Delta X_{B_1}, \dots, \Delta X_{B_k}$ are independent.

The process \mathcal{X} has independent \mathcal{E} -increments if for all $k \in \mathbf{N}^*$ and simple functions $f_1, \dots, f_k \in \mathcal{E}$ with pairwise disjoint supports, the random variables $\mathcal{X}(f_1), \dots, \mathcal{X}(f_k)$ are independent.

The process \mathcal{X} has independent $L(X)$ -increments if the linear process \mathcal{X} exists and for all $k \in \mathbf{N}^*$ and functions $f_1, \dots, f_k \in L(X)$ with pairwise disjoint supports, the random variables $\mathcal{X}(f_1), \dots, \mathcal{X}(f_k)$ are independent.

PROPOSITION 2.3.2. Consider the following statements:

- (i) X has independent \mathcal{C} -increments,
- (ii) X has independent $\mathcal{C}(u)$ -increments,
- (iii) X has independent $\mathcal{B}(X)$ -increments,
- (iv) \mathcal{X} has independent \mathcal{E} -increments,
- (v) \mathcal{X} has independent $L(X)$ -increments.

Then (i) \Leftrightarrow (ii) \Leftrightarrow (iv) and if the stochastic measure ΔX exists, all of them are equivalent.

Proof. Since $\mathcal{C} \subseteq \mathcal{C}(u)$ and $\{\mathbb{1}_U : U \in \mathcal{C}(u)\} \subseteq \mathcal{E}$ due to Remark 1.2.41, we have (iv) \Rightarrow (ii) \Rightarrow (i). According to Proposition 1.2.42, any simple function may be written as a sum of indicators of pairwise disjoint elements of \mathcal{C} , thus (i) \Rightarrow (iv). Hence (i) \Leftrightarrow (ii) \Leftrightarrow (iv).

Now let us suppose that ΔX exists. The implication (iii) \Rightarrow (ii) is trivial, let us suppose (ii) and prove (iii). Let $B_1, \dots, B_k \in \mathcal{B}(X)$ pairwise disjoint. Let $j \in \llbracket 1, k \rrbracket$. Since $\mathbb{1}_{B_j} \in L(X)$, there exists a sequence $(f_{jl})_{l \in \mathbf{N}}$ of simple functions converging \mathbf{m}_X -a.e. to $\mathbb{1}_{B_j}$ as $l \rightarrow \infty$. For all $l \in \mathbf{N}$, instead of considering $\mathbb{1}_{\{f_{jl} \geq 1/2\}}$, we might as well suppose that f_{jl} only takes values in $\{0, 1\}$ and so may be written as $f_{jl} = \mathbb{1}_{U_{jl}}$ where $U_{jl} \in \mathcal{C}(u)$. Hence for all $j \in \llbracket 1, k \rrbracket$, we have for \mathbf{m}_X -a.e. $t \in \mathcal{T}$,

$$\mathbb{1}_{U_{jl} \setminus \bigcup_{i \neq j} U_{il}}(t) \longrightarrow \mathbb{1}_{B_j \setminus \bigcup_{i \neq j} B_i}(t) = \mathbb{1}_{B_j}(t) \quad \text{as } l \rightarrow \infty.$$

Thus a dominated convergence theorem [62, Proposition 7.1.1] yields

$$\left(\Delta X_{U_{1l} \setminus \bigcup_{i \neq 1} U_{il}}, \dots, \Delta X_{U_{kl} \setminus \bigcup_{i \neq k} U_{il}} \right) \xrightarrow{\mathbf{P}} \left(\Delta X_{B_1}, \dots, \Delta X_{B_k} \right) \quad \text{as } l \rightarrow \infty. \quad (2.3.1)$$

According to (ii), the left-hand side of (2.3.1) is made of independent random variables, thus so is the right-hand side. Hence (ii) \Leftrightarrow (iii).

Let us now prove that (v) is equivalent to the other properties. Since ΔX exists, so does \mathcal{X} by Theorem 2.2.10. The implication (v) \Rightarrow (iv) is obvious. Suppose (iii) and let us prove (v). Remark that proving (v) is equivalent to showing that for all $f \in L_{\text{loc}}(X)$ and pairwise disjoint $B_1, \dots, B_k \in \mathcal{B}(X)$, the random variables $\mathcal{X}(f \mathbb{1}_{B_1}), \dots, \mathcal{X}(f \mathbb{1}_{B_k})$ are independent. This formulation is reachable with the same approximation procedure that we used for (ii) \Rightarrow (iii). The equivalence follows. \square

2.3.2 Infinitely divisible processes

Surprisingly enough, having independent increments is quite a restriction and implies a much more precise structure on the distribution of X . In particular, it has to do with being *infinitely divisible*. We shall review here the few general facts about infinitely divisibility that we need afterwards, most of which can be found in [10, 87], where the reader will also find broader and excellent expositions of the subject as a whole.

The theory of infinitely divisible distributions dates back to the pioneering work of Lévy in the 30s and opened a now gigantic field of research with numerous ramifications. Among them, starting with the works of Lee [67] and Maruyama [70], a notion of infinitely divisible process (*ID process*) indexed by a general space T has been developed. This theory mirrors that of Gaussian processes on a general space where the distribution is characterized by a mean and a covariance function. We shall follow on this matter the more modern point of view developed by Rosiński [85].

DEFINITION 2.3.3 (Infinitely divisible distribution, [85, Section 2.2]). *A probability measure μ on $(\mathbf{R}^k, \mathcal{B}(\mathbf{R}^k))$ where $k \in \mathbf{N}^*$ is infinitely divisible if for all $n \in \mathbf{N}^*$, there exists a probability measure μ_n on $(\mathbf{R}^k, \mathcal{B}(\mathbf{R}^k))$ such that $\mu = \mu_n^{*n}$ (where $*$ is the convolution of probability measures) or equivalently in terms of Fourier transform, for all $\xi \in \mathbf{R}^k$, $\widehat{\mu}(\xi) = \widehat{\mu}_n(\xi)^n$.*

A random vector $Z = (Z_1, \dots, Z_k)$ is infinitely divisible if its law is.

A probability measure μ on $(\mathbf{R}^T, \mathcal{B}(\mathbf{R})^{\otimes T})$ for some general set T is infinitely divisible if all its finite-dimensional distributions are.

A stochastic process $Z = \{Z_t : t \in T\}$ is an ID process if its distribution on $(\mathbf{R}^T, \mathcal{B}(\mathbf{R})^{\otimes T})$ is infinitely divisible.

DEFINITION 2.3.4 (Truncation function, [85, Section 2.2]). *A cutoff function is a map $\chi : \mathbf{R}_+ \rightarrow \mathbf{R}_+$ such that $\chi(x) = 1 + o(x)$ as $x \rightarrow 0^+$. For $k \in \mathbf{N}^*$, the associated truncation function $\llbracket \cdot \rrbracket : \mathbf{R}^k \rightarrow \mathbf{R}^k$ is given by*

$$\forall x \in \mathbf{R}^k, \quad \llbracket x \rrbracket = (x_1 \chi(|x_1|), \dots, x_k \chi(|x_k|)).$$

Likewise for a general set T , $\llbracket \cdot \rrbracket : \mathbf{R}^T \rightarrow \mathbf{R}^T$ is given by

$$\forall x \in \mathbf{R}^T, \forall t \in T, \quad \llbracket x \rrbracket(t) = x(t) \chi(|x(t)|).$$

In the sequel, χ will stand for such a cutoff function and $\llbracket \cdot \rrbracket$ for the associated truncation functions. The confusion in the notation should not be an issue once one looks at which set does the argument belong.

EXAMPLE 2.3.5. *There have been quite a few different choices of cutoff functions in the literature, each with its own set of advantages. Here, we mainly have the cutoff function $\chi(\cdot) = \mathbb{1}_{|\cdot| \leq 1}$ in mind since it is the one best suited for the Lévy-Itô decomposition. But we chose to keep the general writing since it is at its core a mere convention and some readers might be more accustomed to other cutoff functions such as $(1 \wedge |\cdot|)^{-1}$ or $(1 + (\cdot)^2)^{-1}$.*

DEFINITION 2.3.6 (Lévy measure, [85, Definition 2.1]). *A Lévy measure on \mathbf{R}^k is a measure ν on $(\mathbf{R}^k, \mathcal{B}(\mathbf{R}^k))$ such that*

$$\nu(\{0_{\mathbf{R}^k}\}) = 0, \tag{2.3.2}$$

$$\forall j \in \llbracket 1, k \rrbracket, \quad \int_{\mathbf{R}^k} (1 \wedge x_j^2) \nu(dx) < \infty. \tag{2.3.3}$$

Likewise, for a general set T , a Lévy measure on \mathbf{R}^T is a measure ν on $(\mathbf{R}^T, \mathcal{B}(\mathbf{R})^{\otimes T})$ such that

$$\forall B \in \mathcal{B}(\mathbf{R})^{\otimes T}, \quad \nu(B) = \nu_*(B \setminus \{0_{\mathbf{R}^T}\}), \quad (2.3.4)$$

$$\forall t \in T, \quad \int_{\mathbf{R}^t} (1 \wedge x(t)^2) \nu(dx) < \infty \quad (2.3.5)$$

where ν_* is the inner measure associated with ν .

REMARKS 2.3.7. \diamond Inner measures are a concept dual to outer measures. More details about them may be found in [35, §14]. An equivalent formulation to (2.3.4) which does not rely on the inner measure is

$$\forall B \in \mathcal{B}(\mathbf{R})^{\otimes T}, \quad \nu(B) = \sup_{\substack{I \subseteq T: \\ \#I < \infty}} \nu(B \setminus \pi_I^{-1}(\{0_{\mathbf{R}^I}\}))$$

where $\pi_I : \mathbf{R}^T \rightarrow \mathbf{R}^I$ is the canonical projection. When $T \simeq \llbracket 1, k \rrbracket$ is finite, (2.3.2) and (2.3.4) are easily shown to be equivalent, but a problem arises for more general T where one might have $\{0_{\mathbf{R}^T}\} \notin \mathcal{B}(\mathbf{R})^{\otimes T}$, hence the more involved condition.

\diamond Another problem of having a possibly uncountable set T is that a Lévy measure ν might not be necessarily σ -finite. However, using [85, Theorem 2.8], ν is still characterized by its finite-dimensional distributions.

THEOREM 2.3.8 (Lévy-Khintchine representation, [87, Theorem 8.1]). A probability measure μ on $(\mathbf{R}^k, \mathcal{B}(\mathbf{R}^k))$ where $k \in \mathbf{N}^*$ is infinitely divisible if and only if there exists a map $\psi : \mathbf{R}^k \rightarrow \mathbf{C}$ such that $\hat{\mu} = e^\psi$ and

$$\forall \xi \in \mathbf{R}^k, \quad \psi(\xi) = i\langle \xi, b \rangle - \frac{1}{2}\langle \xi, \Sigma \xi \rangle + \int_{\mathbf{R}^k} (e^{i\langle \xi, x \rangle} - 1 - i\langle \xi, \llbracket x \rrbracket \rangle) \nu(dx) \quad (2.3.6)$$

where $b \in \mathbf{R}^k$ is a vector, Σ a symmetric non-negative definite $k \times k$ matrix and ν a Lévy measure on \mathbf{R}^k . The map ψ is called the Lévy-Khintchine exponent, b the drift, Σ the covariance matrix and ν the Lévy measure of μ . The triplet (b, Σ, ν) is unique up to the choice of the cutoff function and is called the Lévy-Khintchine triplet of μ .

REMARK 2.3.9. The formula (2.3.6) might appear to be quite abstruse at first, but tells a story somewhat simpler than it looks. First, remark that the Lévy-Khintchine exponent $\psi = \psi_b + \psi_\Sigma + \psi_\nu$ is the sum of three simpler exponents each only depending on one element of the triplet. In other words, the distribution μ may be represented as the sum of three independent random vectors:

1. A random vector with Lévy-Khintchine exponent $\psi_b(\xi) = i\langle \xi, b \rangle$ which is the constant vector equal to b .
2. A random vector with Lévy-Khintchine exponent $\psi_\Sigma(\xi) = -\frac{1}{2}\langle \xi, \Sigma \xi \rangle$ which is a centered Gaussian vector of covariance Σ . It is called the Gaussian part of μ .
3. A random vector with Lévy-Khintchine exponent $\psi_\nu(\xi) = \int_{\mathbf{R}^k} (e^{i\langle \xi, x \rangle} - 1 - i\langle \xi, \llbracket x \rrbracket \rangle) \nu(dx)$ which is a compensated compound Poisson process of intensity measure ν . It is called the Poissonian part of μ . This is by far the most intricate part of an infinitely divisible distribution. For an excellent exposition of Poisson processes, we refer to [58].

This decomposition will be useful to us when tackling the regularity of some specific ID processes (Section 3.5.1).

The Lévy-Khintchine representation for a random vector is actually generalizable to processes as well, provided that one first extends the definition of the Fourier transform to accomodate more general Lévy-Khintchine triplets.

DEFINITION 2.3.10 (Fourier transform). *Let T be a general set and μ a probability measure on $(\mathbf{R}^T, \mathcal{B}(\mathbf{R})^{\otimes T})$. Denote by $\mathbf{R}^{(T)} = \{\xi \in \mathbf{R}^T : \text{supp}(\xi) \text{ is finite}\}$ and*

$$\forall (\xi, \mathbf{z}) \in \mathbf{R}^{(T)} \times \mathbf{R}^T, \quad \langle \xi, \mathbf{z} \rangle = \sum_{t \in \text{supp}(\xi)} \xi(t) z(t).$$

The Fourier transform of μ is the function $\widehat{\mu} : \mathbf{R}^{(T)} \rightarrow \mathbf{C}$ given by

$$\forall \xi \in \mathbf{R}^{(T)}, \quad \widehat{\mu}(\xi) = \int_{\mathbf{R}^T} e^{i\langle \xi, \mathbf{z} \rangle} \mu(d\mathbf{z}).$$

Likewise, the Fourier transform of a T -indexed stochastic process $Z = \{Z_t : t \in T\}$ is the Fourier transform of its distribution \mathbf{P}_Z on $(\mathbf{R}^T, \mathcal{B}(\mathbf{R})^{\otimes T})$, i.e.

$$\forall \xi \in \mathbf{R}^{(T)}, \quad \widehat{\mathbf{P}}_Z(\xi) = \mathbf{E}[e^{i\langle \xi, Z \rangle}].$$

THEOREM 2.3.11 (Lévy-Khintchine representation of ID processes, [85, Corollary 2.9]). *A probability measure μ on $(\mathbf{R}^T, \mathcal{B}(\mathbf{R})^{\otimes T})$ for some general set T is infinitely divisible if and only if there exists a map $\psi : \mathbf{R}^{(T)} \rightarrow \mathbf{C}$ such that $\widehat{\mu} = e^\psi$ and*

$$\forall \xi \in \mathbf{R}^{(T)}, \quad \psi(\xi) = i\langle \xi, \mathbf{b} \rangle - \frac{1}{2}\langle \xi, \Sigma \xi \rangle + \int_{\mathbf{R}^T} (e^{i\langle \xi, x \rangle} - 1 - i\langle \xi, \llbracket x \rrbracket \rangle) \nu(dx). \quad (2.3.7)$$

where $\mathbf{b} \in \mathbf{R}^T$, $\Sigma \in \mathbf{R}^{T \times T}$ is symmetric non-negative definite and ν is a Lévy measure on \mathbf{R}^T . The same terminology as in Theorem 2.3.8 is used for the triplet $(\mathbf{b}, \Sigma, \nu)$, which is also unique.

In particular, for any finite subset $T' \subseteq T$, the finite-dimensional $\mu_{T'}$ induced by μ on $\mathbf{R}^{T'}$ has Lévy-Khintchine triplet $(\mathbf{b}|_{T'}, \Sigma|_{T' \times T'}, \nu_{T'})$.

We now use all this knowledge to compute particular cases of Lévy-Khintchine triplets of interest to us. Those results should already be well-known, but we have not been able to find them in the literature.

LEMMA 2.3.12. (ID process with independent entries) *An ID process $Z = \{Z_t : t \in T\}$ with Lévy-Khintchine triplet $(\mathbf{b}, \Sigma, \nu)$ forms an independent family of random variables if and only if for all $s \neq t$ in T , $\Sigma(s, t) = 0$ and for all finite subset $T' \subseteq T$, $\text{supp}(\nu_{T'}) \subseteq \bigcup_{t \in T'} \mathbf{R}\mathbf{1}_{\{t\}}$ where $\mathbf{R}\mathbf{1}_{\{t\}} = \{a\mathbf{1}_{\{t\}} : a \in \mathbf{R}\}$.*

Remark that for a finite subset $T' \subseteq T$, the family $(\mathbf{1}_{\{t\}})_{t \in T'}$ is nothing but the canonical basis of $\mathbf{R}^{T'}$. In particular, $\bigcup_{t \in T'} \mathbf{R}\mathbf{1}_{\{t\}}$ is the union of all the coordinate axes of $\mathbf{R}^{T'}$.

Proof. The drift corresponding to a constant and the Gaussian case being all too well-known, we suppose that $\mathbf{b} = 0$ and $\Sigma = 0$. Suppose that Z is an independent family and consider a finite

subset $T' \subseteq T$. Then, for all $\xi \in \mathbf{R}^{T'}$,

$$\begin{aligned} \psi_Z(\xi) &= \sum_{t \in T'} \psi_{Z_t}(\xi(t)) \\ &= \sum_{t \in T'} \int_{\mathbf{R}^{(t)}} (e^{i\xi(t)x(t)} - 1 - i\xi(t)\llbracket x(t) \rrbracket) \nu_{\{t\}}(dx) \\ &= \sum_{t \in T'} \int_{\mathbf{R}^{T'}} (e^{i\xi(t)x(t)} - 1 - i\xi(t)\llbracket x(t) \rrbracket) \nu_{T'}(dx) \\ &= \int_{\mathbf{R}^{T'}} (e^{i\langle \xi, x \rangle} - 1 - i\langle \xi, \llbracket x \rrbracket \rangle) \nu_{T'}(E \cap \cdot)(dx) \end{aligned}$$

where $E = \bigcup_{t \in T'} \mathbf{R}\mathbb{1}_{\{t\}}$. We used Theorem 2.3.11 for the first and third equalities. The hypothesis of independence was used for the second. The last one is just a computation. By unicity of the Lévy measure, we have $\nu_{T'} = \nu_{T'}(E \cap \cdot)$ and thus $\text{supp}(\nu_{T'}) \subseteq E$. Conversely, if $\text{supp}(\nu_{T'}) \subseteq E$, one may do the same computation on ψ_Z as above, but in reverse order and actually show that for all $\xi \in \mathbf{R}^{T'}$, $\psi_Z(\xi) = \sum_{t \in T'} \psi_{Z_t}(\xi(t))$ for all finite $T' \subseteq T$. Hence Z is an independent family. \square

Denote by $\mathcal{M}(\mathbf{R})$ the set of (non-negative) Borel measures on \mathbf{R} and by $\varphi_*\mu$ the pushforward of the measure μ by the map φ (whenever they are compatible).

THEOREM 2.3.13 (Characterization of set-indexed processes with independent increments). *The set-indexed process X has independent increments if and only if there exists $b : \mathcal{A} \rightarrow \mathbf{R}$, $\sigma^2 : \mathcal{A} \rightarrow \mathbf{R}_+$ and a kernel $\nu : \mathcal{A} \rightarrow \mathcal{M}(\mathbf{R})$ such that the following holds:*

1. For all $C \in \mathcal{C}$, $\Delta\sigma^2(C) \geq 0$.
2. For all $A \in \mathcal{A}$, ν_A is a Lévy measure on \mathbf{R} and for all $C \in \mathcal{C}$ and $B \in \mathcal{B}(\mathbf{R})$ such that $0 \notin \bar{B}$, $\Delta\nu_C(B) \geq 0$.
3. X is an ID process whose Lévy-Khintchine triplet $(\mathbf{b}, \Sigma, \nu)$ may be expressed for all $A, A' \in \mathcal{A}$ and finite subcollection \mathcal{A}' of \mathcal{A} closed under intersections by

$$\begin{aligned} \mathbf{b}(A) &= b(A), \\ \Sigma(A, A') &= \sigma^2(A \cap A'), \\ \nu_{\mathcal{A}'} &= \varphi_*^{\mathcal{A}'} \left(\sum_{C \in \mathcal{C}^\ell(\mathcal{A}')} (\iota_C)_* \Delta\nu_C \right) \end{aligned} \tag{2.3.8}$$

where $\nu_{\mathcal{A}'}$ is the finite-dimensional Lévy measure induced by ν on $\mathbf{R}^{\mathcal{A}'}$, $\mathcal{C}^\ell(\mathcal{A}')$ has been given in Definition 1.2.38, $\iota_C : \mathbf{R} \hookrightarrow \mathbf{R}^{\mathcal{C}^\ell(\mathcal{A}')}$ is the injection given for all $x \in \mathbf{R}$ by $\iota_C(x) = x\mathbb{1}_{\{C\}}$ and $\varphi^{\mathcal{A}'} : \mathbf{R}^{\mathcal{C}^\ell(\mathcal{A}')} \rightarrow \mathbf{R}^{\mathcal{A}'}$ is the linear map given by

$$\forall x \in \mathbf{R}^{\mathcal{C}^\ell(\mathcal{A}')} , \quad \varphi^{\mathcal{A}'}(x) : A \mapsto \sum_{\substack{C \in \mathcal{C}^\ell(\mathcal{A}'): \\ C \subseteq A}} x(C). \tag{2.3.9}$$

Conversely, a triplet (b, σ^2, ν) which verifies conditions 1 and 2 above uniquely determines through (2.3.8) the Lévy-Khintchine triplet of a set-indexed process with independent increments.

REMARKS 2.3.14.

- ◇ Due to the assured existence of a finitely additive extension (Proposition 1.2.44), Δb (resp. $\Delta\sigma^2$) may be seen as a ‘signed pre-measure’ (resp. ‘pre-measure’).
- ◇ While checking the condition $\Delta\nu_C(B) \geq 0$, the reason why one should restrict their attention to Borel sets $B \in \mathcal{B}(\mathbf{R})$ such that $0 \notin \bar{B}$ is to avoid subtracting infinities together. Indeed, for $A \in \mathcal{A}$, being a Lévy measure for ν_A implies that for such B , $\nu_A(B) < \infty$ but this might not be the case anymore if $0 \in \bar{B}$.
Actually, this fear is unfounded since under that hypothesis, $\Delta\nu_C$ defines a σ -finite pre-measure on the ring $\{B \in \mathcal{B}(\mathbf{R}^*) : 0 \notin \bar{B}\}$, and hence may be uniquely extended to a measure on $\mathcal{B}(\mathbf{R}^*)$. Setting $\Delta\nu_C(\{0\}) = 0$ makes it a (Lévy) measure on $\mathcal{B}(\mathbf{R})$.
- ◇ The morphism $\varphi^{\mathcal{A}'}$ given in (2.3.9) might seem a bit mysterious but just comes from the fact that for all $A \in \mathcal{A}'$, we have the \mathcal{C} -representation $\mathbb{1}_A = \sum_{C \in \mathcal{C}^\ell(\mathcal{A}'): C \subseteq A} \mathbb{1}_C$ (see Proposition 1.2.42).

Proof of Theorem 2.3.13. Suppose that X has independent increments. Let us prove that X is infinitely divisible. Let ϕ be a continuous $\mathcal{A}(u)$ -flow. The projection ΔX^ϕ along ϕ is then a one-dimensional process with independent increments. According to [87, Theorem 9.7], ΔX^ϕ must be an infinitely divisible process. Hence by Theorem 1.3.13, X also is infinitely divisible. Its Lévy-Khintchine triplet $(\mathbf{b}, \Sigma, \nu)$ is thus well-defined by Theorem 2.3.11. In particular, for all $U \in \mathcal{C}(u)$, the random variable ΔX_U is infinitely divisible as well. Denote by $(b(U), \sigma^2(U), \nu_U)$ its Lévy-Khintchine triplet given by Theorem 2.3.8. For all pairwise disjoint elements U, V in $\mathcal{C}(u)$, we have $\Delta X_{U \sqcup V} = \Delta X_U + \Delta X_V$ where ΔX_U and ΔX_V are independent. By unicity of the Lévy-Khintchine triplet, $\sigma^2(U \sqcup V) = \sigma^2(U) + \sigma^2(V)$. By unicity of the increment map (Proposition 1.2.44), $\sigma^2 = \Delta(\sigma^2|_{\mathcal{A}'})$. Hence for all $C \in \mathcal{C}$, $\Delta\sigma^2(C) = \sigma^2(C) \geq 0$, which proves condition 1 for σ^2 in the statement of Theorem 2.3.13. Condition 2 for ν is proven exactly in the same fashion.

Let us show that (2.3.8) holds. Since each element of the triplet corresponds to an independent part of the distribution (Remark 2.3.9), we may cancel two out of the three terms when computing the third one for our purpose. So suppose that $\mathbf{b} = 0$ and $\nu = 0$ and let us establish (2.3.8) for Σ and σ^2 . In this case, X is a centered Gaussian process with covariance Σ . Thus, for all $A, A' \in \mathcal{A}$,

$$\begin{aligned} \Sigma(A, A') &= \text{Cov}(X_A, X_{A'}) \\ &= \text{Cov}(X_{A \cap A'} + \Delta X_{A \setminus A'}, X_{A' \cap A} + \Delta X_{A' \setminus A}) \\ &= \text{Var}(X_{A \cap A'}) && \text{since } X \text{ has independent increments,} \\ &= \sigma^2(A \cap A'). \end{aligned}$$

Now, let us suppose that $\mathbf{b} = 0$ and $\Sigma = 0$ and let us establish (2.3.8) for ν and ν . Let \mathcal{A}' be a finite subcollection of \mathcal{A} . Then, by definition of $\varphi^{\mathcal{A}'}$,

$$(X_A)_{A \in \mathcal{A}'} = \varphi^{\mathcal{A}'}(\Delta X|_{\mathcal{C}^\ell(\mathcal{A}')}). \quad (2.3.10)$$

By definition of $\nu_{\mathcal{A}'}$, the left-hand side of (2.3.10) is an infinitely divisible random vector with triplet $(0, 0, \nu_{\mathcal{A}'})$. As for the right-hand side, when we established condition 2, a by-product was that for all $C \in \mathcal{C}$, ΔX_C has triplet $(0, 0, \Delta\nu_C)$. Moreover, since all the elements of $\mathcal{C}^\ell(\mathcal{A}')$ are pairwise disjoint, $\Delta X|_{\mathcal{C}^\ell(\mathcal{A}'')}$ is a random vector with independent coordinates. According to Lemma 2.3.12, its Lévy-Khintchine triplet is thus $(0, 0, \sum_{C \in \mathcal{C}^\ell(\mathcal{A}')} (\iota_C)_* \Delta\nu_C)$. Hence, by (2.3.10) and the last point in Remark 2.3.14,

$$\nu_{\mathcal{A}'} = \varphi_*^{\mathcal{A}'} \left(\sum_{C \in \mathcal{C}^\ell(\mathcal{A}')} (\iota_C)_* \Delta\nu_C \right). \quad (2.3.11)$$

Conversely, let us consider a triplet (b, σ^2, ν) that checks conditions 1 and 2. Define then (b, Σ, ν) by (2.3.8) and let us show that it is the Lévy-Khintchine triplet of some set-indexed process with independent increments. There is nothing to prove for b . The map Σ is symmetric and for all $k \in \mathbf{N}^*$, $\alpha_1, \dots, \alpha_k \in \mathbf{R}$ and $A_1, \dots, A_k \in \mathcal{A}$,

$$\sum_{1 \leq i, j \leq k} \alpha_i \alpha_j \Sigma(A_i, A_j) = \sum_{1 \leq i, j \leq k} \alpha_i \alpha_j \sigma^2(A_i \cap A_j) = \int_{\mathcal{T}} \left(\sum_{j=1}^k \alpha_j \mathbb{1}_{A_j} \right)^2 d\sigma^2 \geq 0$$

where the integral with respect to σ^2 is well-defined at least for simple functions by Proposition 1.2.44. Thus Σ is non-negative definite.

For ν , one needs to check that it is a well-defined Lévy measure on $\mathbf{R}^{\mathcal{A}}$. By [85, Theorem 2.8], it is enough to check that for all $A \in \mathcal{A}$, $\nu_{\{A\}}$ is a Lévy measure and for all finite subcollections $\mathcal{A}' \subseteq \mathcal{A}''$ of \mathcal{A} , $(\pi_{\mathcal{A}'}^{\mathcal{A}''})_* \nu_{\mathcal{A}''} = \nu_{\mathcal{A}'}$ where $\pi_{\mathcal{A}'}^{\mathcal{A}''} : \mathbf{R}^{\mathcal{A}''} \rightarrow \mathbf{R}^{\mathcal{A}'}$ is the canonical projection.

Let $A \in \mathcal{A}$. By definition, $\nu_{\{A\}} = \nu_A$ and thus is a Lévy measure. Let $\mathcal{A}' \subseteq \mathcal{A}''$ be two finite subcollections of \mathcal{A} . Define the projection

$$p_{\mathcal{A}'}^{\mathcal{A}''} : \mathbf{R}^{\mathcal{C}^{\ell}(\mathcal{A}'')} \longrightarrow \mathbf{R}^{\mathcal{C}^{\ell}(\mathcal{A}')}, \\ x \longmapsto \left[C' \mapsto \sum_{\substack{C'' \in \mathcal{C}^{\ell}(\mathcal{A}'') \\ C'' \subseteq C'}} x(C'') \right]. \quad (2.3.12)$$

Then, since

$$\forall C' \in \mathcal{C}^{\ell}(\mathcal{A}'), \quad \Delta \nu_{C'} = \sum_{\substack{C'' \in \mathcal{C}^{\ell}(\mathcal{A}'') \\ C'' \subseteq C'}} \Delta \nu_{C''},$$

we have

$$(p_{\mathcal{A}'}^{\mathcal{A}''})_* \left(\sum_{C'' \in \mathcal{C}^{\ell}(\mathcal{A}'')} (\iota_{C''})_* \Delta \nu_{C''} \right) = \left(\sum_{C' \in \mathcal{C}^{\ell}(\mathcal{A}')} (\iota_{C'})_* \Delta \nu_{C'} \right). \quad (2.3.13)$$

Moreover, according to (2.3.9), (2.3.12) and the definition of $\pi_{\mathcal{A}'}^{\mathcal{A}''}$, the following diagram commutes:

$$\begin{array}{ccc} \mathbf{R}^{\mathcal{C}^{\ell}(\mathcal{A}'')} & \xrightarrow{\varphi_{\mathcal{A}''}} & \mathbf{R}^{\mathcal{A}''} \\ p_{\mathcal{A}'}^{\mathcal{A}''} \downarrow & & \downarrow \pi_{\mathcal{A}'}^{\mathcal{A}''} \\ \mathbf{R}^{\mathcal{C}^{\ell}(\mathcal{A}')} & \xrightarrow{\varphi_{\mathcal{A}'}} & \mathbf{R}^{\mathcal{A}'} \end{array}$$

Hence, combining (2.3.11) and (2.3.13), we get $(\pi_{\mathcal{A}'}^{\mathcal{A}''})_* \nu_{\mathcal{A}''} = \nu_{\mathcal{A}'}$. It follows that (b, Σ, ν) is a well-defined Lévy-Khintchine triplet. Let us consider an ID process $X = \{X_A : A \in \mathcal{A}\}$ with such a triplet. What is left to show is that X has independent increments. But back-tracking all the previous computations readily shows that for all $k \in \mathbf{N}^*$ and pairwise disjoint $C_1, \dots, C_k \in \mathcal{C}$, the random vector $(\Delta X_{C_1}, \dots, \Delta X_{C_k})$ is infinitely divisible with Lévy-Khintchine triplet

$$\left(\begin{pmatrix} \Delta b(C_1) \\ \vdots \\ \Delta b(C_k) \end{pmatrix}, \begin{pmatrix} \Delta \sigma^2(C_1) & & (0) \\ & \ddots & \\ (0) & & \Delta \sigma^2(C_k) \end{pmatrix}, \sum_{j=1}^k (\iota_{C_j})_* \Delta \nu_{C_j} \right)$$

from which the independence is easily deduced by Lemma 2.3.12. \square

From this theorem, one may deduce a necessary and sufficient condition for a set-indexed process with independent increments to extend to a stochastic measure, improving a well-known result from [83]. Beforehand, we prove a small lemma that eases the proof of Corollary 2.3.16.

LEMMA 2.3.15. Consider random variables Y_0, Y_1, \dots defined on $(\Omega, \mathcal{F}, \mathbf{P})$ and a random variables Z, Z_0, Z_1, \dots defined on another probability space $(\Omega', \mathcal{F}', \mathbf{P}')$. If $(Y_k)_{k \in \mathbf{N}} \stackrel{\text{fdd}}{=} (Z_k)_{k \in \mathbf{N}}$ and $Z_k \xrightarrow{\mathbf{P}'} Z$ as $k \rightarrow \infty$, then there exists $Y \in L^0(\Omega)$ such that $Y_k \xrightarrow{\mathbf{P}} Y$ as $k \rightarrow \infty$.

Proof. Suppose that $(Z_k)_{k \in \mathbf{N}}$ converges in probability. In other words, it is a Cauchy sequence in the complete metric space $L^0(\Omega')$ (Example 2.2.3), meaning that

$$\forall \varepsilon > 0, \exists N \in \mathbf{N} : \forall p, q \geq N, \mathbf{E}'[|Z_p - Z_q| \wedge 1] \leq \varepsilon \quad (2.3.14)$$

where \mathbf{E}' is the expectancy with respect to \mathbf{P}' . If $(Y_k)_{k \in \mathbf{N}} \stackrel{\text{fdd}}{=} (Z_k)_{k \in \mathbf{N}}$, then (2.3.14) implies that $(Y_k)_{k \in \mathbf{N}}$ is a Cauchy sequence in $L^0(\Omega)$. The result follows. \square

COROLLARY 2.3.16 (Extension of a process with independent increments). Suppose that X has independent increments and consider the triplet (b, σ^2, ν) given in Theorem 2.3.13. Then the following statements are equivalent:

- (i) X extends to a stochastic measure ΔX on $\mathcal{B}(X)$,
- (ii) Δb (resp. $\Delta \sigma^2, \Delta \nu(\cdot)$) extends to a signed measure $\Delta \mathbf{b}$ on \mathcal{B} (resp. measure $\Delta \sigma^2$ on \mathcal{B} , measure $\tilde{\nu}$ on $\mathcal{B} \otimes \mathcal{B}(\mathbf{R})$).

We call $(\Delta \mathbf{b}, \Delta \sigma^2, \tilde{\nu})$ the modified triplet of X .

If one only looks for the existence of ΔX ‘in law’, i.e. the extension of the distribution of X on $\mathbf{R}^{\mathcal{A}}$ to a distribution on $\mathbf{R}^{\mathcal{B}(X)}$ that verifies the properties stated in Corollary 2.3.16, then this has essentially been proven by Rajput and Rosiński in [83]. However, it is another story to build ΔX on the same probability space as X and such that $\Delta X|_{\mathcal{A}}$ is actually a version of X .

Proof of Corollary 2.3.16. If X extends to a stochastic measure ΔX , then the fact that the triplet $(\Delta b, \Delta \sigma^2, \Delta \nu)$ extends as well is a consequence of [83, Proposition 2.1(a) and Lemma 2.3].

Conversely, suppose that $(\Delta b, \Delta \sigma^2, \Delta \nu)$ extends to $(\Delta \mathbf{b}, \Delta \sigma^2, \tilde{\nu})$ as intended. Then, by [83, Proposition 2.1(b)], there exists a stochastic process $M = \{M_B : B \in \mathcal{B}(X)\}$ defined on a possibly different probability space $(\Omega', \mathcal{F}', \mathbf{P}')$ such that

- ◊ M is a stochastic measure,
- ◊ M is an ID process with Lévy-Khintchine triplet $(\Delta \mathbf{b}, \Sigma, \nu)$ where Σ and ν are given by formulas very much like (2.3.8) where \mathcal{A} is replaced by $\mathcal{B}(X)$ and $\Delta \nu_c$ by $\tilde{\nu}(B \times \cdot)$.

In particular, we also have $M|_{\mathcal{C}(u)} \stackrel{\text{fdd}}{=} \Delta X$.

Let us build an extension of X using M . For any $B \in \mathcal{B}(X)$, there exists a sequence $(U_k)_{k \in \mathbf{N}}$ in $\mathcal{C}(u)$ such that $M_{U_k} \xrightarrow{\mathbf{P}'} M_B$ as $k \rightarrow \infty$. By Lemma 2.3.15, there exists $Y_B \in L^0(\Omega)$ such that

$\Delta X_{U_k} \xrightarrow{\mathbf{P}} Y_B$ as $k \rightarrow \infty$. Remark that Y_B does not depend on the choice of the sequence $(U_k)_{k \in \mathbf{N}}$ since if $(V_k)_{k \in \mathbf{N}}$ is another one such that $\mathbf{M}_{V_k} \xrightarrow{\mathbf{P}'} \mathbf{M}_B$ and $\Delta X_{V_k} \xrightarrow{\mathbf{P}} Y'_B$ as $k \rightarrow \infty$, then we have

$$\begin{array}{ccc} \mathbf{M}_{U_k} - \mathbf{M}_{V_k} & \xrightarrow[k \rightarrow \infty]{\mathbf{P}'} & 0 \\ \text{law } \parallel & & \parallel \text{ law} \\ \Delta X_{U_k} - \Delta X_{V_k} & \xrightarrow[k \rightarrow \infty]{\mathbf{P}} & Y_B - Y'_B. \end{array}$$

Hence $Y_B = Y'_B$ a.s. In particular, since the definition of Y_B does not depend on the approximating sequence of B , we indeed have $Y|_{\mathcal{A}} = X$.

What is left to show is that $Y = \{Y_B : B \in \mathcal{B}(X)\}$ is a stochastic measure, which would imply by Theorem 2.2.8 that ΔX exists and $Y = \Delta X$. We know that $Y \stackrel{\text{fdd}}{=} \mathbf{M}$. Hence $Y_\emptyset = 0$ a.s. Moreover, if $(B_k)_{k \in \mathbf{N}}$ is a non-decreasing sequence of sets in $\mathcal{B}(X)$ such that $B = \bigcup_{k \in \mathbf{N}} B_k$ belongs to $\mathcal{B}(X)$, then $\mathbf{M}_{B_k} \xrightarrow{\mathbf{P}'} \mathbf{M}_B$ as $k \rightarrow \infty$. By Lemma 2.3.12, $(Y_{B_k})_{k \in \mathbf{N}}$ converges in probability as well. We just need to show that the limit is equal to Y_B . Consider a sequence $(U_k)_{k \in \mathbf{N}}$ in $\mathcal{C}(u)$ such that for all $k \in \mathbf{N}$, $\mathbf{E}'[|\mathbf{M}_{U_k} - \mathbf{M}_{B_k}| \wedge 1] \leq 2^{-k}$. Applying Lemma 2.3.12 once more yields that $Y_{U_k} - Y_{B_k} \xrightarrow{\mathbf{P}} 0$ as $k \rightarrow \infty$. Since $Y_{U_k} \xrightarrow{\mathbf{P}} Y_B$ as $k \rightarrow \infty$ by definition of Y_B , we have $Y_{B_k} \xrightarrow{\mathbf{P}} Y_B$ as $k \rightarrow \infty$. Since Y is also finitely additive, Y is a stochastic measure. The result follows. \square

REMARK 2.3.17. *In the previous proof, one may also apply a general coupling result like [54, Theorem 6.10] to deduce the existence of ΔX from that of \mathbf{M} . However, one needs to be careful and find a Borel subspace of $\mathbf{R}^{\mathcal{B}(X)}$ for which \mathbf{M} has a version with sample paths almost surely residing in it. We kept the proof above since it is more general and we applied the just-described idea in Corollary 2.4.9.*

Once the Lévy-Khintchine triplet of the stochastic measure ΔX has been deduced, it becomes natural to ask the same about the linear process $\mathcal{X} = \{\mathcal{X}(f) : f \in L(X)\}$. Although expressing the Lévy-Khintchine triplet of \mathcal{X} directly from that of X is a bit tedious, using of the modified triplet $(\Delta \mathbf{b}, \Delta \sigma^2, \tilde{\nu})$ eases things.

COROLLARY 2.3.18. *Suppose that X has independent increments and extends to a stochastic measure ΔX . Consider the modified triplet $(\Delta \mathbf{b}, \Delta \sigma^2, \tilde{\nu})$. Then the linear process $\mathcal{X} = \{\mathcal{X}(f) : f \in L(X)\}$ is a well-defined ID process whose Lévy-Khintchine triplet $(\mathbf{b}, \Sigma, \nu)$ may be expressed for all $f, g \in L(X)$, finite subset $\mathbf{f} \subseteq L(X)$ and $B \in \mathcal{B}(\mathbf{R})^{\otimes \mathbf{f}}$ by*

$$\begin{aligned} \mathbf{b}(f) &= \int_{\mathcal{T}} f d(\Delta \mathbf{b}) + \int_{\mathcal{T} \times \mathbf{R}} (\llbracket x f(t) \rrbracket - f(t) \llbracket x \rrbracket) \tilde{\nu}(dt, dx), \\ \Sigma(f, g) &= \int_{\mathcal{T}} f g d(\Delta \sigma^2), \\ \nu_{\mathbf{f}}(B) &= \int_{\mathcal{T} \times \mathbf{R}} \mathbb{1}_B(x, \mathbf{f}(t)) \tilde{\nu}(dt, dx). \end{aligned} \tag{2.3.15}$$

where $\nu_{\mathbf{f}}$ is the Lévy measure induced by ν on $\mathbf{R}^{\mathbf{f}}$ and for all $t \in \mathcal{T}$, $x, \mathbf{f}(t)$ is the element of $\mathbf{R}^{\mathbf{f}}$ that maps $h \in \mathbf{f}$ to $xh(t) \in \mathbf{R}$.

Such a computation when X is a one-dimensional Lévy process has been carried out in [88, Proposition 3.17] for $\chi(\cdot) = \mathbb{1}_{|\cdot| \leq 1}$. The general case is not much harder, but the computation is somewhat different for the Poissonian part due to the absence of stationarity.

Proof. By Theorem 2.2.10, \mathcal{X} is well-defined. Moreover, we know that it takes values in the closure (for the convergence in probability) of $\text{span}(X_A : A \in \mathcal{A})$. Since X is an ID process and ID distributions are closed under weak convergence, we know that \mathcal{X} is an ID process. Hence its Lévy-Khintchine triplet $(\mathbf{b}, \Sigma, \nu)$ is well-defined.

Let us show that (2.3.15) holds. First, remark that each term in (2.3.15) is well-defined due to the characterization of $L(X)$ given by [83, Theorem 3.3] (we shall give a bit more details later in Theorem 2.3.8). By density of \mathcal{E} in $L(X)$ and continuity of \mathcal{X} , we may restrict our attention to simple functions.

Consider $k \in \mathbf{N}^*$ and $\mathbf{f} = (f_1, \dots, f_k)^\top \in \mathcal{E}^k$ and let us compute the Lévy-Khintchine exponent $\psi_{\mathcal{X}(\mathbf{f})}$ of $\mathcal{X}(\mathbf{f}) = (\mathcal{X}(f_1), \dots, \mathcal{X}(f_k))$. Using Proposition 1.2.42, we may write

$$\mathbf{f} = M \begin{pmatrix} \mathbb{1}_{C_1} \\ \vdots \\ \mathbb{1}_{C_l} \end{pmatrix} \quad (2.3.16)$$

where C_1, \dots, C_l are pairwise disjoint elements of \mathcal{C} and M is a $k \times l$ matrix of coefficients. For $\xi \in \mathbf{R}^k$, we have

$$\begin{aligned} \widehat{\mathbf{P}}_{\mathcal{X}(\mathbf{f})}(\xi) &= \mathbf{E} \left[e^{i\langle \xi, \mathcal{X}(\mathbf{f}) \rangle} \right] \\ &= \mathbf{E} \left[e^{i\langle \xi, M(\Delta X_{C_1}, \dots, \Delta X_{C_l}) \rangle} \right] \\ &= \widehat{\mathbf{P}}_{(\Delta X_{C_1}, \dots, \Delta X_{C_l})}(M^\top \xi) \end{aligned}$$

Hence, by independence of $(\Delta X_{C_1}, \dots, \Delta X_{C_l})$,

$$\forall \xi \in \mathbf{R}^k, \quad \psi_{\mathcal{X}(\mathbf{f})}(\xi) = \sum_{j=1}^l \psi_{\Delta X_{C_j}}((M^\top \xi)_j). \quad (2.3.17)$$

In order to make the computation more palatable, we use Remark 2.3.9 and split the drift, Gaussian and Poissonian components apart.

If $(\Delta \mathbf{b}, \Delta \sigma^2, \tilde{\nu}) = (\Delta \mathbf{b}, 0, 0)$, then (2.3.17) reads for all $\xi \in \mathbf{R}^k$,

$$\begin{aligned} \psi_{\mathcal{X}(\mathbf{f})}(\xi) &= \sum_{j=1}^l i(M^\top \xi)_j \Delta \mathbf{b}(C_j) \\ &= i\langle M^\top \xi, \Delta \mathbf{b}(C) \rangle \\ &= i\langle \xi, M \Delta \mathbf{b}(C) \rangle. \end{aligned}$$

Hence, due to (2.3.16), we get

$$\forall \xi \in \mathbf{R}^k, \quad \psi_{\mathcal{X}(\mathbf{f})}(\xi) = i\left\langle \xi, \int_{\mathcal{T}} \mathbf{f} d(\Delta \mathbf{b}) \right\rangle \quad (2.3.18)$$

where the integral of the vector \mathbf{f} is done component-wise.

Suppose that $(\Delta \mathbf{b}, \Delta \sigma^2, \tilde{\nu}) = (0, \Delta \sigma^2, 0)$ and introduce the diagonal matrix $D = (\Delta \sigma^2(C_i \cap C_j))_{1 \leq i, j \leq l}$. Equation (2.3.17) then reads for all $\xi \in \mathbf{R}^k$,

$$\begin{aligned} \psi_{\mathcal{X}(\mathbf{f})}(\xi) &= -\frac{1}{2} \sum_{j=1}^l (M^\top \xi)_j^2 \Delta \sigma^2(C_j) \\ &= -\frac{1}{2} \langle (M^\top \xi), D(M^\top \xi) \rangle \\ &= -\frac{1}{2} \langle \xi, (MDM^\top) \xi \rangle. \end{aligned}$$

Hence, due to (2.3.16), we get

$$\forall \xi \in \mathbf{R}^k, \quad \psi_{\mathcal{X}(f)}(\xi) = -\frac{1}{2} \left\langle \xi, \int_{\mathcal{T}} f f^\top d(\Delta \sigma^2) \xi \right\rangle \quad (2.3.19)$$

where the integral of the matrix $f f^\top$ is done component-wise.

Suppose that $(\Delta \mathbf{b}, \Delta \sigma^2, \tilde{\nu}) = (0, 0, \tilde{\nu})$. Fix a Borel map $\varphi : \mathbf{R} \times \mathbf{R}^k \rightarrow \mathbf{R}_+$ and denote by (e_1, \dots, e_l) the canonical basis of \mathbf{R}^l . Then,

$$\int_{\mathcal{T} \times \mathbf{R}} \varphi(x, f(t)) \tilde{\nu}(dt, dx) = \sum_{j=1}^l \int_{C_j \times \mathbf{R}} \varphi(x, f(t)) \tilde{\nu}(dt, dx).$$

Using (2.3.16), we have $f|_{C_j} = M e_j$ for all $j \in \llbracket 1, l \rrbracket$. Hence

$$\int_{\mathcal{T} \times \mathbf{R}} \varphi(x, f(t)) \tilde{\nu}(dt, dx) = \sum_{j=1}^l \int_{\mathbf{R}} \varphi(x, M e_j) \Delta \nu_{C_j}(dx). \quad (2.3.20)$$

Let $\xi \in \mathbf{R}^k$. Equation (2.3.17) reads

$$\psi_{\mathcal{X}(f)}(\xi) = \sum_{j=1}^l \int_{\mathbf{R}} \left(e^{i(M^\top \xi)_j x} - 1 - i(M^\top \xi)_j \llbracket x \rrbracket \right) \Delta \nu_{C_j}(dx).$$

Since for all $j \in \llbracket 1, l \rrbracket$, $(M^\top \xi)_j = \langle \xi, M e_j \rangle$, we have

$$\psi_{\mathcal{X}(f)}(\xi) = \sum_{j=1}^l \int_{\mathbf{R}} \left(e^{i\langle \xi, x M e_j \rangle} - 1 - i\langle \xi, \llbracket x \rrbracket M e_j \rangle \right) \Delta \nu_{C_j}(dx).$$

Using the usual construction of the integral, we may extend (2.3.20) to $\varphi : \mathbf{R} \times \mathbf{R}^k \ni (x, y) \mapsto e^{i\langle \xi, x, y \rangle} - 1 - i\langle \xi, \llbracket x \rrbracket, y \rangle$ and get

$$\psi_{\mathcal{X}(f)}(\xi) = \int_{\mathcal{T} \times \mathbf{R}} \left(e^{i\langle \xi, x, f(t) \rangle} - 1 - i\langle \xi, \llbracket x \rrbracket, f(t) \rangle \right) \tilde{\nu}(dt, dx).$$

Hence

$$\begin{aligned} \psi_{\mathcal{X}(f)}(\xi) &= i \left\langle \xi, \int_{\mathcal{T} \times \mathbf{R}} (\llbracket x \rrbracket \cdot f(t)) - \llbracket x \rrbracket \cdot f(t) \right\rangle \tilde{\nu}(dt, dx) \\ &\quad + \int_{\mathcal{T} \times \mathbf{R}} \left(e^{i\langle \xi, x, f(t) \rangle} - 1 - i\langle \xi, \llbracket x \rrbracket, f(t) \rangle \right) \tilde{\nu}(dt, dx). \end{aligned} \quad (2.3.21)$$

The result follows from (2.3.18), (2.3.19), (2.3.21) and the unicity of the Lévy-Khintchine triplet. \square

2.3.3 Martingale inequalities

In this section, we investigate some martingale maximal inequalities that can be proven in the case of indexing collections of finite dimension (Definition 1.5.3). Such results are not too surprising since the classical set-indexed setting has always been geared towards generalizing martingale theory. They will be useful to establish regularity results in Chapter 3. For each inequality,

the basic ingredients are the same: import a discrete version of it from the multiparameter case through the finite-dimensional assumption and then, deduce a continuous version using additional regularity of the sample paths.

For this endeavor, we need to borrow a bit more of martingale-related vocabulary from the classical setting [47] and from the multiparameter setting [56]. We also take up on the occasion to establish a link between set-indexed and multiparameter martingales (Theorem 2.3.25).

DEFINITION 2.3.19 (Filtration, adapted from [47, Section 1.4]). A (set-indexed) filtration is a family $(\mathcal{F}_A)_{A \in \mathcal{A}}$ of σ -algebras such that

$$\forall A, A' \in \mathcal{A}, \quad A \subseteq A' \implies \mathcal{F}_A \subseteq \mathcal{F}_{A'}.$$

In the rest of this section, let $(\mathcal{F}_A)_{A \in \mathcal{A}}$ denote such a set-indexed filtration. The classical setting usually imposes more conditions on the filtration, but we will not use have a use for them here.

DEFINITION 2.3.20 (Strong history, [47, Section 1.4]). The strong history associated with $(\mathcal{F}_A)_{A \in \mathcal{A}}$ is the collection $(\mathcal{G}_U^*)_{U \in \mathcal{C}(u)}$ of σ -algebras defined for all $U \in \mathcal{C}(u)$ by $\mathcal{G}_U^* = \bigvee_{A \in \mathcal{A}: A \cap U = \emptyset} \mathcal{F}_A$.

DEFINITION 2.3.21 (Set-indexed martingale, [47, Definition 3.1.1]). The process X is a strong (set-indexed) martingale with respect to $(\mathcal{F}_A)_{A \in \mathcal{A}}$ if for all $A \in \mathcal{A}$, $X_A \in L^1(\Omega, \mathcal{F}_A, \mathbf{P})$ and for all $C \in \mathcal{C}$, $\mathbf{E}[\Delta X_C | \mathcal{G}_C^*] = 0$.

The process X is a (set-indexed) martingale with respect to $(\mathcal{F}_A)_{A \in \mathcal{A}}$ if for all $A \in \mathcal{A}$, $X_A \in L^1(\Omega, \mathcal{F}_A, \mathbf{P})$ and for all $A, A' \in \mathcal{A}$ such that $A \subseteq A'$, $\mathbf{E}[X_{A'} | \mathcal{F}_A] = X_A$.

We may replace in the terminology above ‘martingale’ by ‘submartingale’ (resp. ‘supermartingale’) if the ‘=’ signs are replaced by ‘ \geq ’ (resp. ‘ \leq ’).

By [47, Proposition 3.1.4], any strong martingale is a martingale. There is also a notion of weak martingale [47, Definition 3.1.1], but we will not have a use for it here.

EXAMPLE 2.3.22. If X has independent increments, then it is a strong martingale with respect to its natural filtration where for all $A \in \mathcal{A}$, $\mathcal{F}_A = \sigma(X_{A'} : A' \subseteq A)$. Basically, this is the main setting in which we will use martingale inequalities afterwards. This is why we still chose to include this section here even though a martingale may not have independent increments in general.

We prove a small result about strong martingales that will alleviate the proof of the coming theorem.

LEMMA 2.3.23. If X is a strong martingale (resp. submartingale, supermartingale), then for all $U \in \mathcal{C}(u)$, $\mathbf{E}[\Delta X_U | \mathcal{G}_U^*] = 0$ (resp. ≥ 0 , ≤ 0).

Proof. Suppose that X is a strong martingale and consider $U \in \mathcal{C}(u)$. We write $U = \bigcup_{i=1}^k C_i$ where the C_i ’s are pairwise disjoint elements of \mathcal{C} . Then, since for all $i \in \llbracket 1, k \rrbracket$, $\mathcal{G}_U^* \subseteq \mathcal{G}_{C_i}^*$, we may use the tower rule of conditional expectations to get

$$\mathbf{E}[\Delta X_U | \mathcal{G}_U^*] = \sum_{i=1}^k \mathbf{E}[\mathbf{E}[\Delta X_{C_i} | \mathcal{G}_{C_i}^*] | \mathcal{G}_U^*] = 0.$$

□

On the side of the multiparameter setting, the notion of maringale has been generalized by Cairoli and Walsh as follows.

DEFINITION 2.3.24 (Orthomartingale, [56, Section 2.1]). Let $p \in \mathbf{N}^*$ and consider p filtrations $\mathcal{F}^{(1)}, \dots, \mathcal{F}^{(p)}$ on \mathbf{N} . A process $M = \{M_t : t \in \mathbf{N}^p\}$ is an orthomartingale (resp. orthosubmartingale) with respect to $\mathcal{F}^{(1)}, \dots, \mathcal{F}^{(p)}$ if for all $t = (t_1, \dots, t_p) \in \mathbf{N}^p$ and $j \in \llbracket 1, p \rrbracket$, the process $\{M_{(t_1, \dots, t_{j-1}, s, t_{j+1}, \dots, t_p)} : s \in \mathbf{N}\}$ is a one-dimensional martingale (resp. submartingale) with respect to $\mathcal{F}^{(j)}$.

Actually, there is a link between set-indexed martingales and orthomartingales. We will exploit it by directly importing results from one to the other.

THEOREM 2.3.25. Consider the following statements:

(i) X is a strong martingale with respect to $(\mathcal{F}_A)_{A \in \mathcal{A}}$.

(ii) For any finite subset $\mathcal{A}' \subseteq \mathcal{A}$ closed under intersections and order embedding $\varphi : (\mathcal{A}', \subseteq) \hookrightarrow (\mathbf{N}^p, \preceq)$ where $p \in \mathbf{N}^*$ is such that \mathcal{A}' has poset dimension $\leq p$, the process $X|^\varphi = \{X|_t^\varphi : t \in \mathbf{N}^p\}$ given by

$$\forall t \in \mathbf{N}^p, \quad X|_t^\varphi = \Delta X|_{U|_t^\varphi} \quad \text{where} \quad U|_t^\varphi = \bigcup_{\substack{A \in \mathcal{A}' \\ \varphi(A) \preceq t}} A \quad (2.3.22)$$

is an orthomartingale with respect to $\mathcal{F}^{(1)}, \dots, \mathcal{F}^{(p)}$ where

$$\forall (j, s) \in \llbracket 1, p \rrbracket \times \mathbf{N}, \quad \mathcal{F}_s^{(j)} = \bigvee_{\substack{A \in \mathcal{A}' \\ \varphi(A)_j \preceq s}} \mathcal{F}_A.$$

(iii) X is a martingale with respect to $(\mathcal{F}_A)_{A \in \mathcal{A}}$.

Then (i) \Rightarrow (ii) \Rightarrow (iii). Moreover, this chain of implications still holds if ‘martingale’ is replaced by ‘submartingale’ or ‘supermartingale’.

REMARK 2.3.26. There is a more intuitive, but also less practical way to think about the process $X|^\varphi$. As usual, we may consider that $X|^\varphi$ is indexed by the rectangles $\llbracket 0, t \rrbracket$ ($t \in \mathbf{N}^p$) instead. Then, $X|^\varphi$ is equivalently characterized by

$$\forall t \in \mathbf{N}^p, \quad \Delta X|_{\{t\}}^\varphi = \begin{cases} \Delta X_C & \text{if } t \in \varphi(\mathcal{A}') \text{ and } C \in \mathcal{C}^\ell(\mathcal{A}') \text{ is such that } \mathbf{t}(\varphi^{-1}(t)) \in C, \\ 0 & \text{otherwise.} \end{cases}$$

Since the elements of $\mathcal{C}^\ell(\mathcal{A}')$ are pairwise disjoint, this definition is well-posed. Proving the equivalence with (2.3.22) is a straightforward computation.

Proof of Theorem 2.3.25. We only prove it for martingales, but the other cases only require straightforward adjustments. Suppose (i) and consider an order embedding $\varphi : \mathcal{A}' \hookrightarrow \mathbf{N}^p$ for some finite subset $\mathcal{A}' \subseteq \mathcal{A}$ closed under intersections and $p \in \mathbf{N}^*$. Let us show that $X|^\varphi$ is an orthomartingale. Let $t \in \mathbf{N}^p$ and $j \in \llbracket 1, p \rrbracket$. Due to the inclusion-exclusion formula and the fact that \mathcal{A}' is closed under intersections, $X|_t^\varphi$ is a linear combination of the X_A ’s such that $A \in \mathcal{A}'$ and $\varphi(A) \preceq t$. Hence $X|_t^\varphi$ is integrable and $X|_t^\varphi$ is $\mathcal{F}_{t_j}^{(j)}$ -measurable.

Denote $t' = (t_1, \dots, t_{j-1}, t_j + 1, t_{j+1}, \dots, t_p)$. We claim that

$$\mathcal{F}_{t_j}^{(j)} \subseteq \mathcal{G}_{U|_{t'}^\varphi \cup U|_t^\varphi}^*. \quad (2.3.23)$$

Indeed, consider $A \in \mathcal{A}'$ such that $\varphi(A)_j \leq t_j$. By definition of $U|_t^\varphi$ and $U|_{t'}^\varphi$, we have

$$A \cap (U|_{t'}^\varphi \setminus U|_t^\varphi) = \bigcup_{\substack{A' \in \mathcal{A}'; \\ \varphi(A') \preceq t'}} (A \cap A') \setminus \bigcup_{\substack{A' \in \mathcal{A}'; \\ \varphi(A') \preceq t}} A'. \quad (2.3.24)$$

Moreover, for all $A' \in \mathcal{A}'$ such that $\varphi(A') \preceq t'$,

$$\begin{aligned} \varphi(A \cap A') &= \varphi(A) \wedge \varphi(A') && \text{since } \varphi \text{ is order embedding,} \\ &\preceq \varphi(A) \wedge t' \\ &\preceq t && \text{since } \varphi(A)_j \leq t_j \text{ and for all } i \neq j, t'_i = t_i. \end{aligned}$$

Combining this result with (2.3.24) yields $A \cap (U|_{t'}^\varphi \setminus U|_t^\varphi) = \emptyset$. Hence (2.3.23) holds by definition of $\mathcal{F}_{t_j}^{(j)}$ and $\mathcal{G}_{U|_{t'}^\varphi \setminus U|_t^\varphi}^*$. Then we have

$$\begin{aligned} \mathbf{E}[X|_{t'}^\varphi - X|_t^\varphi | \mathcal{F}_{t_j}^{(j)}] &= \mathbf{E}[\Delta X_{U|_{t'}^\varphi \setminus U|_t^\varphi} | \mathcal{F}_{t_j}^{(j)}] \\ &= \mathbf{E}\left[\mathbf{E}\left[\Delta X_{U|_{t'}^\varphi \setminus U|_t^\varphi} | \mathcal{G}_{U|_{t'}^\varphi \setminus U|_t^\varphi}^*\right] | \mathcal{F}_{t_j}^{(j)}\right] && \text{by (2.3.23),} \\ &= 0 && \text{by Lemma 2.3.23.} \end{aligned}$$

Hence $X|^\varphi$ is an orthomartingale, which proves (i) \Rightarrow (ii).

Now suppose (ii) and let us prove (iii). Let $A, A' \in \mathcal{A}$ such that $A \subseteq A'$. Then, the map $\varphi = \mathbb{1}_{\{A\}}$ is an order embedding from $\{A, A'\}$ to \mathbf{N} . By (ii), we know that $X|^\varphi$ is an orthomartingale with respect to $\mathcal{F}^{(1)}$ where

$$\forall s \in \mathbf{N}, \quad \mathcal{F}_s^{(1)} = \begin{cases} \mathcal{F}_A & \text{if } s = 0, \\ \mathcal{F}_{A'} & \text{if } s \geq 1. \end{cases}$$

It then quickly follows that $X_A \in L^1(\Omega, \mathcal{F}_A, \mathbf{P})$, $X_{A'} \in L^1(\Omega, \mathcal{F}_{A'}, \mathbf{P})$ and $\mathbf{E}[X_{A'} | \mathcal{F}_A] = X_A$. Hence X is a martingale and (ii) \Rightarrow (iii). \square

LEMMA 2.3.27 (Discrete set-indexed maximal inequality). *Let $p \in \mathbf{N}^*$ and suppose that X is a non-negative strong submartingale. There exists a constant $\kappa_{p,1} > 0$ such that for all finite subset $\mathcal{A}' \subseteq \mathcal{A}$ with poset dimension $\leq p$ and $U \in \mathcal{A}'(u)$,*

$$\forall \varepsilon > 0, \quad \mathbf{P}\left(\max_{A \in [\emptyset, U] \cap \mathcal{A}'} |X_A| > \varepsilon\right) \leq \frac{\kappa_{p,1}}{\varepsilon} \left((p-1) + \mathbf{E}[|\Delta X_U| (\ln_+ |\Delta X_U|)^{p-1}]\right) \quad (2.3.25)$$

where $\ln_+ = 0 \vee \ln$ and $\kappa_{p,1}$ may be taken equal to $(e/(e-1))^{p-1}$.

For all $\gamma > 1$, there exists a constant $\kappa_{p,\gamma} > 0$ such that for all finite subset $\mathcal{A}' \subseteq \mathcal{A}$ with poset dimension $\leq p$ and $U \in \mathcal{A}'(u)$,

$$\mathbf{E}\left[\max_{A \in [\emptyset, U] \cap \mathcal{A}'} |X_A|^\gamma\right] \leq \kappa_{p,\gamma} \mathbf{E}[|\Delta X_U|^\gamma] \quad (2.3.26)$$

where $\kappa_{p,\gamma}$ may be taken equal to $(\gamma/(\gamma-1))^{p\gamma}$.

The one-dimensional case is the celebrated Doob's maximal inequality. The multiparameter case is called Cairoli's inequality and its special case $\gamma = 2$ in (2.3.26) is also known as Wichura's inequality [101, Theorem 1].

Proof. Suppose that X is a non-negative strong submartingale and consider a finite subset $\mathcal{A}' \subseteq \mathcal{A}$ and $U \in \mathcal{A}'(u)$. Without loss of generality for what we want to prove, we may suppose that \mathcal{A}' is closed under intersections and that $U = \bigcup_{A \in \mathcal{A}'} A$. Fix an order embedding $\varphi : \mathcal{A}' \hookrightarrow \mathbf{N}^p$ where $p \in \mathbf{N}^*$.

By Theorem 2.3.25, $X|^\varphi$ is a non-negative orthosubmartingale for which we may apply Cairoli's inequality [56, Theorem 2.5.1] and obtain for all $t \in \mathbf{N}^p$,

$$\forall \varepsilon > 0, \quad \mathbf{P} \left(\max_{0 \leq s \leq t} X|_s^\varphi > \varepsilon \right) \leq \frac{\kappa_{p,1}}{\varepsilon} (p-1 + \mathbf{E} [X|_t^\varphi (\ln_+ X|_t^\varphi)^{p-1}]) \quad (2.3.27)$$

where $\kappa_{p,1} = (e/(e-1))^{p-1}$. Then apply (2.3.27) for $t \in \mathbf{N}^p$ big enough so that for all $A \in \mathcal{A}'$, $\varphi(A) \leq t$ and (2.3.25) follows. The inequality (2.3.26) works all the same but for the need to apply another Cairoli's inequality [56, Theorem 2.3.1] instead. \square

As usual, extending those discrete results requires some additional regularity assumption.

DEFINITION 2.3.28 (Outer continuity). *A map $h : \mathcal{A} \rightarrow \mathbf{R}$ is outer continuous if for all non-increasing sequence $(A_k)_{k \in \mathbf{N}}$ in \mathcal{A} , $h(A_k) \rightarrow h(\bigcap_{k \in \mathbf{N}} A_k)$ as $k \rightarrow \infty$.*

THEOREM 2.3.29 (Set-indexed maximal inequality). *Suppose that $p = \dim \mathcal{A} < \infty$ and X is a non-negative strong submartingale with an outer continuous version. Then there exists a constant $\kappa_{p,1} > 0$ such that for all $U \in \mathcal{A}(u)$,*

$$\forall \varepsilon > 0, \quad \mathbf{P} \left(\sup_{A \in [\emptyset, U]} |X_A| > \varepsilon \right) \leq \frac{\kappa_{p,1}}{\varepsilon} (p-1 + \mathbf{E} [|\Delta X_U| (\ln_+ |\Delta X_U|)^{p-1}]) \quad (2.3.28)$$

where $\kappa_{p,1}$ may be taken equal to $(e/(e-1))^{p-1}$.

For all $\gamma > 1$, there exists a constant $\kappa_{p,\gamma} > 0$ such that for all $U \in \mathcal{A}(u)$,

$$\mathbf{E} \left[\sup_{A \in [\emptyset, U]} |X_A|^\gamma \right] \leq \kappa_{p,\gamma} \mathbf{E} [|\Delta X_U|^\gamma] \quad (2.3.29)$$

where $\kappa_{p,\gamma}$ may be taken equal to $(\gamma/(\gamma-1))^{p\gamma}$.

Proof. Let $U \in \mathcal{A}(u)$ and define for all $n \in \mathbf{N}$, $\mathcal{A}'_n = \{A \cap U : A \in \mathcal{A}_n\}$. Remark that due to the separability from above, for all n big enough, we have $U \in \mathcal{A}'_n(u)$.

By outer continuity, we have for all $\varepsilon > 0$,

$$\mathbf{P} \left(\sup_{A \in [\emptyset, U]} |X_A| > \varepsilon \right) = \lim_{n \rightarrow \infty} \mathbf{P} \left(\max_{A \in [\emptyset, U] \cap \mathcal{A}'_n} |X_A| > \varepsilon \right).$$

Hence (2.3.28) follows from its discrete counterpart (2.3.25) and the remark at the beginning of the proof.

For all $\gamma > 1$, we also have by outer continuity and Fatou's lemma,

$$\mathbf{E} \left[\sup_{A \in [\emptyset, U]} |X_A|^\gamma \right] \leq \liminf_{n \rightarrow \infty} \mathbf{E} \left[\max_{A \in [\emptyset, U] \cap \mathcal{A}'_n} |X_A|^\gamma \right]$$

which likewise enables us to prove (2.3.29) from (2.3.26) and the remark at the beginning of the proof. \square

REMARK 2.3.30. If X is a strong martingale and not necessarily a non-negative submartingale, then the maximal inequalities above still work fine. Indeed, by Theorem 2.3.25, we know that for all order embedding φ , $X|^\varphi$ is an orthomartingale. Hence, by a classical argument based on Jensen's inequality, $|X|^\varphi$ is a non-negative orthosubmartingale, which is the only property on X that is used in the proofs (apart from the occasional assumption on the sample paths).

The set-indexed setting also permits another kind of maximal inequality which is localized around some $A \in \mathcal{A}$. And for that, an interesting — but not so surprising phenomenon — appears: controlling the increments $X_A - X_{A'}$ in some neighborhood of A requires a knowledge beyond that neighborhood, knowledge contained in our notion of vicinity $\mathcal{V}(A, \rho)$ (Definition 1.4.6).

COROLLARY 2.3.31 (Localized maximal inequality). *Suppose that $p = \dim \mathcal{A} < \infty$, X is integrable, centered, has an outer continuous version, independent increments and extends to a stochastic measure ΔX . Then for all $\gamma > 1$, there exists a constant $\kappa_{p,\gamma} > 0$ such that for all $A \in \mathcal{A}$ and $\rho > 0$ such that $\mathcal{V}(A, \rho) \in \mathcal{B}(X)$,*

$$\mathbf{E} \left[\sup_{A' \in B_{\mathcal{A}}(A, \rho)} |X_A - X_{A'}|^\gamma \right] \leq \kappa_{p,\gamma} \mathbf{E} [|\Delta X_{\mathcal{V}(A, \rho)}|^\gamma] \quad (2.3.30)$$

where $\kappa_{p,\gamma}$ may be taken equal to $2^\gamma (\gamma/(\gamma-1))^{p\gamma}$.

Proof. Let $A \in \mathcal{A}$, $\rho > 0$, $\gamma > 1$ and $n \in \mathbf{N}$. We will use the notations $\mathcal{V}_n(A, \rho)$, $\overline{\mathcal{V}}_n(A, \rho)$, and $\underline{\mathcal{V}}_n(A, \rho)$ introduced in (1.4.12). Applying (2.3.29) to the strong martingale $Y^{(n)} = \{\Delta X_{A \setminus \underline{\mathcal{V}}_k(A, \rho)} : A \in \mathcal{A}\}$ and $U = \overline{\mathcal{V}}_n(A, \rho)$ where $\underline{\mathcal{V}}(A, \rho) = \bigcap_{k \in \mathbf{N}} \underline{\mathcal{V}}_k(A, \rho)$ gives

$$\mathbf{E} \left[\sup_{\substack{A' \in \mathcal{A}: \\ A' \subseteq \overline{\mathcal{V}}_n(A, \rho)}} |\Delta X_{A' \setminus \underline{\mathcal{V}}_n(A, \rho)}|^\gamma \right] \leq \left(\frac{\gamma}{\gamma-1} \right)^{p\gamma} \mathbf{E} [|\Delta X_{\overline{\mathcal{V}}_n(A, \rho) \setminus \underline{\mathcal{V}}_n(A, \rho)}|^\gamma]. \quad (2.3.31)$$

Using the inequality $|X_A - X_{A'}| \leq |\Delta X_{A \setminus \underline{\mathcal{V}}_n(A, \rho)}| + |\Delta X_{A' \setminus \underline{\mathcal{V}}_n(A, \rho)}|$ together with the convexity of $x \mapsto |x|^\gamma$ yields

$$\forall A' \in B_{\mathcal{A}}(A, \rho), \quad |X_A - X_{A'}|^\gamma \leq 2^{\gamma-1} (|\Delta X_{A \setminus \underline{\mathcal{V}}_n(A, \rho)}|^\gamma + |\Delta X_{A' \setminus \underline{\mathcal{V}}_n(A, \rho)}|^\gamma). \quad (2.3.32)$$

Combining (2.3.31) and (2.3.32) yields

$$\mathbf{E} \left[\sup_{\substack{A' \in B_{\mathcal{A}}(A, \rho): \\ A' \subseteq \overline{\mathcal{V}}_n(A, \rho)}} |X_A - X_{A'}|^\gamma \right] \leq 2^\gamma \left(\frac{\gamma}{\gamma-1} \right)^{p\gamma} \mathbf{E} [|\Delta X_{\overline{\mathcal{V}}_n(A, \rho) \setminus \underline{\mathcal{V}}_n(A, \rho)}|^\gamma]. \quad (2.3.33)$$

Since X has independent increments and $\overline{\mathcal{V}}_n(A, \rho) \setminus \underline{\mathcal{V}}_n(A, \rho) \subseteq \mathcal{V}(A, \rho)$, we know by Proposition 2.3.2 that

$$\mathbf{E} [\Delta X_{\mathcal{V}(A, \rho)} \mid \Delta X_{\overline{\mathcal{V}}_n(A, \rho) \setminus \underline{\mathcal{V}}_n(A, \rho)}] = \Delta X_{\overline{\mathcal{V}}_n(A, \rho) \setminus \underline{\mathcal{V}}_n(A, \rho)}. \quad (2.3.34)$$

Hence, by convexity of $x \mapsto |x|^\gamma$ and Jensen's inequality, we get

$$\mathbf{E} [|\Delta X_{\overline{\mathcal{V}}_n(A, \rho) \setminus \underline{\mathcal{V}}_n(A, \rho)}|^\gamma] \leq \mathbf{E} [|\Delta X_{\mathcal{V}(A, \rho)}|^\gamma]. \quad (2.3.35)$$

Combining (2.3.33) and (2.3.35) yields

$$\forall n \in \mathbf{N}, \quad \mathbf{E} \left[\sup_{\substack{A' \in B_{\mathcal{A}}(A, \rho): \\ A' \subseteq \overline{\mathcal{V}}_n(A, \rho)}} |X_A - X_{A'}|^\gamma \right] \leq 2^\gamma \left(\frac{\gamma}{\gamma-1} \right)^{p\gamma} \mathbf{E} [|\Delta X_{\mathcal{V}(A, \rho)}|^\gamma]. \quad (2.3.36)$$

The result follows by taking $n \rightarrow \infty$ in (2.3.36), Fatou's lemma and outer continuity, just like the proof of (2.3.29). \square

REMARK 2.3.32. *The only place where we used the independence of the increments in the proof of Corollary 2.3.31 is to obtain the martingale-like relation (2.3.34). Unfortunately, it seems not to be a direct consequence of being a strong martingale, even though some additional regularity assumptions in $L^Y(\Omega)$ could also have done the trick. Our proof avoids this caveat, but cannot be adapted to a $L \log^{p-1} L$ inequality like (2.3.28) due to convexity issues.*

2.3.4 A 0-1 law

Among many interesting properties verified by processes with independent increments, one may find Blumenthal's 0-1 law at a good place, e.g. [30, Theorem 8.2.3]. We prove a set-indexed version of such law. The main use we will have for it is to prove that the pointwise Hölder exponent of a set-indexed process with independent increments at some fixed $A \in \mathcal{A}$ must be deterministic (Proposition 3.5.1). Recall that the modular ρ_X has been introduced in (2.2.3) to characterize X -integrable functions.

THEOREM 2.3.33 (Set-indexed 0-1 law). *Suppose that X has independent increments and extends to a stochastic measure ΔX . Let $A \in \mathcal{A}$ and define*

$$\forall \rho > 0, \quad \mathcal{F}_{(A,\rho)} = \sigma(X_A - X_{A'} : A \in \mathcal{A}, 0 < d_{\mathcal{A}}(A, A') < \rho) \quad \text{and} \quad \mathcal{F}_{A^+} = \bigcap_{\rho > 0} \mathcal{F}_{(A,\rho)}.$$

If $\rho_X(\mathbb{1}_{\mathcal{V}(A,\rho)}) \rightarrow 0$ as $\rho \rightarrow 0^+$, then any event in \mathcal{F}_{A^+} has probability either 0 or 1.

Proof. Define the σ -algebra

$$\mathcal{F}_{(A,\infty)} = \bigvee_{\rho > 0} \mathcal{F}_{(A,\rho)}.$$

Since $\mathcal{F}_{A^+} \subseteq \mathcal{F}_{(A,\infty)}$, it is enough to prove that \mathcal{F}_{A^+} is independent from $\mathcal{F}_{(A,\infty)}$. First, remark that the family of cylinders

$$\{X_A - X_{A_1} \in B_1, \dots, X_A - X_{A_k} \in B_k\}$$

where $k \in \mathbf{N}^*$, and for all $j \in \llbracket 1, k \rrbracket$, $A_j \in \mathcal{A}$, $d_{\mathcal{A}}(A, A_j) > 0$ and $B_j \in \mathcal{B}(\mathbf{R})$ is a π -system that generates $\mathcal{F}_{(A,\infty)}$. So by a monotone class argument, it is enough to show that \mathcal{F}_{A^+} is independent from $X_A - X_{A_1}, \dots, X_A - X_{A_k}$ for some fixed $A_1, \dots, A_k \in \mathcal{A}$ such that for all $j \in \llbracket 1, k \rrbracket$, $d_{\mathcal{A}}(A, A_j) > 0$.

Let $\rho > 0$. Since $\mathcal{F}_{A^+} \subseteq \mathcal{F}_{(A,\rho)}$ and X has independent increments, \mathcal{F}_{A^+} is independent from the random variables

$$\{\Delta X_{A \setminus (A_j \cup \mathcal{V}(A,\rho))} - \Delta X_{A_j \setminus (A \cup \mathcal{V}(A,\rho))} : j \in \llbracket 1, k \rrbracket\}.$$

Since $\rho_X(\mathbb{1}_{\mathcal{V}(A,\rho)}) \rightarrow 0$ as $\rho \rightarrow 0^+$, \mathcal{F}_{A^+} is independent from the random variables

$$\left\{ \lim_{\rho \rightarrow 0^+} (\Delta X_{A \setminus (A_j \cup \mathcal{V}(A,\rho))} - \Delta X_{A_j \setminus (A \cup \mathcal{V}(A,\rho))}) = X_A - X_{A_j} : j \in \llbracket 1, k \rrbracket \right\}$$

where the limit holds in probability by Theorem 2.2.10. Hence \mathcal{F}_{A^+} is independent from $\mathcal{F}_{(A,\infty)}$. The result follows. \square

REMARK 2.3.34. *Making use of [62, Proposition 7.1.1 (ii) and (iii)] (which basically is a dominated convergence theorem for \mathcal{X}), the condition $\rho_X(\mathbb{1}_{\mathcal{V}(A,\rho)}) \rightarrow 0$ as $\rho \rightarrow 0^+$ of Theorem 2.3.33 is actually equivalent to $\mathbb{1}_{\mathcal{V}(A,\rho_0)} \in \mathcal{B}(X)$ for some $\rho_0 > 0$ and $\mathbf{m}_X(\bigcap_{\rho > 0} \mathcal{V}(A,\rho)) = 0$.*

2.4 Interlude: set-indexed Lévy processes

In Section 2.3, we studied processes with independent increments. In Section 2.5, we will study processes with stationary increments. Between them, we expose a process that has both of those properties: the set-indexed Lévy process. It has been introduced by Herbin and Merzbach in [41]. A multiparameter version has previously been studied by Bass and Pyke in [16].

We first give a few definitions and examples in Section 2.4.1 before commenting in Section 2.4.2 on the special form the Lévy-Khintchine representation (2.3.7) has in this case.

2.4.1 Definition and examples

DEFINITION 2.4.1 (Set-indexed Lévy process, adapted from [41, Definition 3.1]). *The process X is a set-indexed Lévy process (siLévy) if it verifies the following conditions:*

1. X has independent \mathcal{C} -increments,
2. X is \mathcal{C}_0 -stationary, i.e. for all $k \in \mathbf{N}^*$, $A \subseteq A_1 \subseteq \dots \subseteq A_k$ and $A' \subseteq A'_1 \subseteq \dots \subseteq A'_k$ in \mathcal{A} such that for all $j \in \llbracket 1, k \rrbracket$, $m(A_j \setminus A) = m(A'_j \setminus A')$, we have $(\Delta X_{A_1 \setminus A}, \dots, \Delta X_{A_k \setminus A}) \stackrel{\text{law}}{=} (\Delta X_{A'_1 \setminus A'}, \dots, \Delta X_{A'_k \setminus A'})$.

The property of \mathcal{C}_0 -stationarity will be commented at length in Section 2.5.2. If one were to compare this definition to [41, Definition 3.1], one would note that a stochastic continuity assumption is missing. However, the following result tells us that one may recover ‘for free’ a similar property.

PROPOSITION 2.4.2. *If X is a siLévy, then it is outer continuous in probability, i.e. for all non-increasing sequence $(A_k)_{k \in \mathbf{N}}$ in \mathcal{A} , we have $X_{A_k} \xrightarrow{\mathbf{P}} X_{\bigcap_{j \in \mathbf{N}} A_j}$ as $k \rightarrow \infty$.*

Proof. Consider a non-increasing sequence $(A_k)_{k \in \mathbf{N}}$ in \mathcal{A} and denote $A = \bigcap_{j \in \mathbf{N}} A_j$. By patching together \mathcal{A} -flows from Proposition 1.3.9, we may find a continuous \mathcal{A} -flow ϕ and a non-increasing sequence $(t_k)_{k \in \mathbf{N}}$ in $[0, 1]$ such that $\phi(0) = A$ and for all $k \in \mathbf{N}$, $\phi(t_k) = A_k$. Then, the projection X^ϕ is a $[0, 1]$ -indexed process with independent increments (see Proposition 2.5.2 later). According to [87, Theorem 9.7], we know that

$$X_{t_k}^\phi - X_0^\phi \xrightarrow{\text{law}} 0 \quad \text{as} \quad k \rightarrow \infty.$$

Since the convergence in law to a constant implies convergence in probability, we get

$$X_{A_k} \xrightarrow{\mathbf{P}} X_A \quad \text{as} \quad k \rightarrow \infty.$$

The result follows. □

Remark that in the proof above, [87, Theorem 9.7] is used on a $[0, 1]$ -indexed process instead of a \mathbf{R}_+ -indexed one as required by the statement. But that is a non-issue since one may extend the distribution of X^ϕ to such a process by independence of increments and 1-periodicity.

By now, the reader may already be familiar with several examples of siLévy’s. The simplest one is the *deterministic drift* given by $\{bm(A) : A \in \mathcal{A}\}$ for some $b \in \mathbf{R}$. Another one is the siBm defined in Example 2.2.12. Both of those being ‘jumpsless’ (see [41] for more details), it leaves out a discontinuous part — the Poissonian one from Remark 2.3.9 — that we describe a bit more

precisely here. The presentation is made from a standpoint that will generalize easily to several processes — namely ones with exchangeable increments — that will be encountered later on in Section 2.5.5.

When studying the jump discontinuities of some stochastic process, arises a particular class of processes called *point processes* of independent interest. This class has a long history that dates back to the 70's (see e.g. [77] for an early reference). For modern expositions on the theory of point processes, we refer to [54, Chapter 12] and [64]. We adapt the definitions here to the set-indexed setting. Note that they differ from the classical one [47, Section 8].

DEFINITION 2.4.3 (Set-indexed point process). *A (set-indexed) point process is a process $N = \{\mu(A) : A \in \mathcal{A}\}$ where $\mu : \Omega \rightarrow \mathcal{M}(\mathcal{T})$ is a $\mathbf{N} \cup \{\infty\}$ -valued random measure (Definition 2.2.14).*

REMARK 2.4.4 (Random sets). *Point processes may be used to represent random countable sets Π by considering the measure $\sum_{t \in \Pi} \delta_t$ instead. This point of view has been successfully used in [13, 51] to study the Hölder regularity of \mathbf{R}_+ -indexed Lévy processes. In the Chapter 3, we shall push their methods further to make use of them in the set-indexed setting.*

For a more complete exposition of the theory of random sets, we refer to [64, 75].

DEFINITION 2.4.5 (Set-indexed Poisson process, adapted from [47, Definition 3.4.4]). *Let $\nu \in \mathcal{M}(\mathcal{T})$. A set-indexed Poisson process of intensity ν is a point process $N = \{N_A : A \in \mathcal{A}\}$ with independent \mathcal{C} -increments and such that for all $A \in \mathcal{A}$, $N_A \sim \text{Poi}(\nu(A))$ with the convention that $N_A = \nu(A)$ a.s. whenever $\nu(A) \in \{0, \infty\}$.*

The existence of a set-indexed Poisson process is ensured by [54, Theorem 12.7] even though we will be explaining part of the construction in Proposition 3.3.11.

EXAMPLE 2.4.6 (Set-indexed compound Poisson process). *As expressed in Remark 2.3.9, the Poissonian part of a Lévy process is more than simply a Poisson process. In general, we have what is called a compensated compound Poisson process. Let us explain here how it is defined in the set-indexed setting.*

Consider $\nu \in \mathcal{M}(\mathbf{R})$ such that $\nu(\{0\}) = 0$. Recall that we endowed $\mathbf{R}^ \cup \{\infty\}$ with an indexing collection $\mathcal{A}_{\mathbf{R}^* \cup \{\infty\}}$ in Example 1.2.19. In particular, we may consider that $\nu \in \mathcal{M}(\mathbf{R}^* \cup \{\infty\})$ by specifying that $\nu(\{\infty\}) = 0$. Since the product of indexing collections is still an indexing collection (Proposition 1.2.9), we thus know that $\mathbf{m} \otimes \nu \in \mathcal{M}(\mathcal{T} \times (\mathbf{R}^* \cup \{\infty\}))$. In particular, we may consider a set-indexed Poisson process N of intensity $\mathbf{m} \otimes \nu$.*

If $\int_{\mathbf{R}} (1 \wedge |x|) \nu(dx) < \infty$, we may define a set-indexed compound Poisson process of intensity ν as the process P given by

$$\forall A \in \mathcal{A}, \quad P_A = \int_{\mathcal{T} \times (\mathbf{R}^* \cup \{\infty\})} [x \mathbb{1}_A(t)] N(dt, dx) = \int_{A \times \mathbf{R}} x N(dt, dx) \quad (2.4.1)$$

where the integral converges due to the assumption on ν and [54, Lemma 12.13].

Suppose that instead of $\int_{\mathbf{R}} (1 \wedge |x|) \nu(dx) < \infty$, we have $\int_{\mathbf{R}} (1 \wedge x^2) \nu(dx) < \infty$ (i.e. ν is a Lévy measure on \mathbf{R}) and introduce a set-indexed compensated Poisson process of intensity ν as the process \tilde{N} given by

$$\forall A' \in \mathcal{A} \times \mathcal{A}_{\mathbf{R}^* \cup \{\infty\}}, \quad \tilde{N}_{A'} = N_{A'} - \mathbf{m} \otimes \nu(\{(t, x) \in A' : |x| \leq 1\}). \quad (2.4.2)$$

In that case, we may define a set-indexed compensated compound Poisson process of intensity ν as the process \tilde{P} given by

$$\forall A \in \mathcal{A}, \quad \tilde{P}_A = \int_{\mathcal{T} \times (\mathbf{R}^* \cup \{\infty\})} [x \mathbb{1}_A(t)] \tilde{N}(dt, dx) = \int_{A \times \mathbf{R}} x \tilde{N}(dt, dx) \quad (2.4.3)$$

where the integral converges once more due to [54, Lemma 12.13]. Seeing that \tilde{P} has independent \mathcal{C} -increments is a consequence of the fact that \tilde{N} has independent \mathcal{C} -increments and Proposition 2.3.2. We postpone checking the \mathcal{C}_0 -stationarity to Proposition 2.5.22 when we are a bit more equipped to talk about stationarity.

2.4.2 Weak Lévy-Itô decomposition

In [83], Rajput and Rosiński characterized the space $L(X)$ of X -integrable functions when X has independent increments. We state their result when X is a siLévy since we are going to need it afterwards. For $\gamma \in [0, 2]$, $b \in \mathbf{R}$, $\sigma^2 \in \mathbf{R}_+$ and ν a Lévy measure on \mathbf{R} , denote by $L^\gamma(b, \sigma^2, \nu)$ the Musielak-Orlicz space associated with the function

$$\begin{aligned} \phi_\gamma : \mathcal{T} \times \mathbf{R}_+ &\longrightarrow \mathbf{R}_+ \\ (t, x) &\longmapsto \sup_{\substack{z \in \mathbf{R}: \\ |z| \leq |x|}} \left| by + \int_{\mathbf{R}} (\llbracket yz \rrbracket - z \llbracket y \rrbracket) \nu(dy) \right| + \sigma^2 x^2 \\ &\quad + \int_{\mathcal{T} \times \mathbf{R}} (|xy|^2 \wedge |xy|^\gamma) \nu(dy) \end{aligned} \quad (2.4.4)$$

which does verify the conditions of Definition 2.2.4 due to [83, Lemma 3.1]. The parameter γ is used to characterize the space of X -integrable functions whose stochastic integrals have a moment of order γ .

THEOREM 2.4.7 (Lévy-Khintchine representation). *The following statements are equivalent:*

- (i) X is a siLévy.
- (ii) X has independent increments and its modified triplet $(\Delta \mathbf{b}, \Delta \sigma^2, \tilde{\nu})$ may be expressed as

$$(\Delta \mathbf{b}, \Delta \sigma^2, \tilde{\nu}) = (b\mathbf{m}, \sigma^2 \mathbf{m}, \mathbf{m} \otimes \nu) \quad (2.4.5)$$

where $b \in \mathbf{R}$, $\sigma^2 \in \mathbf{R}_+$ and ν is a Lévy measure on \mathbf{R} .

Under those conditions, the triplet (b, σ^2, ν) is unique and X extends to a linear ID process $\mathfrak{X} = \{\mathfrak{X}(f) : f \in L^0(b, \sigma^2, \nu)\}$ whose Lévy Khintchine triplet $(\mathbf{b}, \Sigma, \nu)$ may be expressed for all $f, g \in L^0(b, \sigma^2, \nu)$, finite subset $f \subseteq L^0(b, \sigma^2, \nu)$ and $B \in \mathcal{B}(\mathbf{R})^{\otimes f}$ by

$$\begin{aligned} \mathbf{b}(f) &= b \int_{\mathcal{T}} f d\mathbf{m} + \int_{\mathcal{T} \times \mathbf{R}} (\llbracket xf(t) \rrbracket - f(t) \llbracket x \rrbracket) (\mathbf{m} \otimes \nu)(dt, dx), \\ \Sigma(f, g) &= \sigma^2 \int_{\mathcal{T}} fg d\mathbf{m}, \\ \nu_f(B) &= \int_{\mathcal{T} \times \mathbf{R}} \mathbb{1}_B(x.f(t)) (\mathbf{m} \otimes \nu)(dt, dx) \end{aligned} \quad (2.4.6)$$

where ν_f is the Lévy measure induced by ν on \mathbf{R}^f and for all $t \in \mathcal{T}$, $x.f(t)$ is the element of \mathbf{R}^f that maps $h \in f$ to $xh(t) \in \mathbf{R}$.

Proof. The equivalence (i) \Leftrightarrow (ii) is essentially the computation that leads to [41, Equation (15)], the only difference being that our choice of cutoff function is more general.

Now, suppose (i) and (ii). Due to [83, Theorem 3.3], we have $L(X) = L^0(b, \sigma^2, \nu)$. The computation of (2.4.6) is then just a joint consequence of (2.3.15) and (2.4.5). \square

COROLLARY 2.4.8. *Suppose that the cutoff function χ is equal to $\mathbb{1}_{|\cdot| \leq 1}$ and consider a Lévy measure ν on \mathbf{R} . Then, the modified triplet of the set-indexed compensated compound Poisson process \tilde{P} of intensity ν is equal to $(0, 0, \mathbf{m} \otimes \nu)$.*

Proof. Consider a Lévy measure ν on \mathbf{R} and the set-indexed compensated Poisson process \tilde{N} from (2.4.2). We already know that \tilde{N} is an ID process on $\mathcal{T} \times (\mathbf{R}^* \cup \{\infty\})$. Using [54, Lemma 12.2], its modified triplet reads $(0, 0, \mathbf{m} \otimes \delta_1 \otimes \nu)$.

Rewriting (2.4.3), the process \tilde{P} is equivalently defined as follows:

$$\forall A \in \mathcal{A}, \quad \tilde{P}_A = \tilde{\mathcal{N}}(f_A) \quad \text{where} \quad f_A : (t, x) \mapsto x \cdot \mathbb{1}_A(t)$$

and $\tilde{\mathcal{N}}$ is the linear process associated with \tilde{N} . Thus, we may compute its Lévy-Khintchine triplet $(\mathbf{b}, \Sigma, \nu)$ using (2.4.6) to get for all $A \in \mathcal{A}$, finite subset $\mathcal{A}' \subseteq \mathcal{A}$ and $B \in \mathcal{B}(\mathbf{R})^{\otimes \mathcal{A}'}$,

$$\begin{aligned} \mathbf{b}(A) &= \int_{\mathcal{T} \times (\mathbf{R}^* \cup \{\infty\}) \times \mathbf{R}} (\llbracket x f_A(t, y) \rrbracket - f_A(t, y) \llbracket x \rrbracket) (\mathbf{m} \otimes \delta_1 \otimes \nu)(dt, dy, dx) \\ &= \int_{\mathcal{T} \times \mathbf{R}} x \mathbb{1}_A(t) (\mathbb{1}_{|x \mathbb{1}_A(t)| \leq 1} - \mathbb{1}_{|x| \leq 1}) (\mathbf{m} \otimes \nu)(dt, dx) \\ &= 0, \end{aligned}$$

$\Sigma = 0$ and

$$\begin{aligned} \nu_{\mathcal{A}'}(B) &= \int_{\mathcal{T} \times (\mathbf{R}^* \cup \{\infty\}) \times \mathbf{R}} \mathbb{1}_B(x \cdot (f_{A'}(t, y))_{A' \in \mathcal{A}'}) (\mathbf{m} \otimes \delta_1 \otimes \nu)(dt, dy, dx) \\ &= \int_{\mathcal{T} \times \mathbf{R}} \mathbb{1}_B(x \cdot (\mathbb{1}_{A'}(t))_{A' \in \mathcal{A}'}) (\mathbf{m} \otimes \nu)(dt, dx). \end{aligned}$$

Thus, using the unicity from Theorem 2.4.7, the modified triplet of \tilde{P} is equal to $(0, 0, \mathbf{m} \otimes \nu)$. \square

COROLLARY 2.4.9 (Weak Lévy-Itô decomposition). *The following statements are equivalent:*

- (i) X is a siLévy.
- (ii) There exists a triplet (b, σ^2, ν) where $b \in \mathbf{R}$, $\sigma^2 \in \mathbf{R}_+$ and ν is a Lévy measure on \mathbf{R} , a siBm W and an independent set-indexed compensated compound Poisson process \tilde{P} of intensity ν such that

$$\forall A \in \mathcal{A}, \quad X_A = b \mathbf{m}(A) + \sigma W_A + \tilde{P}_A. \quad (2.4.7)$$

Under those conditions, the triplet (b, σ^2, ν) is unique and the linear extension \mathcal{X} of X reads

$$\forall f \in L^0(b, \sigma^2, \nu), \quad \mathcal{X}(f) = b \int_{\mathcal{T}} f d\mathbf{m} + \sigma \mathcal{W}(f) + \tilde{\mathcal{P}}(f) \quad (2.4.8)$$

where \mathcal{W} (resp. $\tilde{\mathcal{P}}$) is the linear process associated with the process W (resp. \tilde{P}) in (2.4.7).

To be precise, the existence of W may not be ensured if $\sigma = 0$ without enlarging $(\Omega, \mathcal{F}, \mathbf{P})$ first. But in that case, we would still be able to make sense of (2.4.7) anyway.

The reason why we called it a ‘weak’ representation is because one usually requires a kind of local uniform convergence in the definition of \tilde{P} . Here, since \tilde{P}_A is defined as a stochastic integral, we only know that it is a limit in probability. For now, we do not know much about the sample paths of \tilde{P} . A better result is obtained with Theorem 3.3.8.

Proof. Suppose that the cutoff function χ is equal to $\mathbb{1}_{|\cdot| \leq 1}$. This will ease the computation of the upcoming triplets, but does not impact the generality of the statement since only the value of b is impacted by the choice of χ .

A quick computation shows that the modified triplet of a drift $\{b\mathbf{m}(A) : A \in \mathcal{A}\}$ (resp. a siBb W) is equal to $(b\mathbf{m}, 0, 0)$ (resp. $(0, \sigma^2\mathbf{m}, 0)$). Combining these with Corollary 2.4.8, we know that the modified triplet of the process in the right-hand side of (2.4.7) is equal to $(b\mathbf{m}, \sigma^2\mathbf{m}, \mathbf{m} \otimes \nu)$. Hence the equivalence (i) \Leftrightarrow (ii) would be a direct consequence of Theorem 2.4.7 if the $=$ sign in (2.4.7) were to be replaced by $\stackrel{\text{fdd}}{=}$. More precisely, the implication (ii) \Rightarrow (i) is already proven. Suppose now that (i) holds. Then we get

$$X \stackrel{\text{fdd}}{=} b\mathbf{m} + \sigma W + \tilde{P}$$

for some siBm W and set-indexed compensated compound Poisson process \tilde{P} of intensity ν . In order to conclude, we need a coupling argument. Since $\mathcal{A} = \bigcup_{n \in \mathbf{N}} \mathcal{A}_n$ is countable, the space $\mathbf{R}^{\mathcal{A}}$ endowed with its cylindrical σ -algebra is a Borel space. Thus [54, Theorem 6.10] applies and we may find \mathcal{A} -indexed processes W' and \tilde{P}' such that

$$(X, W', \tilde{P}') \stackrel{\text{fdd}}{=} (b\mathbf{m} + \sigma W + \tilde{P}, W|_{\mathcal{A}}, \tilde{P}|_{\mathcal{A}}). \quad (2.4.9)$$

In particular, (2.4.9) implies that (2.4.7) holds for all $A \in \mathcal{A}$. In order to extend the version to \mathcal{A} , one just needs to use Lemma 2.3.15 and repeat the argument giving Corollary 2.3.16.

The representation (2.4.8) is deduced in the same fashion. \square

2.5 Processes with stationary increments

Apart from having independent increments, another distributional property of processes that is often studied is increment stationarity. For a one-dimensional process $X = \{X_t : t \in \mathbf{R}_+\}$, having *stationary increments* means that

$$\forall \tau \in \mathbf{R}_+, \quad X - X_0 \stackrel{\text{fdd}}{=} \{X_{t+\tau} - X_\tau : t \in \mathbf{R}_+\}. \quad (2.5.1)$$

This definition is adaptable to $[0, 1]$ -indexed processes with obvious restrictions on the possible values for t and τ . Likewise, (2.5.1) also makes sense for a multivariate process.

Our interest for those processes sprang from the integral representations it yields. Indeed, the celebrated Bochner’s theorem [27, Equation (11.1’)] tells that a square-integrable process $X = \{X_t : t \in \mathbf{R}_+\}$ with stationary increments may be written as

$$\forall t \in \mathbf{R}_+, \quad X_t = \int_{\mathbf{R}} \widehat{\mathbb{1}}_{[0,t]} dZ \quad (2.5.2)$$

where $\widehat{(\cdot)}$ is the usual Fourier transform and $Z = \{Z_t : t \in \mathbf{R}\}$ is a square-integrable (complex-valued) process with orthogonal increments.

Actually, Bochner's theorem remains true (and is more natural) for processes with *weakly* stationary increments, *i.e.* where (2.5.1) is replaced by

$$\forall s, t, \tau \in \mathbf{R}_+, \quad \text{Cov}(X_{s+\tau} - X_\tau, X_{t+\tau} - X_\tau) = \text{Cov}(X_s - X_0, X_t - X_0).$$

On the other hand, Kallenberg strengthened (2.5.1) and proved that if X has exchangeable increments (Definition 2.5.11), then it must be a mixture of Lévy processes. Kallenberg's original result [53, Theorem 2.1] is a bit different, but he extended it in [55, Theorem 3.15].

This kind of representation brings us two goods news. First, (2.5.2) for instance may be used to extend X to a linear process $\mathcal{X} = \{\mathcal{X}(f) : f \in L(X)\}$ by the formula

$$\forall f \in L(X), \quad \mathcal{X}(f) = \int_{\mathbf{R}} \hat{f} dZ. \quad (2.5.3)$$

where $L(X) = \{f : \hat{f} \in L(Z)\}$. Secondly, integral representations are a nice way to define more general processes (see *e.g.* [12]) and study their sample paths (see *e.g.* [14, 84]).

Both of those aspects will be explored in the coming sections, but in the set-indexed framework where 'increment stationarity' has possibly several meanings that all generalize the usual one when $\mathcal{T} = \mathbf{R}_+$. Those properties will be studied from the least to the most restrictive one.

2.5.1 General philosophy for set-indexed stationarity

A first element that comes to mind is that (2.5.1) heavily relies on the additive structure on \mathbf{R}_+ to give a meaning to 'stationarity' whereas no such thing exists in the set-indexed framework. Herbin and Merzbach proposed in a series of works [39, 40, 41] an answer to this apparent paradox. Their definition evolved with time (compare [39, Definition 3.5] to [40, Definition 5.1] and [41, Definition 3.1]), but they answered the following requirements (for the latest versions at least):

1. Roughly speaking, increments corresponding to sets of same \mathbf{m} -measure should have the same distribution. What may vary is the collection of sets on which the property holds.
2. Once specified to the case where $\mathcal{T} = \mathbf{R}_+$ and \mathbf{m} is the Lebesgue measure, this should give back the usual definition (2.5.1).

In our attempts to give a satisfactory definition of 'increment stationarity', we shall keep those in mind as well. Moreover, Herbin and Merzbach had another requirement in mind, which was to encompass in their definition the case of the set-indexed fractional Brownian motion (sifBm) W^H defined in (1.6.3). When $\mathcal{T} = \mathbf{R}_+$, the fractional Brownian motion (fBm) is probably known to be the simplest example of a process with stationary increments that still exhibits a long-range dependency (when $H \neq 1/2$) and has been first considered by Kolmogorov in [60]. In this setting, integral representations of the fBm have proven to be useful at the very least in these regards:

1. Knowing the regularity of the kernel and the integrator often leads to understand the regularity of the process itself. This general approach is well illustrated by Rosiński's [84, Theorem 4]. For an example of application closer to the fBm and related processes, we refer to [14].
2. Another use of integral representations is to generalize processes. For instance for the fBm, they have been used to define the class of *multifractional Brownian motions (mBm)*. An historic overview gathering all possible definitions of the mfBm is given in [95].

For all those reasons, it seems to be an interesting goal to study the sifBm from the increment stationarity perspective, especially with a view to obtaining an integral representation for it. Unfortunately, we did not quite manage to do so. What we did manage however is to show that the sifBm actually verifies a stronger increment stationarity property than the one proposed by Herbin and Merzbach (Proposition 2.5.9). We also studied a more restrictive notion of increment stationarity and obtained an integral representation for it (Theorem 2.5.30).

As before, we continue to consider in the sequel a set-indexed process $X = \{X_A : A \in \mathcal{A}\}$.

2.5.2 \mathcal{A} -increment stationarity

As a starting point, let us give the definition of \mathcal{C}_0 -stationarity as given by Herbin and Merzbach. We recall that the notation \mathcal{C}_0 corresponds to $\mathcal{C}_{(1)}$ introduced in Definition 1.2.35.

DEFINITION 2.5.1 (\mathcal{C}_0 -stationarity, [40, Definition 5.1]). *The process X is \mathcal{C}_0 -stationary if for all $k \in \mathbf{N}^*$, $A \subseteq A_1 \subseteq \dots \subseteq A_k$ and $A' \subseteq A'_1 \subseteq \dots \subseteq A'_k$ in \mathcal{A} ,*

$$\forall j \in \llbracket 1, k \rrbracket, \quad \mathbf{m}(A_j \setminus A) = \mathbf{m}(A'_j \setminus A') \implies (\Delta X_{A_1 \setminus A}, \dots, \Delta X_{A_k \setminus A}) \stackrel{\text{law}}{=} (\Delta X_{A'_1 \setminus A'}, \dots, \Delta X_{A'_k \setminus A'}).$$

This definition is actually equivalent to a very natural property linked to projection along geodesic flows (Definition 1.3.6), and related to [40, Proposition 5.4].

PROPOSITION 2.5.2. *The following statements are equivalent:*

- (i) X is \mathcal{C}_0 -stationary.
- (ii) For all $C \in \mathcal{C}_0$ such that $\mathbf{m}(C) = 0$, $\Delta X_C = 0$ a.s. and for all $\phi \in \Phi(\mathcal{A})$, the distribution of $X^\phi - X_0^\phi$ only depends on $\nu(\phi)$.

If those conditions hold, then for all $\phi \in \Phi(\mathcal{A})$, the process X^ϕ has stationary increments in the sense of (2.5.1).

Proof. Suppose (i). Consider $C \in \mathcal{C}_0$ such that $\mathbf{m}(C) = 0$. Since $\mathbf{m}(C) = \mathbf{m}(\emptyset)$, then $\Delta X_C \stackrel{\text{law}}{=} X_\emptyset = 0$. Consider $\phi, \phi' \in \Phi(\mathcal{A})$ such that $\nu(\phi) = \nu(\phi')$ and let us show that $X^\phi - X_0^\phi \stackrel{\text{fdd}}{=} X^{\phi'} - X_0^{\phi'}$. Let $t_1 < \dots < t_k$ in $[0, 1]$, then

$$\forall j \in \llbracket 1, k \rrbracket, \quad \mathbf{m}(\phi(t_j) \setminus \phi(0)) = \nu(\phi) t_j = \nu(\phi') t_j = \mathbf{m}(\phi'(t_j) \setminus \phi'(0)).$$

Thus, since X is \mathcal{C}_0 -stationary,

$$(X_{t_j}^\phi - X_0^\phi)_{1 \leq j \leq k} = (\Delta X_{\phi(t_j) \setminus \phi(0)})_{1 \leq j \leq k} \stackrel{\text{law}}{=} (\Delta X_{\phi'(t_j) \setminus \phi'(0)})_{1 \leq j \leq k} = (X_{t_j}^{\phi'} - X_0^{\phi'})_{1 \leq j \leq k}.$$

Hence $X^\phi - X_0^\phi \stackrel{\text{fdd}}{=} X^{\phi'} - X_0^{\phi'}$ and (ii) follows.

Conversely, suppose (ii). Consider $A = A_0 \subseteq A_1 \subseteq \dots \subseteq A_k$. By Proposition 1.3.9, there exists $\phi \in \Phi(\mathcal{A})$ and $0 = t_0 \leq t_1 \leq \dots \leq t_k = 1$ such that for all $j \in \llbracket 0, k \rrbracket$, $A_j \subseteq \phi(t_j)$ and $\mathbf{m}(\phi(t_j) \setminus A_j) = 0$. By hypothesis, we know that for all $j \in \llbracket 0, k \rrbracket$, $X_{A_j} = X_{t_j}^\phi$. Hence

$$(\Delta X_{A_1 \setminus A}, \dots, \Delta X_{A_k \setminus A}) = (X_{t_1}^\phi - X_0^\phi, \dots, X_{t_k}^\phi - X_0^\phi). \quad (2.5.4)$$

Knowing that ϕ is geodesic, we must have $\nu(\phi) = \mathbf{m}(A_k \setminus A)$ and for all $j \in \llbracket 1, k \rrbracket$, $t_j = \mathbf{m}(A_j \setminus A) / \mathbf{m}(A_k \setminus A)$. In particular, if $A' \subseteq A'_1 \subseteq \dots \subseteq A'_k$ is another sequence in \mathcal{A} such that for all $j \in \llbracket 1, k \rrbracket$, $\mathbf{m}(A_j \setminus A) = \mathbf{m}(A'_j \setminus A')$, then there would exist $\phi' \in \Phi(\mathcal{A})$ such that

$$\nu(\phi) = \nu(\phi') \quad \text{and} \quad (\Delta X_{A'_1 \setminus A'}, \dots, \Delta X_{A'_k \setminus A'}) = (X_{t_1}^{\phi'} - X_0^{\phi'}, \dots, X_{t_k}^{\phi'} - X_0^{\phi'}). \quad (2.5.5)$$

Hence (i) follows from (2.5.4), (2.5.5) and (ii).

So (i) and (ii) are equivalent. Suppose that they both hold and consider $\phi \in \Phi(\mathcal{A})$. Let $\tau \in [0, 1]$ and show that $X^\phi - X_0^\phi|_{[0, 1-\tau]} \stackrel{\text{fdd}}{=} \{X_{t+\tau}^\phi - X_\tau^\phi : t \in [0, 1-\tau]\}$. If $\tau = 1$, then the result is trivial, so we may suppose that $\tau < 1$. Define for all $t \in [0, 1]$, $\phi'(t) = \phi((1-\tau)t)$ and $\phi'_\tau(t) = \phi((1-\tau)t + \tau)$. Both ϕ' and ϕ'_τ are geodesic \mathcal{A} -flows. Moreover, $\nu(\phi') = (1-\tau)\nu(\phi) = \nu(\phi'_\tau)$, which means by (ii) that $X^{\phi'} - X_0^{\phi'} \stackrel{\text{fdd}}{=} X^{\phi'_\tau} - X_0^{\phi'_\tau}$. In particular, we have for all $t_1, \dots, t_k \in [0, 1-\tau]$,

$$(X_{t_j}^\phi - X_0^\phi)_{1 \leq j \leq k} = (X_{t_j/(1-\tau)}^{\phi'} - X_0^{\phi'})_{1 \leq j \leq k} \stackrel{\text{fdd}}{=} (X_{t_j/(1-\tau)}^{\phi'_\tau} - X_0^{\phi'_\tau})_{1 \leq j \leq k} = (X_{t_j+\tau}^\phi - X_\tau^\phi)_{1 \leq j \leq k}.$$

Hence X^ϕ has stationary increments. \square

In the same way that the sifBm is a natural generalization of the fBm (since each of its projection on geodesic flows is a fBm), Proposition 2.5.2 tells us that \mathcal{C}_0 -stationarity is a natural generalization of the increment stationarity to the set-indexed setting.

An issue with this property is that it only concerns elements arranged in increasing order. In particular, no condition is imposed on the distribution of (X_{A_1}, X_{A_2}) when neither $A_1 \subseteq A_2$ nor $A_2 \subseteq A_1$. For set-indexed Lévy processes (siLévy), that is not a big issue since one may take advantage of the independence of increments to actually show [41, Corollary 4.5] that \mathcal{C}_0 -stationarity is equivalent to the stronger \mathcal{C} -increment stationarity (see Definition 2.5.12 later). For the sifBm, this remains an issue since its increments are far from being independent. However, we will circumvent this by showing that the sifBm verifies a stronger property that does not have this disadvantage.

EXAMPLE 2.5.3 (Why stationarity on $\mathcal{A}(u)$ is a dead end). *A first idea that comes to mind is to find a characterization based on $\mathcal{A}(u)$ -flows since we know that projections along those characterize the set-indexed process (Theorem 1.3.13). As mentioned before, this approach has been successfully used for instance in [73] to characterize the siBm.*

However, any such idea is doomed from the outset since it would imply a kind of 'weak $\mathcal{A}(u)$ -stationarity' given by

$$\forall U, V \in \mathcal{A}(u), \quad \mathbf{m}(U) = \mathbf{m}(V) \implies \Delta X_U \stackrel{\text{law}}{=} \Delta X_V. \quad (2.5.6)$$

Let us show that the sifBm cannot verify (2.5.6) in general. Suppose that $\mathcal{T} = \{a + ib \in \mathbf{C} : a, b \in [0, 1] \text{ and } a \times b = 0\}$ is endowed with the indexing collection \mathcal{A} given in Example 1.3.15 and $\mathbf{m} = \mathcal{H}^1(\mathcal{T} \cap \cdot)$ where \mathcal{H}^1 is the Hausdorff measure of order 1 on \mathbf{C} . Let $H \in (0, 1/2)$ and consider a sifBm W^H on \mathcal{A} . Let $U = [0, 1]$ and $V = [0, 1/2] \cup [0, i/2]$. Then,

$$\text{Var}(\Delta W_U^H) = \mathbf{E}[(W_U^H)^2] = \mathbf{m}(U)^{2H} = 1$$

while

$$\text{Var}(\Delta W_V^H) = \mathbf{E}[(W_{[0, 1/2]}^H + W_{[0, i/2]}^H - W_{\{0\}}^H)^2] = 2^{2(1-H)} - 1 \neq 1.$$

Hence, even though $U, V \in \mathcal{A}(u)$ are such that $\mathbf{m}(U) = 1 = \mathbf{m}(V)$, ΔW_U^H and ΔW_V^H do not have the same law. So (2.5.6) cannot hold for the sifBm.

In a nutshell, we want an increment stationarity property that may concern any finite-dimensional distribution of X and which is verified by the sifBm. What Example 2.5.3 explains is that this goal cannot be reached solely through $\mathcal{A}(u)$ -flows. Hence our idea to consider tuples of \mathcal{A} -flows instead, which yield more information than a single $\mathcal{A}(u)$ -flow.

DEFINITION 2.5.4 (Equality in configuration). *Let $k \in \mathbf{N}^*$ and $B_1, \dots, B_k, B'_1, \dots, B'_k \in \mathcal{B}$. We say that (B_1, \dots, B_k) and (B'_1, \dots, B'_k) are equal in configuration if*

$$\forall J \subseteq \llbracket 1, k \rrbracket, \quad m \left(\bigcap_{j \in J} B_j \right) = m \left(\bigcap_{j \in J} B'_j \right).$$

In that case, we write $(B_1, \dots, B_k) \stackrel{(c)}{=} (B'_1, \dots, B'_k)$.

If $\phi_1, \dots, \phi_k, \phi'_1, \dots, \phi'_k$ are \mathcal{B}_m -flows, we say that (ϕ_1, \dots, ϕ_k) and $(\phi'_1, \dots, \phi'_k)$ are equal in configuration if for all t_1, \dots, t_k in $[0, 1]$, $(\phi_1(t_1), \dots, \phi_k(t_k)) \stackrel{(c)}{=} (\phi'_1(t_1), \dots, \phi'_k(t_k))$. In that case, we write $(\phi_1, \dots, \phi_k) \stackrel{(c)}{=} (\phi'_1, \dots, \phi'_k)$.

REMARKS 2.5.5. \diamond Equality in configuration for (B_1, \dots, B_k) and (B'_1, \dots, B'_k) is a natural way to express the fact for those two tuples of sets to be ‘distributed’ in the same way. An equivalent statement that helps support this point of view is the following:

$$(B_1, \dots, B_k) \stackrel{(c)}{=} (B'_1, \dots, B'_k) \iff (\mathbb{1}_{B_1}, \dots, \mathbb{1}_{B_k})_* m = (\mathbb{1}_{B'_1}, \dots, \mathbb{1}_{B'_k})_* m. \quad (2.5.7)$$

Moreover, if the B_j 's are pairwise disjoint, we have

$$(B_1, \dots, B_k) \stackrel{(c)}{=} (B'_1, \dots, B'_k) \iff \forall j \in \llbracket 1, k \rrbracket, \quad m(B_j) = m(B'_j). \quad (2.5.8)$$

\diamond Equality in configuration is easily shown to be an equivalence relation, which enables us to talk about ‘the configuration’ of some tuple (B_1, \dots, B_k) or (ϕ_1, \dots, ϕ_k) when referring to its equivalence class.

DEFINITION 2.5.6 (\mathcal{A} -exchangeability). *The process X is \mathcal{A} -exchangeable if for all $k \in \mathbf{N}^*$, $A_0, \dots, A_k, A'_1, \dots, A'_k$ in \mathcal{A} ,*

$$(A_1, \dots, A_k) \stackrel{(c)}{=} (A'_1, \dots, A'_k) \implies (X_{A_1}, \dots, X_{A_k}) \stackrel{\text{law}}{=} (X_{A'_1}, \dots, X_{A'_k}).$$

DEFINITION 2.5.7 (\mathcal{A} -increment stationarity). *The process X is \mathcal{A} -increment stationary if for all $k \in \mathbf{N}^*$, $A, A_0, \dots, A_k, A', A'_1, \dots, A'_k$ in \mathcal{A} such that $A \subseteq \bigcap_{j=1}^k A_j$ and $A' \subseteq \bigcap_{j=1}^k A'_j$,*

$$(A_1 \setminus A, \dots, A_k \setminus A) \stackrel{(c)}{=} (A'_1 \setminus A', \dots, A'_k \setminus A') \implies (\Delta X_{A_1 \setminus A}, \dots, \Delta X_{A_k \setminus A}) \stackrel{\text{law}}{=} (\Delta X_{A'_1 \setminus A'}, \dots, \Delta X_{A'_k \setminus A'}).$$

Any \mathcal{A} -increment stationary process is obviously \mathcal{A} -exchangeable (take $A = A' = \emptyset$), but the latter will prove to be useful as well and more relevant when comparing to what already exists in the literature. In order to facilitate the comparison with \mathcal{C}_0 -stationarity, we prove the alter ego of Proposition 2.5.2 for \mathcal{A} -increment stationarity.

PROPOSITION 2.5.8. *Consider the following statements:*

(i) X is \mathcal{A} -increment stationary.

- (ii) For all $C \in \mathcal{C}_0$ such that $\mathbf{m}(C) = 0$, $\Delta X_C = 0$ a.s. and for all $k \in \mathbf{N}^*$ and $\phi = (\phi_1, \dots, \phi_k) \in \Phi(\mathcal{A})^k$ such that $\phi_1(0) = \dots = \phi_k(0)$, the distribution of the process $X^\phi - X_0^\phi = (X^{\phi_1} - X_0^{\phi_1}, \dots, X^{\phi_k} - X_0^{\phi_k})$ only depends on the configuration of $\phi \setminus \phi(0) = (\phi_1 \setminus \phi_1(0), \dots, \phi_k \setminus \phi_k(0))$.
- (iii) For all $k \in \mathbf{N}^*$ and $\phi = (\phi_1, \dots, \phi_k) \in \Phi(\mathcal{A})^k$ such that $\phi_1(0) = \dots = \phi_k(0)$, the multivariate process $X^\phi = (X^{\phi_1}, \dots, X^{\phi_k})$ has stationary increments in the sense of (2.5.1).

Then we have (i) \Rightarrow (ii) \Rightarrow (iii).

Remark that for all $\phi, \phi' \in \Phi(\mathcal{A})$, we have $\phi \setminus \phi(0) \stackrel{(c)}{=} \phi' \setminus \phi'(0)$ if and only if $\nu(\phi) = \nu(\phi')$, so Proposition 2.5.8 is in a sense a generalization of Proposition 2.5.2. The converse (i) \Leftarrow (ii) should also be true, but proving it requires a multivariate version of Proposition 1.3.9, which we could not show.

Proof of Proposition 2.5.8. Let us suppose (i). The fact that for all $C \in \mathcal{C}_0$, $\mathbf{m}(C) = 0$ implies $\Delta X_C = 0$ a.s. follows from Proposition 2.5.2. Consider $\phi = (\phi_1, \dots, \phi_k) \in \Phi(\mathcal{A})^k$ such that $\phi_1(0) = \dots = \phi_k(0)$ and $t_1, \dots, t_l \in [0, 1]$. Then, we have

$$(X_{t_j}^\phi - X_0^\phi)_{1 \leq j \leq l} = (\Delta X_{\phi_i(t_j) \setminus \phi_i(0)})_{\substack{1 \leq i \leq k \\ 1 \leq j \leq l}}$$

whose distribution, according to (i), only depends on the configuration of $(\phi_i(t_j) \setminus \phi_i(0))_{i,j}$. Hence the distribution of $X^\phi - X_0^\phi$ only depends on the configuration of $\phi \setminus \phi(0)$, which proves (ii).

Proving (ii) \Rightarrow (iii) only requires to copy the corresponding part in the proof of Proposition 2.5.2 since the ‘shift and dilatation trick’ that takes place also preserves the configuration. \square

PROPOSITION 2.5.9. For any $H \in (0, 1/2]$, the sifBm W^H with Hurst index H is \mathcal{A} -increment stationary.

Proof. Let $H \in (0, 1/2]$. Since W^H is a centered Gaussian process, we just need to prove that the quantity $\text{Cov}(\Delta W_{A_1 \setminus A}^H, \Delta W_{A_2 \setminus A}^H)$ depends only on the configuration of $(A_1 \setminus A, A_2 \setminus A)$ where $A, A_1, A_2 \in \mathcal{A}$ and $A \subseteq A_1 \cap A_2$. We have for such A, A_1, A_2 ,

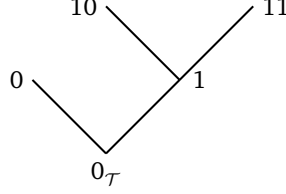
$$\begin{aligned} 2 \text{Cov}(\Delta W_{A_1 \setminus A}^H, \Delta W_{A_2 \setminus A}^H) &= 2 \mathbf{E}[(W_{A_1}^H - W_A^H)(W_{A_2}^H - W_A^H)] \\ &= \mathbf{m}(A_1 \Delta A)^{2H} + \mathbf{m}(A \Delta A_2)^{2H} - \mathbf{m}(A_1 \Delta A_2)^{2H} - \mathbf{m}(A \Delta A)^{2H} \end{aligned}$$

Since $\mathbf{m}(A_1 \Delta A_2) = \mathbf{m}((A_1 \setminus A) \Delta (A_2 \setminus A))$, we obtain

$$\text{Cov}(\Delta W_{A_1 \setminus A}^H, \Delta W_{A_2 \setminus A}^H) = \frac{1}{2} [\mathbf{m}(A_1 \setminus A)^{2H} + \mathbf{m}(A_2 \setminus A)^{2H} - \mathbf{m}((A_1 \setminus A) \Delta (A_2 \setminus A))^{2H}] \quad (2.5.9)$$

which indeed only depends on the configuration of $(A_1 \setminus A, A_2 \setminus A)$. \square

EXAMPLE 2.5.10 (\mathcal{C}_0 -stationarity and \mathcal{A} -exchangeability $\not\Rightarrow$ \mathcal{A} -increment stationarity). A natural question is whether or not saying that a process X is \mathcal{A} -increment stationary is equivalent to say that X is \mathcal{C}_0 -stationary and \mathcal{A} -exchangeable. The direct implication is obviously true. As for the converse, let us construct a counter-example. Consider that \mathcal{T} is the series gluing of $(\mathcal{T}_u)_{u \in \mathcal{U}}$ along the discrete tree $\mathcal{U} = \{0_{\mathcal{U}}, 0, 1, 10, 11\}$ where $\mathcal{T}_{0_{\mathcal{U}}} = \{0_{\mathcal{T}}\}$ and for all $u \in \mathcal{U} \setminus \{0_{\mathcal{U}}\}$, $\mathcal{T}_u = [0, 1]$. This is a special case of Example 1.2.25 and is illustrated in Figure 2.1.

Figure 2.1: Continuous tree \mathcal{T} obtained from $\mathcal{U} = \{0_{\mathcal{U}}, 0, 1, 10, 11\}$.

Consider the canonical projection $\pi : \bigsqcup_{u \in \mathcal{U}} \mathcal{T}_u \rightarrow \mathcal{T}$ just like in Proposition 1.2.23. Define $\mathbf{m} = \pi_* \mathbf{m}'$ where for all $u \in \mathcal{U}$, $\mathbf{m}'(\mathcal{T}_u \cap \cdot)$ is the usual Lebesgue measure on \mathcal{T}_u . Denote by \mathcal{A}_L (resp. \mathcal{A}_R) the indexing collection of $\pi(\mathcal{T}_0)$ (resp. $\pi(\mathcal{T}_1 \sqcup \mathcal{T}_{10} \sqcup \mathcal{T}_{11})$). Remark that the indexing collection \mathcal{A} of \mathcal{T} verifies $\mathcal{A} = \mathcal{A}_L \cup \mathcal{A}_R$ and $\mathcal{A}_L \cap \mathcal{A}_R = \{\emptyset, \emptyset'\}$.

Let $H \in (0, 1/2)$ and define $X = \{X_A : A \in \mathcal{A}\}$ such that $X|_{\mathcal{A}_L}$ and $X|_{\mathcal{A}_R}$ are two independent sifBm with Hurst index H .

A straightforward computation shows that for any $\phi \in \Phi(\mathcal{A})$, $v(\phi)^{-H}(X^\phi - X_0^\phi)$ is a $[0, 1]$ -indexed fBm. Thus, according to Proposition 2.5.2, X is \mathcal{C}_0 -stationary.

Let us show that X is also \mathcal{A} -stationary. Since it is a Gaussian process, we just need to consider $A_1, A_2, A'_1, A'_2 \in \mathcal{A}$ such that $(A_1, A_2) \stackrel{(c)}{=} (A'_1, A'_2)$ and show that $\text{Cov}(X_{A_1}, X_{A_2}) = \text{Cov}(X_{A'_1}, X_{A'_2})$.

- ◊ **1st case** — $(A_1, A_2) \in \emptyset$ where $\emptyset = (\mathcal{A}_R \times \mathcal{A}_L) \cup (\mathcal{A}_L \times \mathcal{A}_R)$, i.e. A_1 and A_2 belong to ‘opposite sides’ of the indexing collection. In particular, we know that X_{A_1} and X_{A_2} are independent, hence $\text{Cov}(X_{A_1}, X_{A_2}) = 0$. Moreover, one may easily check that

$$(A_1, A_2) \in \emptyset \iff \mathbf{m}(A_1 \cap A_2) = 0.$$

Hence, in this particular case, the condition that $(A_1, A_2) \in \emptyset$ depends solely on the configuration of (A_1, A_2) . Thus, since $(A_1, A_2) \stackrel{(c)}{=} (A'_1, A'_2)$,

$$(A_1, A_2) \in \emptyset \iff (A'_1, A'_2) \in \emptyset. \quad (2.5.10)$$

In particular, we also know that $X_{A'_1}$ and $X_{A'_2}$ are independent and $\text{Cov}(X_{A'_1}, X_{A'_2}) = 0$.

Hence $\text{Cov}(X_{A_1}, X_{A_2}) = \text{Cov}(X_{A'_1}, X_{A'_2})$.

- ◊ **2nd case** — $(A_1, A_2) \notin \emptyset$. Due to (2.5.10), $(A'_1, A'_2) \notin \emptyset$ as well. In particular, both (A_1, A_2) and (A'_1, A'_2) belong to $\mathcal{A}_R^2 \cup \mathcal{A}_L^2$. Since $X|_{\mathcal{A}_R}$ and $X|_{\mathcal{A}_L}$ are sifBm, we already know that $\text{Cov}(X_{A_1}, X_{A_2}) = \text{Cov}(X_{A'_1}, X_{A'_2})$ by Proposition 2.5.9.

Hence X is both \mathcal{C}_0 -stationary and \mathcal{A} -exchangeable. Let us show however that X cannot be \mathcal{A} -increment stationary. Denote $A_1 = [0_{\mathcal{T}}, 0]$, $A_2 = [0_{\mathcal{T}}, 1] = A'$, $A'_1 = [0_{\mathcal{T}}, 10]$ and $A'_2 = [0_{\mathcal{T}}, 11]$. Then, we have $(A_1, A_2) \stackrel{(c)}{=} (A'_1 \setminus A', A'_2 \setminus A')$ but $\text{Cov}(X_{A_1}, X_{A_2}) = 0$ and, according to (2.5.9), $\text{Cov}(\Delta X_{A'_1 \setminus A'}, \Delta X_{A'_2 \setminus A'}) = 1 - 2^{2H-1} \neq 0$.

2.5.3 \mathcal{C} -exchangeability

Alas, our approach in Section 2.5.2 to obtain a representation for stationary processes has not been successful so far. So we choose to take a different approach here. Namely, we study a

stronger stationarity condition for which it has been possible for us to find a representation theorem (Theorem 2.5.30). As before, this condition stems from a well-known one-dimensional one.

DEFINITION 2.5.11 (One-dimensional exchangeable increments, adapted from [55, Section 1.3]). A $[0, 1]$ -indexed process $Z = \{Z_t : t \in [0, 1]\}$ has exchangeable increments if for all $k \in \mathbf{N}^*$ and k pairwise disjoint intervals $I_j = (s_j, t_j]$ for $j \in \llbracket 1, k \rrbracket$ (resp. $I'_j = (s'_j, t'_j]$ for $j \in \llbracket 1, k \rrbracket$) of $[0, 1]$,

$$\forall j \in \llbracket 1, k \rrbracket, \quad t_k - s_k = t'_k - s'_k \implies (X_{t_j} - X_{s_j})_{1 \leq j \leq k} \stackrel{\text{law}}{=} (X_{t'_j} - X_{s'_j})_{1 \leq j \leq k}.$$

The definition given in [55, Section 1.3] is actually the following — weaker — one. A sequence $(Y_k)_{0 \leq k < n}$ where $n \in \mathbf{N} \cup \{\infty\}$ is *exchangeable* if its distribution is invariant under permutations, i.e.

$$\forall \sigma \in \mathfrak{S}(\mathbf{N} \cap [0, n)), \quad (Y_k)_{0 \leq k < n} \stackrel{\text{fdd}}{=} (Y_{\sigma(k)})_{0 \leq k < n}. \quad (2.5.11)$$

A $[0, 1]$ -indexed process $Z = \{Z_t : t \in [0, 1]\}$ is then said to have exchangeable increments if for all $n \in \mathbf{N}^*$, the sequence $(Z_{(k+1)/n} - Z_{k/n})_{0 \leq k < n}$ is exchangeable. However this definition is better suited for \mathbf{Q}_+ -indexed processes (we refer to the numerous related results in [55]). Fortunately, in the right-continuous case, [55, Theorem 1.15] tells they are equivalent. This is the reason why we will stick to the terminology of Definition 2.5.11.

DEFINITION 2.5.12 (\mathcal{D} -exchangeability). Let \mathcal{D} be a subset of $\mathcal{C}(u)$. The process X is \mathcal{D} -exchangeable if for all $k \in \mathbf{N}^*$ and $D_1, \dots, D_k, D'_1, \dots, D'_k \in \mathcal{D}$,

$$(D_1, \dots, D_k) \stackrel{(c)}{=} (D'_1, \dots, D'_k) \implies (\Delta X_{D_1}, \dots, \Delta X_{D_k}) \stackrel{\text{law}}{=} (\Delta X_{D'_1}, \dots, \Delta X_{D'_k}).$$

Likewise, X is \mathfrak{B}_m -exchangeable if X extends to a stochastic measure $\Delta X = \{\Delta X_B : B \in \mathfrak{B}_m\}$ and for all $k \in \mathbf{N}^*$ and $B_1, \dots, B_k, B'_1, \dots, B'_k \in \mathfrak{B}_m$,

$$(B_1, \dots, B_k) \stackrel{(c)}{=} (B'_1, \dots, B'_k) \implies (\Delta X_{B_1}, \dots, \Delta X_{B_k}) \stackrel{\text{law}}{=} (\Delta X_{B'_1}, \dots, \Delta X_{B'_k}).$$

When $\mathcal{T} = [0, 1]$ and m is the Lebesgue measure, being \mathcal{D} -exchangeable for any $\mathcal{D} \in \{\mathcal{C}, \mathcal{C}(u), \mathfrak{B}_m\}$ is equivalent to have exchangeable increments (see Proposition 2.5.13).

The next result should be compared with Proposition 2.5.8. A consequence is that as long as $\mathcal{D} \supseteq \mathcal{C}$, \mathcal{D} -increment stationarity (defined in the same way as in Definition 2.5.7 where \mathcal{A} is replaced by \mathcal{D}) is equivalent to \mathcal{D} -exchangeability, which is the reason why we did not introduce this notion here contrary to the previous section.

PROPOSITION 2.5.13. Consider the following statements:

- (i) X is \mathcal{C} -exchangeable.
- (ii) X is $\mathcal{C}(u)$ -exchangeable.
- (iii) For all $C \in \mathcal{C}$ such that $m(C) = 0$, $\Delta X_C = 0$ a.s. and for all $\phi \in \Phi(\mathcal{A}(u))$, the process ΔX^ϕ has exchangeable increments and its distribution only depends on $\nu(\phi)$.
- (iv) X is \mathfrak{B}_m -exchangeable.

Then we have (i) \Leftrightarrow (ii) \Leftrightarrow (iii) and if X extends to a stochastic measure, all of them are equivalent.

Before proceeding to the proof of Proposition 2.5.13, we need to prove a technical lemma.

LEMMA 2.5.14. *Let $U, U' \in \mathcal{C}(u)$ such that $\mathbf{m}(U) = \mathbf{m}(U')$. There exists $k \in \mathbf{N}^*$ and $C_1, \dots, C_k, C'_1, \dots, C'_k$ in \mathcal{C} such that $U = \bigsqcup_{j=1}^k C_j$, $U' = \bigsqcup_{j=1}^k C'_j$ and for all $j \in \llbracket 1, k \rrbracket$, $\mathbf{m}(C_j) = \mathbf{m}(C'_j)$.*

Proof. We proceed by induction. We initiate with the following:

$$U = U_0 = \bigsqcup_{j=1}^{k_0} C_{0j}, \quad \mathcal{L}_0 = \emptyset, \quad U' = U'_0 = \bigsqcup_{j=1}^{k'_0} C'_{0j} \quad \text{and} \quad \mathcal{L}'_0 = \emptyset$$

for some $k_0, k'_0 \in \mathbf{N}$ and C_{01}, \dots, C_{0k_0} (resp. $C'_{01}, \dots, C'_{0k'_0}$) pairwise disjoint in \mathcal{C} . Remark that if $k_0 k'_0 = 0$, then $\mathbf{m}(U) = \mathbf{m}(U') = 0$ and the result is trivial because the representation in Lemma 2.5.14 holds with

$$\forall j \in \llbracket 1, k_0 \vee k'_0 \rrbracket, \quad C_j = \begin{cases} C_{0j} & \text{if } j \leq k_0, \\ \emptyset & \text{if } j > k_0 \end{cases} \quad \text{and} \quad C'_j = \begin{cases} C'_{0j} & \text{if } j \leq k'_0, \\ \emptyset & \text{if } j > k'_0 \end{cases}$$

Now, suppose that for some $i \in \mathbf{N}$, there is $(k_i, k'_i) \in \mathbf{N}^2$ such that the following predicate holds:

$$\mathcal{H}_i : \quad \text{"} U_i = \bigsqcup_{j=1}^{k_i} C_{ij}, \quad \mathcal{L}_i = (C_1, \dots, C_i), \quad U'_i = \bigsqcup_{j=1}^{k'_i} C'_{ij} \quad \text{and} \quad \mathcal{L}'_i = (C'_1, \dots, C'_i) \text{"}$$

where the C_{ij} 's and C_i 's (resp. C'_{ij} 's and C'_i 's) are pairwise disjoint elements of \mathcal{C} and

$$U = U_i \sqcup \bigsqcup_{C \in \mathcal{L}_i} C, \quad U' = U'_i \sqcup \bigsqcup_{C' \in \mathcal{L}'_i} C' \quad \text{and} \quad \forall j \in \llbracket 1, i \rrbracket, \quad \mathbf{m}(C_j) = \mathbf{m}(C'_j)."$$

Now suppose that $k_i k'_i \neq 0$ and $\mathbf{m}(C_{ik_i}) \leq \mathbf{m}(C'_{ik'_i})$. Then denote $C_{i+1} = C_{ik_i}$. By Proposition 1.3.11, we may find $C'_{i+1} \in \mathcal{C}$ such that $C'_{i+1} \subseteq C'_{ik'_i}$, $C'_{ik'_i} \setminus C'_{i+1} \in \mathcal{C}$ and $\mathbf{m}(C'_{i+1}) = \mathbf{m}(C_{i+1})$. Then, denote $k_{i+1} = k_i - 1$, $k'_{i+1} = k'_i$, and

$$\forall j \in \llbracket 1, k_{i+1} \rrbracket, \quad C_{i+1j} = C_{ij} \quad \text{and} \quad \forall j \in \llbracket 1, k'_{i+1} \rrbracket, \quad C'_{i+1j} = \begin{cases} C'_{ij} & \text{if } j < k'_i, \\ C'_{ik'_i} \setminus C'_{i+1} & \text{if } j = k'_i. \end{cases}$$

If $\mathbf{m}(C_{ik_i}) \geq \mathbf{m}(C'_{ik'_i})$, then we would have defined $C'_{i+1} = C_{ik'_i}$ and applied Proposition 1.3.11 to C_{ik_i} instead. In any case, we constructed a sequence $(k_i, k'_i)_i$ such that

$$\left(\mathcal{H}_i \quad \text{and} \quad k_i k'_i \neq 0 \right) \implies \left(\mathcal{H}_{i+1} \quad \text{and} \quad k_{i+1} + k'_{i+1} < k_i + k'_i \right).$$

Hence the induction terminates at some $l \in \mathbf{N}$ such that $k_l k'_l = 0$ and we are back at the case examined at the beginning of the proof. The result follows. \square

Proof of Proposition 2.5.13. We first prove (i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i).

Suppose (i). Consider $U_1, \dots, U_k, U'_1, \dots, U'_k$ in $\mathcal{C}(u)$ such that $(U_1, \dots, U_k) \stackrel{\text{(i)}}{\sim} (U'_1, \dots, U'_k)$. Instead of making intersections and set differences, we might as well suppose that the U_i 's are pairwise disjoint. In particular, for all $i \neq j$, $\mathbf{m}(U'_i \cap U'_j) = 0$ by equality in configuration. For $j \in \llbracket 1, k \rrbracket$, denote $V_j = U'_j \setminus \bigcup_{i \neq j} U'_i$. Since $V_j \setminus U'_j$ (resp. $U'_j \setminus V_j$) is an element of $\mathcal{C}(u)$ of \mathbf{m} -measure zero, we may write it as a disjoint union of elements in \mathcal{C} of \mathbf{m} -measure zero.

Since X is \mathcal{C} -exchangeable, we know that $\Delta X_{V_j \setminus U'_j} = 0$ a.s. (resp. $\Delta X_{U'_j \setminus V_j} = 0$ a.s.). Thus $\Delta X_{V_j} = \Delta X_{U'_j}$ a.s.. So without loss of generality, we may also suppose that the U'_i 's are pairwise disjoint. To summarize, we now have $U_1, \dots, U_k, U'_1, \dots, U'_k$ in $\mathcal{C}(u)$ such that U_1, \dots, U_k (resp. U'_1, \dots, U'_k) are pairwise disjoint and for all $j \in \llbracket 1, k \rrbracket$, $\mathbf{m}(U_j) = \mathbf{m}(U'_j)$ and we want to prove that $(\Delta X_{U_1}, \dots, \Delta X_{U_k}) \stackrel{\text{law}}{=} (\Delta X_{U'_1}, \dots, \Delta X_{U'_k})$. This result readily follows from Lemma 2.5.14 and \mathcal{C} -exchangeability. Hence (ii) holds.

Suppose (ii). The ‘separability’ property is proven as in Proposition 2.5.8. Consider $\phi \in \Phi(\mathcal{A}(u))$, $k \in \mathbb{N}^*$ and k pairwise disjoint intervals $I_j = (s_j, t_j]$ of $[0, 1]$ for $j \in \llbracket 1, k \rrbracket$. Since X is $\mathcal{C}(u)$ -exchangeable and the $(\phi(t_j) \setminus \phi(s_j))_{1 \leq j \leq k}$ are pairwise disjoint elements of $\mathcal{C}(u)$ of respective \mathbf{m} -measure $\nu(\phi)|t_j - s_j|$, the distribution of $(X_{t_j} - X_{s_j})_{1 \leq j \leq k}$ only depends on $\nu(\phi)$ and $(t_j - s_j)_{1 \leq j \leq k}$. Hence (iii) holds.

Suppose (iii). Consider $C_1, \dots, C_k, C'_1, \dots, C'_k \in \mathcal{C}$ such that $(C_1, \dots, C_k) \stackrel{(c)}{=} (C'_1, \dots, C'_k)$. Similarly to the proof of (i) \Rightarrow (ii), we might as well suppose that the C_j 's (resp. C'_j 's) are pairwise disjoint. Then, define \mathcal{A} as the collection of all intersections made from elements of \mathcal{A} used in the extremal representations of the C_j 's and order them consistently, i.e. $\mathcal{A} = \{A_1, \dots, A_l\}$ where for all $i, j \in \llbracket 1, l \rrbracket$, $A_i \subseteq A_j$ implies $i \leq j$. For all $i \in \llbracket 1, l \rrbracket$, denote $U_i = \bigcup_{j=1}^i A_j$. By Proposition 1.3.12, there exists $\phi \in \Phi(\mathcal{A}(u))$ and $0 = t_1 \leq \dots \leq t_l \leq 1$ such that for all $i \in \llbracket 1, l \rrbracket$, $U_i \subseteq \phi(t_i)$ and $\mathbf{m}(\phi(t_i) \setminus U_i) = 0$. Similarly to the proof of Proposition 2.5.8, we may show that

$$(\Delta X_{U_1}, \dots, \Delta X_{U_l}) = (\Delta X_{t_1}^\phi, \dots, \Delta X_{t_l}^\phi) \quad \text{a.s.} \quad (2.5.12)$$

Since \mathcal{A} has been consistently ordered, we have

$$\mathcal{C}^\ell(\mathcal{A}) = \{U_i \setminus U_{i-1} : 1 \leq i \leq l\} \quad \text{with } U_0 = \emptyset.$$

Thus, by definition of $\mathcal{C}^\ell(\mathcal{A})$, we know that there exists a family $(\mathcal{I}_j)_{1 \leq j \leq k}$ of pairwise disjoint subsets of $\llbracket 1, l \rrbracket$ such that

$$\forall j \in \llbracket 1, k \rrbracket, \quad C_j = \bigsqcup_{i \in \mathcal{I}_j} (U_i \setminus U_{i-1}). \quad (2.5.13)$$

Combining (2.5.12) and (2.5.13), we get

$$(\Delta X_{C_j})_{1 \leq j \leq k} = \left(\sum_{i \in \mathcal{I}_j} (\Delta X_{t_i}^\phi - \Delta X_{t_{i-1}}^\phi) \right)_{1 \leq j \leq k} \quad \text{a.s.} \quad (2.5.14)$$

Define the sequence $(s_j)_{0 \leq j \leq k}$ by

$$s_0 = 0 \quad \text{and} \quad \forall j \in \llbracket 1, k \rrbracket, \quad s_j - s_{j-1} = \sum_{i \in \mathcal{I}_j} (t_i - t_{i-1}). \quad (2.5.15)$$

Since ΔX^ϕ has exchangeable increments, we may deduce from (2.5.14) and (2.5.15) that

$$(\Delta X_{C_j})_{1 \leq j \leq k} \stackrel{\text{law}}{=} (\Delta X_{s_j}^\phi - \Delta X_{s_{j-1}}^\phi)_{1 \leq j \leq k} \quad (2.5.16)$$

Then define a new geodesic $\mathcal{A}(u)$ -flow ψ for all $t \in [0, 1]$ by $\psi(t) = \phi(s_k t)$. Combining (2.5.13) and (2.5.15), we may rewrite (2.5.14) to get

$$(\Delta X_{C_j})_{1 \leq j \leq k} \stackrel{\text{law}}{=} (\Delta X_{u_j}^\psi - \Delta X_{u_{j-1}}^\psi)_{1 \leq j \leq k} \quad (2.5.17)$$

where

$$v(\psi) = \mathbf{m} \left(\bigsqcup_{j=1}^k C_j \right), \quad u_0 = 0 \quad \text{and} \quad \forall j \in \llbracket 1, k \rrbracket, \quad u_j - u_{j-1} = \mathbf{m}(C_j)/v(\psi). \quad (2.5.18)$$

In particular, (2.5.18) shows that the representation (2.5.17) only depends on the configuration of (C_1, \dots, C_k) . Thus, we get

$$\left(\Delta X_{C_j} \right)_{1 \leq j \leq k} \stackrel{\text{law}}{=} \left(\Delta X_{C'_j} \right)_{1 \leq j \leq k}$$

from which (i) follows.

Suppose that X extends to a stochastic measure $\Delta \mathbf{X}$. The implication (iv) \Rightarrow (ii) is trivial since $\mathcal{C}(u) \subseteq \mathcal{B}_m$. The converse (ii) \Rightarrow (iv) is just a consequence of the fact that any $\Delta \mathbf{X}_B$ for $B \in \mathcal{B}_m$ is the limit in probability of a sequence $(\Delta X_{U_k})_{k \in \mathbf{N}}$ where $(U_k)_{k \in \mathbf{N}} \in \mathcal{C}(u)^{\mathbf{N}}$. \square

2.5.4 Functional exchangeability

We saw that a \mathcal{C} -exchangeable process is also \mathcal{B}_m -exchangeable provided that it extends to a stochastic measure (Proposition 2.5.13). We also know that if a process extends to a stochastic measure, then it has an associated linear process (Theorem 2.2.10). This section just states the stationarity property that such a linear process inherits.

DEFINITION 2.5.15 (Functional exchangeability). *The process \mathcal{X} is \mathcal{E} -exchangeable if for all $k \in \mathbf{N}^*$ and simple functions $f_1, \dots, f_k, f'_1, \dots, f'_k$ in \mathcal{E} ,*

$$(f_1, \dots, f_k)_* \mathbf{m} = (f'_1, \dots, f'_k)_* \mathbf{m} \implies (\mathcal{X}(f_1), \dots, \mathcal{X}(f_k)) \stackrel{\text{law}}{=} (\mathcal{X}(f'_1), \dots, \mathcal{X}(f'_k)).$$

The process \mathcal{X} is $L(X)$ -exchangeable if it extends to a linear process $\mathcal{X} = \{\mathcal{X}(f) : f \in L(X)\}$ and for all $k \in \mathbf{N}^$ and $f_1, \dots, f_k, f'_1, \dots, f'_k$ in $L(X)$,*

$$(f_1, \dots, f_k)_* \mathbf{m} = (f'_1, \dots, f'_k)_* \mathbf{m} \implies (\mathcal{X}(f_1), \dots, \mathcal{X}(f_k)) \stackrel{\text{law}}{=} (\mathcal{X}(f'_1), \dots, \mathcal{X}(f'_k)).$$

PROPOSITION 2.5.16. *Suppose that X extends to a linear process. Consider the following statements:*

- (i) X is \mathcal{C} -exchangeable.
- (ii) \mathcal{X} is \mathcal{E} -exchangeable.
- (iii) \mathcal{X} is $L(X)$ -exchangeable.

Then we have (i) \Leftrightarrow (ii) and if X extends to a stochastic measure, all of them are equivalent.

Proof. Suppose (i). Consider $k \in \mathbf{N}^*$, $f = (f_1, \dots, f_k)$ and $f' = (f'_1, \dots, f'_k)$ in \mathcal{E}^k such that $f_* \mathbf{m} = f'_* \mathbf{m}$. Consider the finite subset of \mathbf{R}^k given by $E = \{a \in \mathbf{R}^k : \mathbf{m}(f = a) > 0\}$. Since $f_* \mathbf{m} = f'_* \mathbf{m}$, we also have $E = \{a \in \mathbf{R}^k : \mathbf{m}(f' = a) > 0\}$ and

$$f = \sum_{a \in E} a \mathbb{1}_{\{f=a\}} \quad \text{and} \quad f' = \sum_{a \in E} a \mathbb{1}_{\{f'=a\}} \quad \mathbf{m}\text{-a.e.}$$

where $(\{f = a\})_{a \in E}$ (resp. $(\{f' = a\})_{a \in E}$) are pairwise disjoint sets such that for all $a \in E$, $\mathbf{m}(f = a) = \mathbf{m}(f' = a)$. Thus $(\{f = a\})_{a \in E} \stackrel{(c)}{=} (\{f' = a\})_{a \in E}$. Since f and f' are simple

functions, we also know that those sets belong to $\mathcal{C}(u)$. Hence, since X is $\mathcal{C}(u)$ -exchangeable (Proposition 2.5.13), we get $\mathcal{X}(f) \stackrel{\text{law}}{=} \mathcal{X}(f')$ from which (ii) follows. The converse (ii) \Rightarrow (i) is trivial.

Now, suppose that X extends to a stochastic measure and (i). Consider $k \in \mathbf{N}^*$, $f = (f_1, \dots, f_k)$ and $f' = (f'_1, \dots, f'_k)$ in $L(X)^k$ such that $f_*\mathbf{m} = f'_*\mathbf{m}$. We denote for all $n \in \mathbf{N}$,

$$f_n = \sum_{k=-n2^n}^{n2^n} k2^{-n} \mathbb{1}_{\{k2^{-n} \leq f < (k+1)2^{-n}\}} \quad \text{and} \quad f'_n = \sum_{k=-n2^n}^{n2^n} k2^{-n} \mathbb{1}_{\{k2^{-n} \leq f' < (k+1)2^{-n}\}}.$$

Since $f_*\mathbf{m} = f'_*\mathbf{m}$, we know that for all $n \in \mathbf{N}$, $(f_n)_*\mathbf{m} = (f'_n)_*\mathbf{m}$. By Proposition 2.5.13, we know that X is \mathcal{B}_m -exchangeable. Mimicking the proof of (i) \Rightarrow (ii), we show that for all $n \in \mathbf{N}$, $\mathcal{X}(f_n) \stackrel{\text{law}}{=} \mathcal{X}(f'_n)$. Taking $n \rightarrow \infty$ gives $\mathcal{X}(f) \stackrel{\text{law}}{=} \mathcal{X}(f')$, from which (iii) follows. For the same reason as before, the implication (iii) \Rightarrow (ii) is trivial. \square

EXAMPLE 2.5.17 (Series representation for $L^2(0, 1)$ -exchangeable processes). *As a quick illustration and trailer for the next section, let us see what happens for such processes when $\mathcal{T} = [0, 1]$ and \mathbf{m} is the Lebesgue measure. Define for all $n \in \mathbf{N}$ the function*

$$\begin{aligned} e_n : [0, 1] &\longrightarrow \mathbf{R} \\ t &\longmapsto \sum_{k=0}^{2^n} (-1)^{k+1} \mathbb{1}_{k2^{-n} \leq t < (k+1)2^{-n}}. \end{aligned}$$

It is easy to see that under \mathbf{m} , $(e_n)_{n \in \mathbf{N}}$ is an iid sequence of Rademacher variables. In particular, if X is \mathcal{C} -exchangeable, we know by Proposition 2.5.16 that the sequence $(\mathcal{X}(e_n))_{n \in \mathbf{N}}$ is exchangeable and hence is conditionally iid by de Finetti's theorem (see e.g. [55, Theorem 1.1]).

In particular, if X has an associated linear process \mathcal{X} on $L^2(0, 1)$, we have the series representation

$$\forall f \in L^2(0, 1), \quad \mathcal{X}(f) = \sum_{n=0}^{\infty} \langle f, e_n \rangle \mathcal{X}(e_n)$$

where the sum converges almost surely (it is a sum of conditionally independent variables that converges in probability).

This representation could be extended to a much more general setting as long as one possesses an appropriate 'Haar basis' $(e_n)_{n \in \mathbf{N}}$.

2.5.5 Examples of \mathcal{C} -exchangeable processes

Due to Proposition 2.5.13 and Example 2.5.3, it is clear that the sifBm is not \mathcal{C} -exchangeable in general. However, there are interesting stochastic processes that do verify this property.

Due to [41, Corollary 4.5], a siLévy X is such that for all $C, C' \in \mathcal{C}$ such that $\mathbf{m}(C) = \mathbf{m}(C')$, $\Delta X_C \stackrel{\text{law}}{=} \Delta X_{C'}$. Using the independence of \mathcal{C} -increments, we readily deduce that X is \mathcal{C} -exchangeable.

Since a siLévy has independent increments, it also has a Lévy-Khintchine representation. Using the \mathcal{C} -exchangeability, we may deduce that the triplet in Corollary 2.3.16 takes the form $(b\mathbf{m}, \sigma^2\mathbf{m}, \mathbf{m} \otimes \nu)$ where $b \in \mathbf{R}$, $\sigma^2 \in \mathbf{R}_+$ and ν is a Lévy measure on \mathbf{R} . This computation has been carried out in greater details in [41, Equation (15)].

However, there is no reason in general for a \mathcal{C} -exchangeable process to have independent increments. And a neat way to generate some is by mixing siLévy's. In general, the resulting process does not have independent increments as the examples below will show.

EXAMPLE 2.5.18 (Random drift). For any real-valued random variable $b \in L^0(\Omega)$, the process $\{b \mathbf{m}(A) : A \in \mathcal{A}\}$ is \mathcal{C} -exchangeable.

EXAMPLE 2.5.19 (Set-indexed Brownian bridge). Suppose that $0 < \mathbf{m}(\mathcal{T}) < \infty$. The set-indexed Brownian bridge (siBb) is a centered Gaussian process $\widehat{W} = \{\widehat{W}_A : A \in \mathcal{A}\}$ with covariance function given by

$$\forall A, A' \in \mathcal{A}, \quad \text{Cov}(\widehat{W}_A, \widehat{W}_{A'}) = \mathbf{m}(A \cap A') - \frac{\mathbf{m}(A)\mathbf{m}(A')}{\mathbf{m}(\mathcal{T})}. \quad (2.5.19)$$

It is straightforward to check that if W is a siBm, then it is possible to consider a siBb as a siBm with a random drift as follows:

$$\forall A \in \mathcal{A}, \quad \widehat{W}_A = W_A - \frac{\mathbf{m}(A)}{\mathbf{m}(\mathcal{T})} \Delta W_{\mathcal{T}} \quad (2.5.20)$$

where ΔW is the stochastic measure defined on \mathfrak{B}_m which extends W . Its name comes from the fact that for any geodesic flow ϕ such that $\mathbf{m}(\phi(0)) = 0$ and $\phi(1) = \mathcal{T}$, the process $v(\phi)^{-1/2} \Delta \widehat{W}^\phi$ is the usual $[0, 1]$ -indexed Brownian bridge. A similar — but not equivalent — process may be found in [4, Equation (4.3.8)] under the name set-indexed pinned Brownian sheet.

Let us check that the siBb is \mathcal{C} -exchangeable. Using the representation formula (2.5.20), we get for all $C, C' \in \mathcal{C}$,

$$\text{Cov}(\Delta \widehat{W}_C, \Delta \widehat{W}_{C'}) = \mathbf{m}(C \cap C') - \frac{\mathbf{m}(C)\mathbf{m}(C')}{\mathbf{m}(\mathcal{T})}$$

which only depends on the configuration of (C, C') . The \mathcal{C} -exchangeability of the siBb follows.

In particular, a set-indexed Poisson process of intensity \mathbf{m} is a siLévy, and thus \mathcal{C} -exchangeable. However, a \mathcal{C} -exchangeable process has no reason whatsoever to have independent increments, so a mixture of Poisson processes (whose existence is once more ensured by [54, Theorem 12.7]) should also be considered.

DEFINITION 2.5.20 (Set-indexed Cox process). Let ν be a random measure on \mathcal{T} . A set-indexed Cox process (siCox) of intensity ν is a point process $N = \{N_A : A \in \mathcal{A}\}$ such that, conditionally on ν , N is a set-indexed Poisson process of intensity ν .

EXAMPLE 2.5.21 (Set-indexed compound Cox process). If instead in Example 2.4.6, ν was a random measure on \mathcal{T} such that $\nu(\{0\}) = 0$ and $\int_{\mathbf{R}} (1 \wedge |x|) \nu(dx) < \infty$ or $\int_{\mathbf{R}} (1 \wedge x^2) \nu(dx) < \infty$ a.s., then putting a set-indexed Cox process of intensity $\mathbf{m} \otimes \nu$ in (2.4.1) (resp. (2.4.2), (2.4.3)) instead of N would define a set-indexed compound Cox process (resp. compensated Cox process, compensated compound Cox process) of intensity ν .

PROPOSITION 2.5.22. All the processes defined in Examples 2.4.6 and 2.5.21 are \mathcal{C} -exchangeable.

Proof. The rest being similar, we only prove it for the compound Poisson process P in (2.4.1). Denote by \mathcal{N} the linear process associated to N . We have

$$\forall C \in \mathcal{C}, \quad \Delta P_C = \mathcal{N}(f_C) \quad \text{where} \quad f_C : (t, x) \mapsto x \mathbf{1}_C(t).$$

Consider $C_1, \dots, C_k, C'_1, \dots, C'_k \in \mathcal{C}$ such that $(C_1, \dots, C_k) \stackrel{(c)}{=} (C'_1, \dots, C'_k)$. Due to the definition of f_C , we get

$$(f_{C_1}, \dots, f_{C_k})_* (\mathbf{m} \otimes \nu) = (f_{C'_1}, \dots, f_{C'_k})_* (\mathbf{m} \otimes \nu).$$

By Proposition 2.5.16, this actually implies that

$$(\Delta P_{C_1}, \dots, \Delta P_{C_k}) \stackrel{\text{law}}{=} (\Delta P_{C'_1}, \dots, \Delta P_{C'_k}).$$

Thus P is \mathcal{C} -exchangeable. \square

One could be tempted to think that the jump discontinuities of a \mathcal{C} -exchangeable process may always be expressed in terms of Cox processes as in Example 2.5.21. However, that is not true when $0 < \mathbf{m}(\mathcal{T}) < \infty$ where there is another way to define a compensated point process.

EXAMPLE 2.5.23 (Another compensated point process). *Suppose that $0 < \mathbf{m}(\mathcal{T}) < \infty$ and consider a random measure $J = \sum_{j=0}^{\infty} \delta_{J_j}$ on \mathbf{R}^* such that*

$$\int_{\mathbf{R}} x^2 J(dx) = \sum_{j=0}^{\infty} J_j^2 < \infty \quad \text{a.s.}$$

as well as a sequence $(\tau_j)_{j \in \mathbf{N}}$ of iid random variables on \mathcal{T} with distribution $\mathbf{m}(\cdot)/\mathbf{m}(\mathcal{T})$ and independent from J . Define the set-indexed process \tilde{M} by

$$\forall A' \in \mathcal{A} \times \mathcal{A}_{\mathbf{R}^* \cup \{\infty\}}, \quad \tilde{M}_{A'} = \sum_{j=0}^{\infty} \left(\left(\delta_{\tau_j} - \frac{\mathbf{m}}{\mathbf{m}(\mathcal{T})} \right) \otimes \delta_{J_j} \right) (A') \quad (2.5.21)$$

where $\mathcal{A}_{\mathbf{R}^* \cup \{\infty\}}$ is the indexing collection from Example 1.2.19 and the sum is well-defined since it is almost surely finite for all A' .

Let us define the set-indexed compensated Kallenberg point process (sicK) of intensity J by

$$\forall A \in \mathcal{A}, \quad \tilde{Q}_A = \int_{\mathcal{T} \times (\mathbf{R}^* \cup \{\infty\})} [x \mathbb{1}_A(t)] \tilde{M}(dt, dx) = \int_{A \times \mathbf{R}} x \tilde{M}(dt, dx). \quad (2.5.22)$$

where the name follows from a terminology suggested by Kingman in [57, Section 7].

Let us show that \tilde{Q} is a well-defined process. In the following, everything is conditioned with respect to J so that it may be treated as a constant. A direct application of Theorem 2.2.8 shows that \tilde{M} extends to a stochastic measure. Denote by $\tilde{\mathcal{M}}$ the associated linear process. Consider $A \in \mathcal{A}$ and denote $f : (t, x) \mapsto x \mathbb{1}_A(t)$. We only need to prove that $f \in L(\tilde{M})$. Introduce the following discretizations:

$$\forall k \in \mathbf{N}, \forall x \in \mathbf{R}, \quad x^{(k)} = \frac{\lfloor 2^k x \rfloor}{2^k} \wedge k \vee (-k) \quad \text{and} \quad f^{(k)} : (t, x) \mapsto x^{(k)} \mathbb{1}_A(t).$$

The functions $f^{(k)}$ are simple and converge pointwise to f as $k \rightarrow \infty$. Thus, by Definition 2.2.9, we need to show that

$$\rho_{\tilde{M}}(f^{(k)} - f^{(l)}) = \sup_g \mathbf{E} \left[\left| \tilde{\mathcal{M}}((f^{(k)} - f^{(l)})g) \right| \wedge 1 \right] \rightarrow 0 \quad \text{as} \quad k, l \rightarrow \infty \quad (2.5.23)$$

where $\rho_{\tilde{M}}$ is the modular induced by \tilde{M} defined in (2.2.3) and the supremum is over all simple functions g on $\mathcal{T} \times (\mathbf{R}^* \cup \{\infty\})$. Consider such a function g . Using Proposition 1.2.42, we write a \mathcal{C} -representation of g as follows

$$g = \sum_{i=1}^n a_i \mathbb{1}_{C_i \times I_i}$$

where for all i , $|a_i| \leq 1$, $C_i \in \mathcal{C}$, I_i is an interval included in \mathbf{R}^* (we may drop the ∞ since $f_A^{(k)}(\infty) = 0$) and the $C_i \times I_i$'s are pairwise disjoint.

Denoting by \tilde{M}_j the j^{th} term in the sum (2.5.21) and $\tilde{\mathcal{M}}_j$ the associated linear process, we have

$$\tilde{\mathcal{M}}(f^{(k,l)}g) = \sum_{j=0}^{\infty} \tilde{\mathcal{M}}_j(f^{(k,l)}g)$$

where $f^{(k,l)} = f^{(k)} - f^{(l)}$ and the terms are independent since we conditioned with respect to J . Hence

$$\text{Var}\left(\tilde{\mathcal{M}}(f^{(k,l)}g)\right) = \sum_{j=0}^{\infty} \sum_{1 \leq i_1, i_2 \leq n} a_{i_1} a_{i_2} \text{Cov}\left(\tilde{\mathcal{M}}_j(f^{(k,l)}\mathbb{1}_{C_{i_1} \times I_{i_1}}), \tilde{\mathcal{M}}_j(f^{(k,l)}\mathbb{1}_{C_{i_2} \times I_{i_2}})\right). \quad (2.5.24)$$

A tedious but straightforward computation yields for all i_1, i_2 ,

$$\text{Cov}\left(\tilde{\mathcal{M}}_j(f^{(k,l)}\mathbb{1}_{C_{i_1} \times I_{i_1}}), \tilde{\mathcal{M}}_j(f^{(k,l)}\mathbb{1}_{C_{i_2} \times I_{i_2}})\right) \leq (J_j^{(k,l)})^2 \mathbb{1}_{J_j^{(k,l)} \in I_{i_1}} \times \frac{\mathbf{m}(A \cap C_{i_1})}{\mathbf{m}(\mathcal{T})} \mathbb{1}_{i_1=i_2} \quad (2.5.25)$$

where $J_j^{(k,l)} = J_j^{(k)} - J_j^{(l)}$. Putting (2.5.25) back into (2.5.24) yields

$$\text{Var}\left(\tilde{\mathcal{M}}(f^{(k,l)}g)\right) \leq \sum_{j=0}^{\infty} \sum_{i=1}^n a_i^2 (J_j^{(k,l)})^2 \times \frac{\mathbf{m}(A \cap C_i)}{\mathbf{m}(\mathcal{T})}$$

Since for all i , $|a_i| \leq 1$, we get

$$\text{Var}\left(\tilde{\mathcal{M}}(f^{(k,l)}g)\right) \leq \frac{\mathbf{m}(A)}{\mathbf{m}(\mathcal{T})} \sum_{j=0}^{\infty} (J_j^{(k,l)})^2$$

Since $\sum_{j=0}^{\infty} J_j^2 < \infty$, we have $\sum_{j=0}^{\infty} (J_j^{(k,l)})^2 \rightarrow 0$ as $k, l \rightarrow \infty$, and thus

$$\sup_g \text{Var}\left(\tilde{\mathcal{M}}(f^{(k,l)}g)\right) \rightarrow 0 \quad \text{as} \quad k, l \rightarrow \infty. \quad (2.5.26)$$

Hence (2.5.23) follows from (2.5.26) and the fact that $\mathbf{E}[|Z| \wedge 1] \leq \text{Var}(Z)$ for a centered random variable: the process \tilde{Q} is well-defined.

Let us now prove that \tilde{Q} is \mathcal{C} -exchangeable. Integrating in (2.5.22) gives

$$\forall A \in \mathcal{A}, \quad \tilde{Q}_A = \sum_{j=0}^{\infty} J_j (\mathbb{1}_{\tau_j \in A} - \frac{\mathbf{m}(A)}{\mathbf{m}(\mathcal{T})})$$

which is actually closer to the expression usually employed in the one-dimensional case (see [53, Equation (2.1)] or [55, Equation (13)]). Then, still conditionally on J , \tilde{Q} is an infinite linear combination of independent set-indexed processes $\{\mathbb{1}_{\tau_j \in A} - \mathbf{m}(\cdot)/\mathbf{m}(\mathcal{T})\}$ which are easily shown to be \mathcal{C} -exchangeable.

2.5.6 Representation theorems

In this section, we investigate several representations that one could deduce from a stationarity assumption on the process X . Our biggest regret is that none of those obtained so far encompasses the case of the sifBm .

Analogy with a discrete case

Let us start by giving an illustration of what \mathcal{A} -exchangeability and \mathcal{A} -increment stationarity would mean in the special case of the infintary tree. The reason why we briefly focus on this ‘off-track’ example is because our notions reduce to more familiar ones in this setting, *i.e.* stationarity will be expressed as invariance of the distribution under some group action. Moreover, this example has also been studied as a ‘toy model’ by Aldous in [6, Section 13].

As hinted in Example 2.5.17, the starting point is the celebrated de Finetti’s theorem [55, Theorem 1.1] which tells that any exchangeable infinite sequence of random variables is conditionally iid. Aldous and Hoover [5, 6, 45] then reframed and generalized it to come up with the more general notion of *partial exchangeability* and the representation theorem for exchangeables arrays [55, Corollary 7.23] that now bears their name. This result has seen successive generalizations in several directions, see *e.g.* [11, 24, 25, 52], even some going far beyond classical probability theory [43].

Since this literature mostly focuses on the discrete case, we will commit a slight infringement and forget until the end of the section the hypotheses on \mathbf{m} (Definition 1.3.1) and $d_{\mathcal{A}}$ (Definition 1.4.1) that make our setting continuous. We do so because we found interesting to see how our notions would fit in this landscape and hopefully yield more intuition.

For now and until further notice, suppose that $\mathcal{T} = \bigcup_{k \in \mathbf{N}} \mathbf{N}^k$ is the *infintary tree* rooted at $0_{\mathcal{T}}$. It is endowed with its usual partial order \preccurlyeq and indexing collection \mathcal{A} from Example 1.2.24. We recall that a generic element $t \in \mathcal{T}$ is denoted $t = t_1 \dots t_k$ for $k \in \mathbf{N}$ and $t_1, \dots, t_k \in \mathbf{N}$ together with the convention that $t = 0_{\mathcal{T}}$ if $k = 0$. Moreover, two vertices $s = s_1 \dots s_k$ and $t = t_1 \dots t_l$ in \mathcal{T} may be concatenated to give the vertex $st = s_1 \dots s_k t_1 \dots t_l$.

DEFINITION 2.5.24 (Morphisms on \mathcal{T}). *A monomorphism of \mathcal{T} is an injective order embedding $\varphi : (\mathcal{T}, \preccurlyeq) \rightarrow (\mathcal{T}, \preccurlyeq)$. The set of all monomorphisms on \mathcal{T} is denoted by $\text{Mon}(\mathcal{T})$.*

An automorphism of \mathcal{T} is a bijective order embedding $\varphi : (\mathcal{T}, \preccurlyeq) \rightarrow (\mathcal{T}, \preccurlyeq)$. The set of all automorphisms on \mathcal{T} is denoted by $\text{Aut}(\mathcal{T})$.

REMARKS 2.5.25. \diamond *These notions of morphisms are consistent with the Definition 1.2.4 of morphism of indexing semilattices. For more insight on the automorphisms on a tree, we refer to [97].*

- \diamond *$\text{Mon}(\mathcal{T})$ (resp. $\text{Aut}(\mathcal{T})$) is a monoid (resp. group) under composition.*
- \diamond *\mathcal{T} may be seen as a directed graph (\mathcal{T}, E) where \mathcal{T} is the set of vertices and the set of edges E is such that for all $s, t \in \mathcal{T}$, $(s, t) \in E$ if and only if $s \prec t$ and there is no $u \in \mathcal{T}$ such that $s \prec u \prec t$. In particular, everything that is stated may be translated using graph-theoretic vocabulary. But since we chose to stick with posets since it is the basic structure pertaining to the thesis.*
- \diamond *With the previous remark in mind, what has been told so far is a special case considered by Aldous in [6, (12.17)]. However, we split apart from what he goes on doing in Section 13 in *loc. cit.* where he studies a notion of exchangeability on \mathcal{T} , but as a non-directed graph.*

To go with the infintary tree \mathcal{T} , we denote by \mathbf{m} the counting measure on \mathcal{T} . Remark that Definitions 2.5.6 and 2.5.7 still makes sense for such \mathcal{T} and \mathbf{m} , even though Proposition 2.5.8 fails.

PROPOSITION 2.5.26 (\mathcal{A} -exchangeability on the infintary tree). *Consider the following statements:*

- (i) *X is \mathcal{A} -exchangeable.*

(ii) The distribution of the \mathcal{T} -indexed process $X \circ A(\cdot) = \{X_{A(t)} : t \in \mathcal{T}\}$ is invariant under the action of $\text{Aut}(\mathcal{T})$, i.e.

$$\forall \varphi \in \text{Aut}(\mathcal{T}), \quad X \circ A(\cdot) \stackrel{\text{fdd}}{=} \{X_{A(\varphi(t))} : t \in \mathcal{T}\}.$$

(iii) There exists a collection of iid random variables $\{U_t : t \in \mathcal{T}\}$ uniform on $[0, 1]$ and a measurable map $f : \bigcup_{0 \leq k \leq n} [0, 1]^k \rightarrow \mathbf{R}$ with the convention that $[0, 1]^0 = \{0\}$ and such that

$$\forall A \in \mathcal{A}, \quad X_A = f((U_t)_{t \in A}) \quad (2.5.27)$$

and $f(0) = 0$.

Then we have (i) \Leftrightarrow (ii) \Leftrightarrow (iii).

We conjecture that the implication (ii) \Rightarrow (iii) in Proposition 2.5.26 holds true as well. This implication is very close to [52, Theorem 3.2] but does not quite fall under the scope of their assumptions unfortunately. For now, we only know that the equivalence holds if one cuts \mathcal{T} at a finite height. We refer to [52, Example 2.1(c)] for more details.

Proof. Suppose (i). Consider $\varphi \in \text{Aut}(\mathcal{T})$, $t^1, \dots, t^k \in \mathcal{T}$ and let us show that $(X_{A(t^1)}, \dots, X_{A(t^k)}) \stackrel{\text{law}}{=} (X_{A(\varphi(t^1))}, \dots, X_{A(\varphi(t^k))})$. Since X is \mathcal{A} -exchangeable, it is enough to show that $(A(t^1), \dots, A(t^k)) \stackrel{(c)}{=} (A(\varphi(t^1)), \dots, A(\varphi(t^k)))$. Let $J \subseteq \llbracket 1, k \rrbracket$. Then, repeatedly using the fact that φ is an automorphism, we get

$$\begin{aligned} m\left(\bigcap_{j \in J} A(\varphi(t^j))\right) &= m\left(A\left(\bigwedge_{j \in J} \varphi(t^j)\right)\right) \\ &= \#\{s \in \mathcal{T} : s \preceq \varphi\left(\bigwedge_{j \in J} t^j\right)\} \\ &= \#\{s \in \mathcal{T} : s \preceq \bigwedge_{j \in J} t^j\} \\ &= m\left(\bigcap_{j \in J} A(t^j)\right). \end{aligned}$$

Hence (ii) follows.

Conversely, suppose (ii). Consider $A(s^1), \dots, A(s^k), A(t^1), \dots, A(t^k)$ in \mathcal{A} such that $(A(s^1), \dots, A(s^k)) \stackrel{(c)}{=} (A(t^1), \dots, A(t^k))$ and let us prove that $(X_{A(s^1)}, \dots, X_{A(s^k)}) \stackrel{\text{law}}{=} (X_{A(t^1)}, \dots, X_{A(t^k)})$. Using (ii), it is enough to find $\varphi \in \text{Aut}(\mathcal{T})$ such that for all $j \in \llbracket 1, k \rrbracket$, $\varphi(s^j) = t^j$. Denote for all $j \in \llbracket 1, k \rrbracket$, $s^j = s_1^j \dots s_{l_j}^j$ and $t^j = t_1^j \dots t_{l_j}^j$ where $l_j \in \mathbf{N}$ is indeed the same for s^j and t^j since $l_j = \#A(s^j) = \#A(t^j)$. We build by induction on $i \in \mathbf{N}$ an automorphism $\varphi_i \in \text{Aut}(\mathcal{T})$ such that for all $j \in \llbracket 1, k \rrbracket$ and $l \leq i \wedge l_j$, $\varphi_i(s_1^j \dots s_l^j) = t_1^j \dots t_l^j$. For that, remark that an automorphism φ on \mathcal{T} is fully characterized once given a permutation of \mathbf{N} on each vertex of \mathcal{T} , i.e. a map $\sigma \in \mathfrak{S}(\mathbf{N})^{\mathcal{T}}$ where for all $t \in \mathcal{T}$, $\sigma(t)$ indicates how $\varphi = \varphi_\sigma$ permutes the children of t . This is explained in greater details in [11, p.811] and [52, Example 2.3(d)].

Setting $\varphi_0 = \text{id}_{\mathcal{T}}$ is enough for the initialization. Now, suppose that $\varphi_i = \varphi_{\sigma_i}$ exists and let us define $\varphi_{i+1} = \varphi_{\sigma_{i+1}}$. For that, denote $\mathcal{T}_{i-1} = \bigcup_{k < i} \mathbf{N}^k$ the subtree of \mathcal{T} of height $i-1$ (it may be empty) and define $\sigma_{i+1}|_{\mathcal{T}_{i-1}} = \sigma_i|_{\mathcal{T}_{i-1}}$ and for all $j \in \llbracket 1, k \rrbracket$, $\sigma_{i+1}(s_1^j \dots s_i^j)$ is any permutation of \mathbf{N} such that $\sigma_{i+1}(s_1^j \dots s_i^j)(s_{i+1}^j) = t_{i+1}^j$. The only way for this permutation to be inconsistently defined would be to have some $l \neq j$ such that $s_1^j \dots s_{i+1}^j = s_1^l \dots s_{i+1}^l$ but $t_{i+1}^j \neq t_{i+1}^l$. But this cannot happen since $\#(A(s^j) \cap A(s^l)) = \#(A(t^j) \cap A(t^l))$. Hence σ_{i+1} is well-defined and $\varphi_{\sigma_{i+1}}$ verifies the required property. Hence (i) follows.

Suppose (iii), i.e. that the representation (2.5.27) holds for some f and $\{U_t : t \in \mathcal{T}\}$ as stated above. Consider $t^1, \dots, t^k \in \mathcal{T}$ and $\varphi \in \text{Aut}(\mathcal{T})$. Since φ is an automorphism, we have

$$\left((U_t)_{t \preceq t^j} \right)_{\substack{1 \leq j \leq k \\ 1 \leq i \leq i_j}} \stackrel{\text{law}}{=} \left((U_s)_{s \preceq \varphi(t^j)} \right)_{\substack{1 \leq j \leq k \\ 1 \leq i \leq i_j}}.$$

Hence $(X_{A(t^1)}, \dots, X_{A(t^k)}) \stackrel{\text{law}}{=} (X_{A(\varphi(t^1))}, \dots, X_{A(\varphi(t^k))})$, from which (ii) follows. \square

PROPOSITION 2.5.27 (\mathcal{A} -increment stationarity on the infinitary tree). *The following statements are equivalent:*

- (i) X is \mathcal{A} -increment stationary.
- (ii) For $\mathcal{T}^* = \mathcal{T} \setminus \{0_{\mathcal{T}}\}$, the distribution of the process $Y = \{\Delta X_{\{t\}} : t \in \mathcal{T}^*\}$ is invariant under the action of $\text{Mon}(\mathcal{T})$, i.e.

$$\forall \varphi \in \text{Mon}(\mathcal{T}), \quad Y \stackrel{\text{fdd}}{=} \{Y_{\varphi(t)} : t \in \mathcal{T}^*\}.$$

Remark that for $t = t_1 \dots t_j \in \mathcal{T}^*$, we have $\{t\} = A(t) \setminus A(t_1 \dots t_{j-1})$, so $\Delta X_{\{t\}}$ is well-defined in Proposition 2.5.27.

Proof. Suppose (i). Consider $\varphi \in \text{Mon}(\mathcal{T})$ and $t^1, \dots, t^k \in \mathcal{T}^*$. Denote for all $j \in \llbracket 1, k \rrbracket$, $t^j = t_1^j \dots t_{i_j}^j$ where $i_j \in \mathbb{N}^*$. Since $\varphi \in \text{Mon}(\mathcal{T})$, we may find $\varphi' \in \text{Aut}(\mathcal{T})$ such that

$$\forall t \in \mathcal{T}, \quad \varphi(t) = \varphi(0_{\mathcal{T}})\varphi'(t). \quad (2.5.28)$$

Since φ' is an automorphism, we have

$$\left(A(t_1^j \dots t_{i_j}^j) \right)_{\substack{1 \leq j \leq k \\ 1 \leq i \leq i_j}} \stackrel{(c)}{=} \left(A(\varphi'(t_1^j \dots t_{i_j}^j)) \right)_{\substack{1 \leq j \leq k \\ 1 \leq i \leq i_j}}. \quad (2.5.29)$$

Combining (2.5.28) and (2.5.29) yields

$$\left(A(t_1^j \dots t_{i_j}^j) \right)_{\substack{1 \leq j \leq k \\ 1 \leq i \leq i_j}} \stackrel{(c)}{=} \left(A(\varphi(t_1^j \dots t_{i_j}^j)) \setminus A(\varphi(0_{\mathcal{T}})) \right)_{\substack{1 \leq j \leq k \\ 1 \leq i \leq i_j}}.$$

Since X is \mathcal{A} -increment stationary, it follows that

$$\left(X_{A(t_1^j \dots t_{i_j}^j)} \right)_{\substack{1 \leq j \leq k \\ 1 \leq i \leq i_j}} \stackrel{\text{law}}{=} \left(\Delta X_{A(\varphi(t_1^j \dots t_{i_j}^j)) \setminus A(\varphi(0_{\mathcal{T}}))} \right)_{\substack{1 \leq j \leq k \\ 1 \leq i \leq i_j}}. \quad (2.5.30)$$

Taking differences between successive terms (with respect to the order of \mathcal{T}) in (2.5.30) yields

$$\left(Y_{t_1^j \dots t_{i_j}^j} \right)_{\substack{1 \leq j \leq k \\ 1 \leq i \leq i_j}} \stackrel{\text{law}}{=} \left(Y_{\varphi(t_1^j \dots t_{i_j}^j)} \right)_{\substack{1 \leq j \leq k \\ 1 \leq i \leq i_j}}$$

from which (ii) follows.

Conversely, suppose (ii). Consider $A(s^1), \dots, A(s^k), A(t), A(t^1), \dots, A(t^k)$ in \mathcal{A} such that $A(t) \subseteq \bigcap_{j=1}^k A(t^j)$ and $(A(s^1), \dots, A(s^k)) \stackrel{(c)}{=} (A(t^1) \setminus A(t), \dots, A(t^k) \setminus A(t))$. Write $u^1, \dots, u^k \in \mathcal{T}$ such that for all $j \in \llbracket 1, k \rrbracket$, $t^j = tu^j$. In particular, we have $(A(s^1), \dots, A(s^k)) \stackrel{(c)}{=} (A(u^1), \dots, A(u^k))$. So, using the construction by induction from the proof of Proposition 2.5.26, we may find $\varphi' \in \text{Aut}(\mathcal{T})$ such

that for all $j \in \llbracket 1, k \rrbracket$, $\varphi'(s^j) = u^j$. By concatenation, $\varphi = t\varphi'$ defines a monomorphism such that for all $j \in \llbracket 1, k \rrbracket$, $\varphi(s^j) = t^j$. Hence, (ii) implies

$$(Y_{s^j})_{1 \leq j \leq k} \stackrel{\text{law}}{=} (Y_{t^j})_{1 \leq j \leq k}. \quad (2.5.31)$$

Making summations along increasing paths in (2.5.31) yields

$$(X_{A(s^j)})_{1 \leq j \leq k} \stackrel{\text{law}}{=} (\Delta X_{A(t^j) \setminus A(t)})_{1 \leq j \leq k}$$

from which (i) follows. \square

In Proposition 2.5.27, some could raise the issue that $\text{Mon}(\mathcal{T})$ is not a group per say, but only a monoid, and they would be right. The reason is that the infinitary tree is not closed under translation. However, it is possible to extend Y to a bigger tree — on which a group acts this time — in the same way a stationary \mathbf{N} -indexed process may be extended to a stationary \mathbf{Z} -indexed one. This tree is the $(1, \omega)$ -regular tree defined as

$$\vec{\mathcal{T}} = \bigcup_{k \in \mathbf{Z}} \mathbf{N}^{(-\infty, k] \cap \mathbf{Z}}.$$

Even though $\vec{\mathcal{T}}$ is not exactly a tree according to Neveu's definition, its poset structure still makes it 'tree-like' in the sense that each node $t = \dots t_{k-1} t_k \in \vec{\mathcal{T}}$ has one 'ancestor' — $\dots t_{k-2} t_{k-1}$ — and countably many 'children' — all the $\dots t_{k-1} t_k s$ for $s \in \mathbf{N}$ — hence the name of $(1, \omega)$ -regular tree. The infinitary tree \mathcal{T} is embeddable into $\vec{\mathcal{T}}$ by picking a vertex $\rho = \dots \rho_{-1} \rho_0 \in \mathbf{N}^{(-\infty, 0] \cap \mathbf{Z}}$ and using the order embedding $\mathcal{T} \ni t = t_1 \dots t_k \mapsto \rho t = \dots \rho_{-1} \rho_0 t_1 \dots t_k \in \vec{\mathcal{T}}$. Then, we may show that \mathcal{A} -increment stationarity for X is equivalent to Y having an extension to $\vec{\mathcal{T}}$ whose distribution is invariant under the action of $\text{Aut}(\vec{\mathcal{T}})$.

Unfortunately, we have not been able to establish in Proposition 2.5.27 an Aldous-Hoover type of representation for \mathcal{A} -increment stationary processes in the spirit of (2.5.27).

Let us now turn our attention towards an analog of Proposition 2.5.26 for \mathcal{C} -exchangeable processes.

PROPOSITION 2.5.28 (\mathcal{C} -exchangeability on the infinitary tree). *The following statements are equivalent:*

- (i) X is \mathcal{C} -exchangeable.
- (ii) The distribution of the \mathcal{T} -indexed process $Y = \{\Delta X_{\{t\}} : t \in \mathcal{T}\}$ is invariant under the action of $\mathfrak{S}(\mathcal{T})$, i.e.

$$\forall \sigma \in \mathfrak{S}(\mathcal{T}), \quad Y \stackrel{\text{fdd}}{=} \{\Delta X_{\{\sigma(t)\}} : t \in \mathcal{T}\}.$$

- (iii) There exists a collection of iid random variables $\{U_\emptyset\} \cup \{U_t : t \in \mathcal{T}\}$ uniform on $[0, 1]$ and a measurable map $f : [0, 1]^2 \rightarrow \mathbf{R}$ such that

$$\forall A \in \mathcal{A}, \quad X_A = \sum_{t \in A} f(U_\emptyset, U_t). \quad (2.5.32)$$

Proof. This thread of equivalences is just a barely hidden application of de Finetti's theorem and the related Aldous-Hoover representation.

Suppose (i). Since $\{\{t\} : t \in \mathcal{T}\} \subseteq \mathcal{C}$, the process Y may be seen as an exchangeable sequence. Hence (ii) holds.

Conversely, if (ii) holds, then X is $\{\{t\} : t \in \mathcal{T}\}$ -exchangeable. Since any element in \mathcal{C} is a finite union of singletons, X is actually \mathcal{C} -exchangeable. Hence (i) and (ii) are equivalent.

Suppose (ii). By [55, Lemma 7.1], we may find a collection of iid random variables $\{U_\emptyset\} \cup \{U_t : t \in \mathcal{T}\}$ uniform on $[0, 1]$ and a measurable map $f : [0, 1]^2 \rightarrow \mathbf{R}$ such that for all $t \in \mathcal{T}$, $\Delta X_{\{t\}} = f(U_\emptyset, U_t)$. The representation (2.5.32) — and hence (iii) — immediately follows.

Conversely, if (iii) holds, then we have for all $t \in \mathcal{T}$, $\Delta X_{\{t\}} = f(U_\emptyset, U_t)$ and thus, the distribution of Y is invariant under permutations. Hence (ii) and (iii) are equivalent. \square

REMARK 2.5.29. *The two representations (2.5.27) and (2.5.32) were given in a form that makes them easier to compare. This indicates how much more restrictive being \mathcal{C} -exchangeable is compared to being \mathcal{A} -exchangeable (which is not too surprising since any process is \mathcal{A} -exchangeable when $\mathcal{T} = \mathbf{R}_+$). This also shows that any representation for discrete \mathcal{A} -increment stationary processes should lie 'somewhere in between' (2.5.27) and (2.5.32). We wish we had one, but that remains to be proven.*

We hope that this interlude in the discrete world helped to clarify our new notions of stationarity. In the continuous case, it is still not clear to us whether or not they are equivalent to invariance under the action of some measure-preserving order embeddings.

Representations of \mathcal{C} -exchangeable processes

Let us put aside the infinitary tree of the previous section and go back to the more general case. We prove a generalization of Kallenberg's [55, Theorem 3.15] which is a representation for one-dimensional processes with exchangeable increments.

THEOREM 2.5.30 (Representation of \mathcal{C} -exchangeable processes when $m(\mathcal{T}) < +\infty$). *Suppose that $0 < m(\mathcal{T}) < +\infty$. The following statements are equivalent:*

- (i) X is \mathcal{C} -exchangeable and outer continuous in probability (Definition 2.4.1).
- (ii) There exists a random triplet $(b, \sigma^2, J) : \Omega \rightarrow \mathbf{R} \times \mathbf{R}_+ \times \mathcal{M}(\mathbf{R}^*)$ such that J is \mathbf{N} -valued, $\int_{\mathbf{R}^*} x^2 J(dx) < \infty$ a.s., an independent siBb \widehat{W} (Example 2.5.19) and a sicK \widetilde{Q} of intensity J (Example 2.5.23) such that $\widetilde{Q} \perp_J (b, \sigma^2, \widehat{W})$ and

$$\forall A \in \mathcal{A}, \quad X_A = b m(A) + \sigma \widehat{W}_A + \widetilde{Q}_A. \quad (2.5.33)$$

Under those conditions, the distribution of (b, σ^2, J) is unique and characterizes the distribution of X . It is called the directing triplet of X .

Proof. The implication (ii) \Rightarrow (i) has already been carried out through Examples 2.5.19 and 2.5.23.

Conversely, suppose (i). We first prove the result in the special case when $\mathcal{T} \in \mathcal{A}(u)$. Repeating the coupling argument giving (2.4.7) indicates that in order to prove (2.5.33), it is enough to show that

$$X \stackrel{\text{fdd}}{=} b m + \sigma \widehat{W} + \widetilde{Q}. \quad (2.5.34)$$

for some siBb \widehat{W} and sicK \widetilde{Q} of intensity J as described in (ii).

Fix $\phi \in \Phi_{\max}(\mathcal{A}(u)) = \{\phi \in \Phi(\mathcal{A}(u)) : \mathbf{m}(\phi(0)) = 0 \text{ and } \phi(1) = \mathcal{T}\}$. By Proposition 2.5.8, we know that X^ϕ is a $[0, 1]$ -indexed process with exchangeable increments. Thus, by [55, Theorem 3.15], there exists a random triplet $(b(\phi), \sigma^2(\phi), J(\phi))$, an independent Brownian bridge \widehat{W}' and a $[0, 1]$ -indexed sicK \widetilde{Q}' of intensity $J(\phi)$ such that $\widetilde{Q}' \perp_{J(\phi)} (b(\phi), \sigma^2(\phi), \widehat{W}')$ and

$$\{\Delta X_t^\phi : t \in [0, 1]\} \stackrel{\text{fdd}}{=} \{b(\phi)t + \sigma(\phi)\widehat{W}'_t + \widetilde{Q}'_t : t \in [0, 1]\}. \quad (2.5.35)$$

Consider an siBb \widehat{W} independent from $(b(\phi), \sigma^2(\phi), J(\phi))$ and a sicK \widetilde{Q} of intensity $J(\phi)$ such that $\widetilde{Q} \perp_{J(\phi)} (b(\phi), \sigma^2(\phi), \widehat{W})$ and define the set-indexed process Y by

$$\forall A \in \mathcal{A}, \quad Y_A = b(\phi)\mathbf{m}(A) + \sigma(\phi)\widehat{W}_A + \widetilde{Q}_A.$$

Let us show that $X \stackrel{\text{fdd}}{=} Y$, which will be enough to prove (2.5.34). Let $C \in \mathcal{C}$ such that $\mathbf{m}(C) = 0$. Since X (resp. Y) is \mathcal{C} -exchangeable and $\mathbf{m}(C) = \mathbf{m}(\emptyset)$, we have $\Delta X_C \stackrel{\text{law}}{=} X_\emptyset = 0$ a.s. (resp. $\Delta Y_C \stackrel{\text{law}}{=} Y_\emptyset = 0$ a.s.). Hence, by Theorem 1.3.13 and Remark 1.3.14, it is enough to show that

$$\forall \psi \in \Phi_{\max}(\mathcal{A}(u)), \quad \Delta X^\psi \stackrel{\text{fdd}}{=} \Delta Y^\psi. \quad (2.5.36)$$

We already know by (2.5.35) that $\Delta X^\phi \stackrel{\text{fdd}}{=} \Delta Y^\phi$. Let $\psi \in \Phi_{\max}(\mathcal{A}(u))$. Applying [55, Theorem 3.15] once more similarly yields a triplet $(b(\psi), \sigma^2(\psi), J(\psi))$ independent from \widehat{W} such that

$$\Delta X^\psi \stackrel{\text{fdd}}{=} \Delta Z^\psi \quad \text{where} \quad \forall A \in \mathcal{A}, \quad Z_A = b(\psi)\mathbf{m}(A) + \sigma(\psi)\widehat{W}_A + \widetilde{R}_A$$

and \widetilde{R} is a sicK of intensity $J(\psi)$ such that $\widetilde{R} \perp_{J(\psi)} (b(\psi), \sigma^2(\psi), \widehat{W})$. Since the distribution of ΔY^ψ (resp. ΔZ^ψ) is characterized by the distribution of $(b(\phi), \sigma^2(\phi), J(\phi))$ (resp. $(b(\psi), \sigma^2(\psi), J(\psi))$), it is enough to show that

$$(b(\phi), \sigma^2(\phi), J(\phi)) \stackrel{\text{law}}{=} (b(\psi), \sigma^2(\psi), J(\psi)). \quad (2.5.37)$$

For all $n \in \mathbf{N}$ and $j \in \llbracket 0, n! \rrbracket$, denote $\Delta_{j,n}^\phi = \Delta X_{j/n!}^\phi - \Delta X_{(j-1)/n!}^\phi$ and

$$(b^{(n)}(\phi), \gamma^{(n)}(\phi), J^{(n)}(\phi)) = \left(\Delta X_1^\phi, \sum_{j=1}^{n!} (\Delta_{j,n}^\phi)^2, \sum_{j=1}^{n!} \delta_{\Delta_{j,n}^\phi} \right). \quad (2.5.38)$$

The notations $\Delta_{j,n}^\psi$ and $(b^{(n)}(\psi), \gamma^{(n)}(\psi), J^{(n)}(\psi))$ are defined similarly. By [55, Theorem 3.8], we know that there exists a non-negative random variable $\gamma(\phi)$ such that

$$\begin{cases} (b^{(n)}(\phi), \gamma^{(n)}(\phi), J^{(n)}(\phi)) \xrightarrow{\text{law}} (b(\phi), \gamma(\phi), J(\phi)) & \text{as } n \rightarrow \infty \\ \sigma^2(\phi) = \gamma(\phi) - \int_{\mathbf{R}} x^2 J(\phi)(dx) \end{cases} \quad (2.5.39)$$

and likewise for ψ . Since $\phi, \psi \in \Phi_{\max}(\mathcal{A}(u))$, we have $v(\phi) = \mathbf{m}(\mathcal{T}) = v(\psi)$ and

$$\forall n \in \mathbf{N}, \quad \left(\phi\left(\frac{j}{n!}\right) \setminus \phi\left(\frac{j-1}{n!}\right) \right)_{1 \leq j \leq n!} \stackrel{(c)}{=} \left(\psi\left(\frac{j}{n!}\right) \setminus \psi\left(\frac{j-1}{n!}\right) \right)_{1 \leq j \leq n!}. \quad (2.5.40)$$

By Proposition 2.5.16, we know that X is $\mathcal{C}(u)$ -exchangeable, thus (2.5.40) implies that

$$\forall n \in \mathbf{N}, \quad (\Delta_{j,n}^{\phi})_{1 \leq j \leq n!} \stackrel{\text{law}}{=} (\Delta_{j,n}^{\psi})_{1 \leq j \leq n!}. \quad (2.5.41)$$

Due to the definition (2.5.38) of the triplets, (2.5.41) implies that

$$\forall n \in \mathbf{N}, \quad (b^{(n)}(\phi), \gamma^{(n)}(\phi), J^{(n)}(\phi)) \stackrel{\text{law}}{=} (b^{(n)}(\psi), \gamma^{(n)}(\psi), J^{(n)}(\psi)). \quad (2.5.42)$$

Hence, (2.5.37) follows from (2.5.39) and (2.5.42), which concludes the proof when $\mathcal{T} \in \mathcal{A}(u)$.

Now suppose that $\mathcal{T} \notin \mathcal{A}(u)$. Since $\mathcal{T}_n \uparrow \mathcal{T}$, we may use Proposition 1.3.12 to find $\phi \in \Phi(\mathcal{B}_m)$ such that $\mathbf{m}(\phi(0)) = 0$, for all $t \in [0, 1)$, $\phi(t) \in \mathcal{A}(u)$ and $\phi(1) = \mathcal{T}$. Using this flow, the proof would go as before provided that one is able to 'extend X by continuity at \mathcal{T} ', i.e. to show that there exists a random variable Z such that $\Delta X_U \xrightarrow{\mathbf{P}} Z$ as $U \in \mathcal{A}(u) \rightarrow \mathcal{T}$ in (\mathcal{B}_m, d_m) . Let us show that. By (1.4.7), we have

$$\forall U, V \in \mathcal{A}(u), \quad |\Delta X_U - \Delta X_V| \leq |\Delta X_{U \setminus V}| + |\Delta X_{V \setminus U}|. \quad (2.5.43)$$

Fix $\psi \in \Phi(\mathcal{A})$ such that $\mathbf{m}(\psi(0)) = 0$. Then for $U, V \in \mathcal{A}(u)$,

$$\Delta X_{U \setminus V} \stackrel{\text{law}}{=} X_{m(U \setminus V)/v(\psi)}^{\psi} \xrightarrow{\mathbf{P}} 0 \quad \text{as} \quad d_m(U, V) \rightarrow 0 \quad (2.5.44)$$

where the equality in law is by $\mathcal{C}(u)$ -exchangeability and the convergence by outer continuity in probability of X^{ψ} at 0. Since $\mathbf{m}(\mathcal{T}) < \infty$, we have $d_m(U, V) \rightarrow 0$ as $U, V \rightarrow \mathcal{T}$ in (\mathcal{B}_m, d_m) . Hence (2.5.44) yields

$$\Delta X_{U \setminus V} \xrightarrow{\mathbf{P}} 0 \quad \text{as} \quad U, V \in \mathcal{A}(u) \rightarrow \mathcal{T} \text{ in } (\mathcal{B}_m, d_m). \quad (2.5.45)$$

Combining (2.5.43) and (2.5.45), we get

$$|\Delta X_U - \Delta X_V| \xrightarrow{\mathbf{P}} 0 \quad \text{as} \quad U, V \in \mathcal{A}(u) \rightarrow \mathcal{T} \text{ in } (\mathcal{B}_m, d_m)$$

from which the result follows by completeness of $L^0(\Omega)$. \square

Making use of the same shenanigans, we may show the following generalization of Bühlmann's theorem [55, Theorem 1.19].

THEOREM 2.5.31 (Representation of \mathcal{C} -exchangeable processes when $\mathbf{m}(\mathcal{T}) = +\infty$). *Suppose that $\mathbf{m}(\mathcal{T}) = +\infty$. The following statements are equivalent:*

(i) X is \mathcal{C} -exchangeable and outer continuous in probability.

(ii) *There exists a random triplet $(b, \sigma, \nu) : \Omega \rightarrow \mathbf{R} \times \mathbf{R}_+ \times \mathcal{M}(\mathbf{R}^*)$ such that $\int_{\mathbf{R}^*} (1 \wedge x^2) \nu(dx) < \infty$ a.s., an independent siBm W (Example 2.2.12) and a set-indexed compensated compound Cox process \tilde{P} of intensity ν (Example 2.5.21) such that $\tilde{P} \perp_{\nu} (b, \sigma, W)$ and*

$$\forall A \in \mathcal{A}, \quad X_A = b\mathbf{m}(A) + \sigma W_A + \tilde{P}_A. \quad (2.5.46)$$

Under those conditions, the distribution of (b, σ, ν) is unique and characterizes the distribution of X . It is called the directing triplet of X .

COROLLARY 2.5.32 (Integration with respect to a \mathcal{C} -exchangeable process). *Suppose that X has the representation (2.5.34). Then, X extends to a stochastic measure and we have the representation*

$$\forall f \in L(X), \quad \mathcal{X}(f) = b \int_{\mathcal{T}} f \, d\mathbf{m} + \sigma \mathcal{W}(f) + \tilde{\mathcal{P}}(f) \quad (2.5.47)$$

where \mathcal{W} (resp. $\tilde{\mathcal{P}}$) is the linear process associated with W (resp. \tilde{P}) and

$$L(X) = \left\{ f \in \mathbf{R}^{\mathcal{T}} : \text{for } \mathbf{P}\text{-a.e. } \omega \in \Omega, f \in L^0(b(\omega), \sigma^2(\omega), \nu(\omega)) \right\}. \quad (2.5.48)$$

Proof. The representation (2.5.47) is obtained from (2.5.46) in the same way (2.4.8) is obtained from (2.4.7). What is left to establish is (2.5.48). The direct inclusion is a consequence of (2.5.47). For the converse inclusion, consider $f \in \mathbf{R}^{\mathcal{T}}$ such that for \mathbf{P} -a.e. $\omega \in \Omega$, $f \in L^0(b(\omega), \sigma^2(\omega), \nu(\omega))$ as well as a deterministic sequence $(f^{(j)})_{j \in \mathbf{N}}$ of simple functions converging pointwise to f . Denote for all $j, k \in \mathbf{N}$, $f^{(j,k)} = f^{(j)} - f^{(k)}$. By Theorem 2.3.11, we know that

$$\sup_{\substack{g \in \mathcal{E}: \\ |g| \leq 1}} \mathbf{E} \left[|\mathcal{X}(f^{(j,k)})g| \wedge 1 \mid b, \sigma^2, \nu \right] \xrightarrow{\text{a.s.}} 0 \quad \text{as } j, k \rightarrow \infty.$$

Taking expectation and using the dominated convergence theorem above yields

$$\mathbf{E} \left[\sup_{\substack{g \in \mathcal{E}: \\ |g| \leq 1}} \mathbf{E} \left[|\mathcal{X}(f^{(j,k)})g| \wedge 1 \mid b, \sigma^2, \nu \right] \right] \rightarrow 0 \quad \text{as } j, k \rightarrow \infty.$$

Hence, permuting supremum and expectation yields

$$\sup_{\substack{g \in \mathcal{E}: \\ |g| \leq 1}} \mathbf{E} \left[|\mathcal{X}(f^{(j,k)})g| \wedge 1 \right] \rightarrow 0 \quad \text{as } j, k \rightarrow \infty.$$

Thus $f \in L(X)$ and the result follows. \square

2.6 Perspectives

As we have done for Chapter 1 before, we give some directions in which the work of this chapter could be developed.

2.6.1 Characteristics for stochastic measures

When $\mathcal{T} = \mathbf{R}_+$, it is well-known that ‘good integrators’ are semimartingales. In [62, Theorem 9.4.1], Kwapień and Woyczyński described a way to define a triplet (b, σ^2, ν) for a general semimartingale X and used it to give a description of the space $L(X)$ that extends Rajput and Rosiński’s result for siLévy’s. It would be interesting to be able to extend those results in order to fully characterize ‘good integrators’ for a general \mathcal{T} and subsequently describe the space of integrable functions.

2.6.2 Towards functional martingale inequalities

Another interesting direction would be to try establish functional maximal inequalities of the form

$$\mathbf{E} \left[\sup_{f \in \mathcal{F}} |\mathcal{X}(f)|^\gamma \right] \leq \kappa_{\mathcal{F}, \gamma} \mathbf{E} [|\mathcal{X}(f^*)|^\gamma]$$

for some subset $\mathcal{F} \subseteq L(X)$, $\gamma > 1$, $\kappa_{\mathcal{F}, \gamma} > 0$ and $f^* = \sup_{f \in \mathcal{F}} f$. This would relegate the inequalities obtained in Section 2.3.3 to the special case $\mathcal{F} = \{\mathbf{1}_A : A \in [\emptyset, U]\}$ for $U \in \mathcal{A}(u)$.

As an application, this kind of inequality could be used to prove a strong version of the functional Lévy-Itô decomposition (2.4.8).

2.6.3 More representation theorems

As we stated before in Section 2.5.6, we would love to see whether one may be able to establish a representation theorem for \mathcal{A} -increment stationary processes, being for a general \mathcal{T} or the infinitary tree. Maybe an easier task would be to first prove the implication (ii) \Rightarrow (iii) in Proposition 2.5.26, which is much closer to the existing results we are aware of.

2.7 Perspectives (français)

2.7.1 Caractéristiques pour les mesures stochastiques

Lorsque $\mathcal{T} = \mathbf{R}_+$, il est bien connu que les "bons intégrateurs" sont les semimartingales. Kwapien and Woyczyński [62, Theorem 9.4.1] donnent une manière de définir un triplet (b, σ^2, ν) pour une semimartingale quelconque X et l'utilisent pour donner une description de l'espace $L(X)$ qui rejoint les travaux de Rajput et Rosiński dans le cas des processus de Lévy indexés par des ensembles. Il serait intéressant de pouvoir étendre ces résultats afin de complètement caractériser les "bons intégrateurs" pour un \mathcal{T} général et décrire l'espace des fonctions intégrables correspondant.

2.7.2 Vers des inégalités fonctionnelles de type martingale

Une autre piste est d'essayer d'établir une inégalité maximale fonctionnelle de la forme

$$\mathbf{E} \left[\sup_{f \in \mathcal{F}} |\mathcal{X}(f)|^\gamma \right] \leq \kappa_{\mathcal{F}, \gamma} \mathbf{E} [|\mathcal{X}(f^*)|^\gamma]$$

pour un sous-ensemble $\mathcal{F} \subseteq L(X)$ donné, $\gamma > 1$, $\kappa_{\mathcal{F}, \gamma} > 0$ et $f^* = \sup_{f \in \mathcal{F}} f$. Les inégalités obtenues en Section 2.3.3 se réduiraient alors au cas particulier où $\mathcal{F} = \{ \mathbb{1}_A : A \in [\emptyset, U] \}$ pour un $U \in \mathcal{A}(u)$ donné.

L'application que nous avons en la tête serait d'obtenir une version forte de la représentation de Lévy-Itô fonctionnelle (2.4.8).

2.7.3 Davantage de théorèmes de représentation

Comme nous l'avons mentionné en Section 2.5.6, nous serions ravis de pouvoir obtenir un théorème de représentation pour les processus à \mathcal{A} -accroissements stationnaires, que ce soit pour un \mathcal{T} général ou bien l'arbre binaire infini. Peut-être qu'une première tâche pourrait être de démontrer l'implication (ii) \Rightarrow (iii) de la Proposition 2.5.26, beaucoup plus proche des résultats existants dont nous avons connaissance.

Sample path properties of generalized processes 3

Contents

3.1	Introduction	108
3.2	The continuous space $\mathcal{C}^\Phi(\mathcal{A})$	108
3.2.1	Projective limit of Polish spaces	108
3.2.2	The space $\mathcal{C}^\Phi(\mathcal{A})$ and its topology	109
3.2.3	Some processes living in $\mathcal{C}^\Phi(\mathcal{A})$	110
3.2.4	A link with flows	114
3.3	The càdlàg space $\mathcal{D}^\Phi(\mathcal{A})$	116
3.3.1	A quick review of $\mathcal{D}(0,1)$	117
3.3.2	Blind spots of geodesic flows	117
3.3.3	The space $\mathcal{D}^\Phi(\mathcal{A})$ and its topology	120
3.3.4	A Lévy-Itô decomposition	121
3.4	Hölder exponents for generalized processes	124
3.4.1	Set-indexed pointwise Hölder exponents	124
3.4.2	Regularity of generic OJI functions	127
3.5	Hölder regularity of the primitive with respect to a siLévy	132
3.5.1	Divide and conquer	132
3.5.2	Hölder regularity	133
3.5.3	d_γ -localized regularity	142
3.5.4	Examples and applications	143
3.6	Perspectives	147
3.6.1	Improve on the finite-dimensional assumption	147
3.6.2	Towards a uniform upper bound for the primitive process	148
3.6.3	Regularity of \mathcal{C} -exchangeable processes	148
3.7	Perspectives (français)	149
3.7.1	Améliorer l'hypothèse de dimension finie	149
3.7.2	Vers une borne supérieure uniforme de la régularité de la primitive	149
3.7.3	Régularité des processus \mathcal{C} -échangeables	150

3.1 Introduction

When one wants to study sample path properties of a stochastic process $X = \{X_A : A \in \mathcal{A}\}$, it is not enough to consider X as a random function in $\mathbf{R}^{\mathcal{A}}$ since the cylindrical σ -algebra of $\mathbf{R}^{\mathcal{A}}$ only harbors events concerning at most countably many values of X at the same time. It is thus not proper to ask questions such that: "What is the probability of X being continuous?". To our knowledge, there are two (mainly equivalent) ways to circumvent this issue:

1. Once X is defined on $\mathbf{R}^{\mathcal{A}}$ through its finite-dimensional distributions, find a version with better sample paths by means of Kolmogorov-Čentsov or a like-minded theorem.
2. Directly define X as a random object in a functional space \mathcal{H} having nicer topological properties than $\mathbf{R}^{\mathcal{A}}$.

As pointed out, those approaches do not differ much in the mathematical details and mainly rely on building smooth enough discretizations $X_n \in \mathcal{H}$ of X and show two things:

1. The distribution of X_n — as a probability measure \mathbf{P}_{X_n} on \mathcal{H} — converges weakly to some distribution. To this effect, Prohorov's theorem (see e.g. [18, Theorems 5.1 and 5.2]) gives a practical equivalent condition to the relative compactness of $\{\mathbf{P}_{X_n} : n \in \mathbf{N}\}$ called *tightness* provided that \mathcal{H} is Polish (i.e. a complete and separable metric space).
2. The finite-dimensional distributions of X_n converge to those of X . Combined with the previous condition, this would imply that X indeed has sample paths in \mathcal{H} , provided that the Borel σ -algebra of \mathcal{H} corresponds to the trace of the cylindrical σ -algebra of $\mathbf{R}^{\mathcal{A}}$ on \mathcal{H} .

Following this philosophy, we introduce in Sections 3.2 and 3.3 functional spaces where a lot of processes introduced in Chapter 2 live. Once a version with well-defined sample paths is secured, we may begin to talk about pathwise Hölder regularity, which we do in Section 3.4. We then apply all of the above in Section 3.5 to study the regularity of the integral of a deterministic function $f : \mathcal{T} \rightarrow \mathbf{R}$ with respect to a siLévy X .

3.2 The continuous space $\mathcal{C}^\Phi(\mathcal{A})$

For a start, we introduce a space of 'locally continuous' functions. The main goal is to present ideas that will carry over to Section 3.3 where we tackle càdlàg functions.

3.2.1 Projective limit of Polish spaces

We start by giving a general result about Polish spaces. Although we suspect it being already well-known, we could not find any reference stating exactly what we wished for. This result will later be applied to see several functional spaces E as Polish spaces, a condition craved by probabilists when they want to look at E -valued random variables and their weak convergence. For a marvelous treatment of the general theory, we refer to [18].

THEOREM 3.2.1 (Projective limit of Polish spaces). *Consider a projective system of Polish spaces, i.e. a set E and sequences $(E_n)_{n \in \mathbf{N}}$, $(\pi_n)_{n \in \mathbf{N}}$ and $(\pi_m^n)_{m \leq n}$ such that the following holds:*

1. For all $n \in \mathbf{N}$, E_n is a Polish space and the maps $\pi_m^n : E_n \rightarrow E_m$ ($m \leq n$) are continuous.
2. (Identification). The maps $\pi_n : E \rightarrow E_n$ ($n \in \mathbf{N}$) are such that the map $E \ni x \mapsto (\pi_n(x))_{n \in \mathbf{N}} \in \prod_{n \in \mathbf{N}} E_n$ is injective.

3. (Compatibility). For all $l \leq m \leq n$, the following diagrams commute:

$$\begin{array}{ccc} E & & \\ \pi_n \downarrow & \searrow \pi_m & \\ E_n & \xrightarrow{\pi_m^n} & E_m \end{array} \quad \begin{array}{ccccc} & & \pi_l^n & & \\ & & \curvearrowright & & \\ E_n & \xrightarrow{\pi_m^n} & E_m & \xrightarrow{\pi_l^n} & E_l. \end{array}$$

Then, the set E is Polish for the topology induced by metric

$$\forall x, y \in E, \quad d(x, y) = \sum_{n=0}^{\infty} 2^{-n} (d_n(\pi_n(x), \pi_n(y)) \wedge 1) \quad (3.2.1)$$

where for all $n \in \mathbf{N}$, d_n is any metric on E_n that induces its topology.

Proof. By identification, E may be isometrically embedded into $\prod_{n \in \mathbf{N}} E_n$ as

$$E \simeq \left\{ x \in \prod_{n \in \mathbf{N}} E_n : \forall m \leq n, \pi_m^n(x_n) = x_m \right\}. \quad (3.2.2)$$

By [22, Chapitre IX, §6, Proposition 1], we already know that $\prod_{n \in \mathbf{N}} E_n$ is Polish and that we only need to show that E is a closed subset of $\prod_{n \in \mathbf{N}} E_n$. Rewriting (3.2.2) yields

$$E \simeq \bigcap_{m \leq n} (\pi_n \circ \pi_m^n, \pi_m)^{-1}(\Delta_m)$$

where for all $m \in \mathbf{N}$, $\Delta_m = \{(x, x) : x \in E_m\}$ is the diagonal in E_m . Since all the projections are continuous — either by definition or by construction — and the diagonals are closed, E is closed. The result follows. \square

3.2.2 The space $C^\Phi(\mathcal{A})$ and its topology

Denote for all $n \in \mathbf{N}$,

$$\mathcal{T}_n = [\emptyset, \mathcal{T}_n] = \{A \in \mathcal{A} : A \subseteq \mathcal{T}_n\} \quad (3.2.3)$$

where we recall that $\mathcal{T}_n = \bigcup_{A \in \mathcal{A}_n} A \in \mathcal{A}(u)$. By Proposition 1.4.10, we know that $(\mathcal{T}_n, d_{\mathcal{A}})$ is compact for all $n \in \mathbf{N}$. In particular, the space $C^\Phi(\mathcal{T}_n) = C^0(\mathcal{T}_n; \mathbf{R})$ of continuous maps from \mathcal{T}_n to \mathbf{R} endowed with the norm given by

$$\forall h \in C^\Phi(\mathcal{T}_n), \quad \|h\|_{\infty, \mathcal{T}_n} = \sup_{A \in \mathcal{T}_n} |h(A)|$$

is a separable Banach space (see [9, Lemmas 3.97 and 3.99]). In turn, this enables us to endow the space

$$C^\Phi(\mathcal{A}) = \left\{ h \in \mathbf{R}^{\mathcal{A}} : \forall n \in \mathbf{N}, h|_{\mathcal{T}_n} \in C^\Phi(\mathcal{T}_n) \right\} \quad (3.2.4)$$

with the projective limit topology given by Theorem 3.2.1 applied to the sequence $(C^\Phi(\mathcal{T}_n))_{n \in \mathbf{N}}$, the maps $\pi_m^n = (\cdot)|_{\mathcal{T}_m}$ and $\pi_n = (\cdot)|_{\mathcal{T}_n}$ ($m \leq n$). In particular, $C^\Phi(\mathcal{A})$ is a Polish space. As to why we used ‘ Φ ’ in the notation, we refer to the coming Section 3.2.4.

Not all functions in $C^\Phi(\mathcal{A})$ are continuous. Indeed, even though \mathcal{A} is σ -compact (Proposition 1.4.10), it is not locally compact in general. This has already been illustrated in Example 1.4.11. We will make further use of it in Example 3.2.3 to give a process which has sample paths in $C^\Phi(\mathcal{A})$ while not being continuous.

3.2.3 Some processes living in $\mathcal{C}^\Phi(\mathcal{A})$

We give some instances when Gaussian processes have sample paths in $\mathcal{C}^\Phi(\mathcal{A})$. Examples and counter-examples are then given.

PROPOSITION 3.2.2. *Suppose that $d_{\mathcal{A}} = d_m$ and $\dim \mathcal{A} < \infty$. Then, the following processes have a version in $\mathcal{C}^\Phi(\mathcal{A})$: the siBm W^H for any $H \in (0, 1/2]$, the siBb \widehat{W} and if $f : \mathcal{T} \rightarrow \mathbf{R}$ is a measurable map such that for all $n \in \mathbf{N}$, $\|f\|_{\infty, \mathcal{T}_n} = \sup_{t \in \mathcal{T}_n} |f(t)| < \infty$, the primitive process of f with respect to the siBm W , i.e. the process Y given by*

$$\forall A \in \mathcal{A}, \quad Y_A = \int_A f dW.$$

Proof. Consider $H \in (0, 1/2]$, $n \in \mathbf{N}$ and let us show that $W^H|_{\mathcal{T}_n}$ has a version in $\mathcal{C}^\Phi(\mathcal{T}_n)$.

Denote by \mathcal{H} the log-entropy (Definition 1.5.11) of (\mathcal{T}_n, d_H) where for all $A, A' \in \mathcal{A}$, $d_H(A, A') = \|W_A^H - W_{A'}^H\|_{L^2(\Omega)}$. Using the covariance of W^H , we have for all $A, A' \in \mathcal{A}$,

$$\begin{aligned} \mathbf{E}[(W_A^H - W_{A'}^H)^2] &= \text{Var}(W_A^H) + \text{Var}(W_{A'}^H) - 2\text{Cov}(W_A^H, W_{A'}^H) \\ &= \mathbf{m}(A)^{2H} + \mathbf{m}(A')^{2H} - (\mathbf{m}(A)^{2H} + \mathbf{m}(A')^{2H} - \mathbf{m}(A\Delta A')^{2H}) \\ &= \mathbf{m}(A\Delta A')^{2H}. \end{aligned}$$

Hence $d_H = (d_m)^H$. Using this relation and the estimate on the log-entropy of (\mathcal{T}_n, d_m) from Theorem 1.5.12, we get

$$\int_0^1 \mathcal{H}^{1/2}(\varepsilon) d\varepsilon < \infty.$$

Thus, by a classical result from Dudley (see [28, Corollary 2.3] or [4, Theorem 1.3.5]), $W^H|_{\mathcal{T}_n}$ has a continuous version. Hence

$$\mathbf{P}(\forall n \in \mathbf{N}, W^H|_{\mathcal{T}_n} \in \mathcal{C}^\Phi(\mathcal{T}_n)) = 1$$

which proves that W has a version in $\mathcal{C}^\Phi(\mathcal{A})$.

The version for the siBb \widehat{W} is then just a consequence of the one for the siBm and its representation (2.5.20) in terms of a siBm. As for the primitive process Y , the hypothesis on f enables us to show that for all $n \in \mathbf{N}$, we have

$$\forall A, A' \in \mathcal{T}_n, \quad \mathbf{E}[(Y_A - Y_{A'})^2] = \int_{A\Delta A'} f^2 d\mathbf{m} \leq \|f\|_{\infty, \mathcal{T}_n}^2 d_m(A, A')$$

which may then be used to do the same entropy argument as we did for W^H . \square

EXAMPLE 3.2.3 (Why Proposition 3.2.2 fails for $\mathcal{C}^0(\mathcal{A}; \mathbf{R})$). *One of the reasons why we considered $\mathcal{C}^\Phi(\mathcal{A})$ instead of $\mathcal{C}^0(\mathcal{A}; \mathbf{R})$ is because even the siBm does not have continuous sample paths in general. It is known for some specific indexing collections [4, Theorem 1.4.5], but let us build another example.*

Consider the continuous tree $\mathcal{T} \simeq \{0_\tau\} \sqcup ((0, 1] \times \mathbf{N})$ from Example 1.4.11 endowed with its 'Lebesgue' measure \mathbf{m} . Remark that the metric d_m actually corresponds to the SNCF metric considered in that very example. Moreover, checking that $\dim \mathcal{A} = 2$ is straightforward.

Let us show that a siBm W cannot have a version in $C^0(\mathcal{A}; \mathbf{R})$. For all $j \in \mathbf{N}$, denote $A_j = A(\sigma_j^2, j) \in \mathcal{A}$ for some $\sigma_j \in (0, 1]$ to be determined. To be more wordy about A_j , it is the segment that lies on the j^{th} 'branch' of \mathcal{T} , of which $0_{\mathcal{T}}$ is an extremity and has length σ_j^2 .

Since for all integers $j \neq k$, $\mathbf{m}(A_j \cap A_k) = \mathbf{m}(\{0_{\mathcal{T}}\}) = 0$, the W_{A_j} 's are independent Gaussian variables. According to Borel's 0-1 law, we get

$$\limsup_{j \rightarrow \infty} W_{A_j} > 1 \text{ a.s.} \iff \sum_{j=0}^{\infty} \mathbf{P}(W_{A_j} > 1) = \infty. \quad (3.2.5)$$

A quick study of function shows that for all $u \in \mathbf{R}_+$, $e^{-u^2/2} \geq u^2 e^{-u^3}$. Integrating this inequality yields

$$\forall x \in \mathbf{R}_+, \int_x^{+\infty} e^{-u^2/2} du \geq \frac{1}{3} e^{-x^3}.$$

In particular, since for all $j \in \mathbf{N}$, $W_{A_j} \sim N(0, \sigma_j^2)$, we get

$$\forall j \in \mathbf{N}, \mathbf{P}(W_{A_j} > 1) \geq \frac{1}{3\sqrt{2\pi}} e^{-\sigma_j^{-3}}.$$

Taking for instance $\sigma_j = (\ln j)^{-1/3}$, we obtain the divergence of the series in (3.2.5). Thus $\limsup_{j \rightarrow \infty} W_{A_j} > 1$ a.s. Moreover, since $d_{\mathbf{m}}(A_j, \emptyset) = \sigma_j^2 \rightarrow 0$ as $j \rightarrow \infty$ and $W_{\emptyset} = 0$ a.s., we obtain the a.s. discontinuity of W at \emptyset .

EXAMPLE 3.2.4 (Why Proposition 3.2.2 may fail when $\dim \mathcal{A} = \infty$). Assuming that \mathcal{A} is finite dimensional is also not superfluous. We give here an example for which $\dim \mathcal{A} = \infty$ and where the siBm has no version in $C^\Phi(\mathcal{A})$.

Consider the Hilbert's cube $\mathcal{T} = \prod_{j \in \mathbf{N}^*} [0, a_j]$ where for all $j \in \mathbf{N}^*$, $a_j = j^{-3/2}$. By Proposition 1.2.9, we know that \mathcal{T} may be endowed with the product indexing collection \mathcal{A} . In that case, since $\mathcal{T} \in \mathcal{A}$, we have $C^\Phi(\mathcal{A}) = C^0(\mathcal{A}; \mathbf{R})$. Let us define the measure \mathbf{m} by

$$\mathbf{m} = \sum_{j=1}^{\infty} \delta_0^{\otimes j-1} \otimes \text{Leb}|_{[0, a_j]} \otimes \delta_0 \otimes \delta_0 \otimes \dots$$

Basically, \mathbf{m} only charges the 'coordinate axes' of \mathcal{T} on which it corresponds to the usual Lebesgue measure. In particular, we have

$$\forall A \in \mathcal{A}, \mathbf{m}(A) = \sum_{j=1}^{\infty} \mathbf{t}(A)_j \quad (3.2.6)$$

where $\mathbf{t}(A)_j \in [0, a_j]$ is the j^{th} component of the tip $\mathbf{t}(A)$ of A (Definition 1.2.1). Using (3.2.6), one may readily check that a necessary and sufficient condition for \mathbf{m} to be compatible (Definition 1.3.1) is for the series $\sum_j a_j$ to converge, which happens to be true for our choice of $(a_j)_j$.

Let us consider a sequence $(W^j)_{j \in \mathbf{N}^*}$ of independent \mathbf{R}_+ -indexed Brownian motions. A straightforward computation shows that the set-indexed process $W = \{W_A : A \in \mathcal{A}\}$ given by

$$\forall A \in \mathcal{A}, W_A = \sum_{j=1}^{\infty} W_{\mathbf{t}(A)_j}^j \quad (3.2.7)$$

is well-defined as a limit in $L^2(\Omega)$ and actually is a siBm with control measure \mathbf{m} .

Let us show that W is not continuous at \emptyset for the metric $d_{\mathcal{A}} = d_m$ (which, due to (3.2.6), corresponds to the ℓ^1 -metric). For that, consider a \mathbf{R}_+ -indexed Brownian motion B and define

$$\tau = \sup \left\{ t > 0 : \forall s \leq t, \max_{0 \leq u \leq s} B_u > s^{2/3} \right\} \quad (3.2.8)$$

together with the convention $\sup \emptyset = +\infty$. Due to the asymptotic behavior of B at 0 (e.g. [30, Theorem 8.8.1]) we know that $\mathbf{P}(\tau > 0) = 1$. In particular, we may find $\delta > 0$ such that

$$\mathbf{P}(\tau > \delta) > 1/2. \quad (3.2.9)$$

We introduce the following

$$\begin{aligned} \forall j \in \mathbf{N}^*, \quad \tau^j &= \sup \left\{ t > 0 : \forall s \leq t, \max_{0 \leq u \leq s} W_u^j > s^{2/3} \right\} \\ \Omega^* &= \bigcap_{k \in \mathbf{N}^*} \liminf_{n \rightarrow \infty} \left\{ \sum_{j=k}^{2^n} \mathbb{1}_{\{\max_{0 \leq s \leq a_j} W_s^j > a_j^{2/3}\}} > 2^{n-1} \right\}. \end{aligned}$$

Since the τ^j 's are iid copies of τ , we may apply the strong law of large numbers to get

$$\frac{1}{n} \sum_{j=1}^n \mathbb{1}_{\{\tau^j > \delta\}} \xrightarrow{\text{a.s.}} \mathbf{P}(\tau > \delta) \quad \text{as } n \rightarrow \infty.$$

Hence, cutting the sum at some k and extracting along the subsequence $(2^n)_{n \in \mathbf{N}}$, the following holds with probability one:

$$\forall k \in \mathbf{N}^*, \quad 2^{-n} \sum_{j=k}^{2^n} \mathbb{1}_{\{\tau^j > \delta\}} \longrightarrow \mathbf{P}(\tau > \delta) \quad \text{as } n \rightarrow \infty. \quad (3.2.10)$$

Since $\sum_{j \in \mathbf{N}^*} a_j < \infty$, there exists $j_0 \in \mathbf{N}^*$ such that for all $j \geq j_0$, $a_j < \delta$. In particular, we get

$$\forall j \geq j_0, \quad \{\tau^j > \delta\} \subseteq \left\{ \max_{0 \leq s \leq a_j} W_s^j > a_j^{2/3} \right\}. \quad (3.2.11)$$

Using (3.2.10) and (3.2.11), the following holds with probability one:

$$\forall k \in \mathbf{N}^*, \quad \liminf_{n \rightarrow \infty} \left(2^{-n} \sum_{j=k}^{2^n} \mathbb{1}_{\{\max_{0 \leq s \leq a_j} W_s^j > a_j^{2/3}\}} \right) \geq \mathbf{P}(\tau > \delta). \quad (3.2.12)$$

Using (3.2.9) and (3.2.12), we get $\mathbf{P}(\Omega^*) = 1$.

We now wish to exhibit a sequence $(A^k)_{k \in \mathbf{N}}$ in \mathcal{A} converging to \emptyset such that

$$\forall \omega \in \Omega^*, \quad \limsup_{k \rightarrow \infty} \sup_{A \in [\emptyset, A^k]} W_A(\omega) > \kappa \text{ a.s.} \quad (3.2.13)$$

for some constant $\kappa > 0$. Since $\mathbf{P}(\Omega^*) = 1$, this will immediately imply that W has almost surely discontinuous sample paths at \emptyset .

For all $k \in \mathbf{N}^*$, define $t^k = (0, \dots, 0, a_k, a_{k+1}, \dots) \in \mathcal{T}$ and $A^k = A(t^k)$. Due to (3.2.6) and the fact that $\sum_{j \in \mathbf{N}^*} a_j < \infty$, we get

$$d_m(A^k, \emptyset) = \sum_{j=k}^{\infty} a_j \longrightarrow 0 \quad \text{as } k \rightarrow \infty. \quad (3.2.14)$$

Fix $k \in \mathbf{N}^*$. We have

$$\begin{aligned} \sup_{A \in [\emptyset, A^k]} W_A &= \sup_{s \in A^k} \sum_{j=k}^{\infty} W_{s_j}^j \\ &= \sum_{j=k}^{\infty} \max_{0 \leq s_j \leq a_j} W_{s_j}^j \\ &\geq \sum_{j=k}^{\infty} a_j^{2/3} \mathbf{1}_{\{\max_{0 \leq s \leq a_j} W_s^j > a_j^{2/3}\}}. \end{aligned}$$

Hence

$$\forall n \in \mathbf{N}^*, \quad \sup_{A \in [\emptyset, A^k]} W_A \geq \sum_{j=k}^{2^n} a_j^{2/3} \mathbf{1}_{\{\max_{0 \leq s \leq a_j} W_s^j > a_j^{2/3}\}}. \quad (3.2.15)$$

Let $\omega \in \Omega^*$. By definition of Ω^* , there exists $n_k(\omega) \in \mathbf{N}^*$ such that

$$\sum_{j=k}^{2^{n_k(\omega)}} \mathbf{1}_{\{\max_{0 \leq s \leq a_j} W_s^j(\omega) > a_j^{2/3}\}} > 2^{n_k(\omega)-1}. \quad (3.2.16)$$

Hence (3.2.16) implies that at least 2^{n_k-1} terms in the sum from (3.2.15) are non-zero. Since $a_j^{2/3} = 1/j$ is decreasing in j , we may take the last 2^{n_k-1} as a lower bound, i.e.

$$\sup_{A \in [\emptyset, A^k]} W_A(\omega) \geq \sum_{j=2^{n_k(\omega)-1}}^{2^{n_k(\omega)}} j^{-1} \geq \ln 2,$$

which implies (3.2.13). Hence W has almost surely discontinuous (and also unbounded actually) sample paths at \emptyset , which explains why we supposed $\dim \mathcal{A} < \infty$ in Proposition 3.2.2.

EXAMPLE 3.2.5 (Why the conditions in Proposition 3.2.2 are not sharp). We should however mention that the previous example may be tweaked so that the siBm W on the Hilbert's cube has a version in $C^\Phi(\mathcal{A})$ even though we still have $\dim \mathcal{A} = \infty$. To show that, we keep all the notations from Example 3.2.4 apart from the sequence $(a_j)_{j \in \mathbf{N}^*}$ that will be chosen later on.

Let us show that W almost surely has continuous sample paths. Denote the function ϖ known as Lévy's modulus of continuity by

$$\varpi(0) = 0 \quad \text{and} \quad \forall \delta \in (0, e^{-1}], \quad \varpi(\delta) = \sqrt{2\delta \ln(1/\delta)}$$

where $[0, e^{-1}]$ is the interval on which ϖ is increasing. By [30, Theorem 8.4.2], we have for the one-dimensional Brownian motion B ,

$$\sup_{\substack{u, v \in [0, 1]: \\ |u-v| \leq \delta}} \frac{|B_u - B_v|}{\varpi(\delta)} \xrightarrow{a.s.} 1 \quad \text{as} \quad \delta \rightarrow 0^+$$

Hence, there exists a random variable such that $\mathbf{P}(\eta > 0) = 1$ and the following holds with probability one:

$$\forall u, v \in [0, 1], \quad |u - v| \leq \eta \implies |B_u - B_v| \leq 2\varpi(|s - t|). \quad (3.2.17)$$

Since the W^j 's are all iid copies of B , we may consider a sequence $(\eta^j)_{j \in \mathbf{N}^*}$ which are iid copies of η such that (3.2.17) holds for W^j and η^j instead of B and η . Since for all $j \in \mathbf{N}^*$, $\mathbf{P}(\eta^j > 0) = 1$, we

may choose $(a_j)_{j \in \mathbf{N}^*}$ in $(0, e^{-1}]$ converging sufficiently fast towards 0 so that

$$\sum_{j=1}^{\infty} a_j < \infty, \quad (3.2.18)$$

$$\sum_{j=1}^{\infty} \varpi(a_j) < \infty, \quad (3.2.19)$$

$$\sum_{j=1}^{\infty} \mathbf{P}(\eta^j < a_j) < \infty. \quad (3.2.20)$$

The condition (3.2.18) ensures that all the processes are well-defined (see Example 3.2.4). By Borel-Cantelli's lemma applied to (3.2.20), there exists a random integer n_0 such that the following holds with probability one:

$$\forall j > n_0, \quad a_j \leq \eta^j. \quad (3.2.21)$$

Let $\varepsilon > 0$. We have

$$\forall A(s), A(t) \in \mathcal{A}, \quad \forall n \geq n_0, \quad |W_{A(s)} - W_{A(t)}| \leq \sum_{j=1}^n |W_{s_j}^j - W_{t_j}^j| + \sum_{j=n+1}^{\infty} |W_{s_j}^j - W_{t_j}^j|.$$

Thus, according to (3.2.17) and (3.2.21), the following holds with probability one:

$$\forall A(s), A(t) \in \mathcal{A}, \quad \forall n \geq n_0, \quad |W_{A(s)} - W_{A(t)}| \leq \sum_{j=1}^n |W_{s_j}^j - W_{t_j}^j| + \sum_{j=n+1}^{\infty} \varpi(a_j). \quad (3.2.22)$$

By (3.2.19), we may find a (random) integer $n \geq n_0$ big enough so that $\sum_{j>n+1} \varpi(a_j) < \varepsilon/2$ a.s. Putting that back into (3.2.22), we get that the following holds with probability one:

$$\forall A(s), A(t) \in \mathcal{A}, \quad \forall n \geq n_0, \quad |W_{A(s)} - W_{A(t)}| \leq \sum_{j=1}^n |W_{s_j}^j - W_{t_j}^j| + \frac{\varepsilon}{2}. \quad (3.2.23)$$

Since for all $j \leq n$, the maps $\mathcal{A} \ni A(t) \mapsto t_j \in [0, a_j]$ and W^j are continuous, (3.2.23) actually shows that

$$\mathbf{P}(\Omega_\varepsilon) = 1 \quad \text{where} \quad \Omega_\varepsilon = \bigcup_{\delta > 0} \bigcap_{\substack{A, A' \in \mathcal{A}: \\ d_m(A, A') \leq \delta}} \{|W_{A(s)} - W_{A(t)}| \leq \varepsilon\}.$$

Hence $\mathbf{P}\left(\bigcap_{\varepsilon \in \mathbf{Q}_+^*} \Omega_\varepsilon\right) = 1$, i.e. W almost surely has continuous sample paths.

3.2.4 A link with flows

Up to this point, some might wonder why we chose to denote $\mathcal{C}^\Phi(\mathcal{A})$ with a ‘ Φ ’ in exponent, notation reserved so far for geodesic flows. This section is devoted to unearth the link between $\mathcal{C}^\Phi(\mathcal{A})$ and geodesic flows. This link will then be exploited in Section 3.3 to define a càdlàg counterpart to $\mathcal{C}^\Phi(\mathcal{A})$. As hinted in Section 1.3.2, we will mainly use the metric d_m since it is the one that works best with flows.

LEMMA 3.2.6. *For all $n \in \mathbf{N}$, the space $\Phi(\mathcal{T}_n)$ is compact once endowed with the topology of uniform convergence. In particular, the space $\mathcal{C}^0(\Phi(\mathcal{T}_n); \mathcal{C}^0([0, 1]; \mathbf{R}))$ is Polish for the topology of uniform convergence.*

REMARK 3.2.7. Theorem 3.2.1 and Lemma 3.2.6 could be used to endow $\Phi(\mathcal{A})$ with a Polish space structure by means of the projections defined as follows. For $n \in \mathbf{N}$ and $\phi \in \Phi(\mathcal{A})$, denote $v(\pi_n \phi) = m(\phi(1) \cap \mathcal{T}_n)$ and

$$\forall t \in [0, 1], \quad \pi_n \phi(t) = \begin{cases} \bigcap_{\substack{s \in [0, 1]: \\ m(\phi(s) \cap \mathcal{T}_n) > v(\pi_n \phi)t}} (\phi(s) \cap \mathcal{T}_n) & \text{if } t < 1, \\ \phi(1) \cap \mathcal{T}_n & \text{if } t = 1 \end{cases}$$

where we check that $\pi_n \phi$ is indeed a geodesic \mathcal{T}_n -flow with speed $v(\pi_n)$. The projections π_m^n are then defined similarly.

Proof of Lemma 3.2.6. Let $n \in \mathbf{N}$. Since $\Phi(\mathcal{T}_n)$ is a subset of $C^0([0, 1]; \mathcal{T}_n)$, we will be using Arzelà-Ascoli theorem. We refer for instance to [63, Theorem 3.1] and even though the functions must be Banach-valued in the statement, it is not an issue here since \mathcal{T}_n embeds itself isometrically into $L^1(\mathbf{m})$ by (1.3.10). Thus we need to show that for all $t \in [0, 1]$, $\{\phi(t) : \phi \in \Phi(\mathcal{T}_n)\}$ is totally bounded and that $\Phi(\mathcal{T}_n)$ is equicontinuous and closed.

First, we have for all $t \in \mathcal{T}$, $\{\phi(t) : \phi \in \Phi(\mathcal{T}_n)\} \subseteq \mathcal{T}_n$, which is compact by Proposition 1.4.10. Moreover, for all $s, t \in [0, 1]$ and $\phi \in \Phi(\mathcal{T}_n)$, we have

$$\begin{aligned} d_m(\phi(s), \phi(t)) &= v(\phi) |s - t| \\ &= d_m(\phi(0), \phi(1)) |s - t| \\ &\leq \text{diam}(\mathcal{T}_n) |s - t| \end{aligned}$$

which settles the equicontinuity. Hence $\Phi(\mathcal{T}_n)$ is totally bounded.

What is left is to show that $\Phi(\mathcal{T}_n)$ is closed. Consider a sequence $(\phi_k)_{k \in \mathbf{N}}$ converging to some $\phi \in C^0([0, 1]; \mathcal{T}_n)$. Since for all $k \in \mathbf{N}$, $v(\phi_k) = d_m(\phi_k(0), \phi_k(1))$ and for all $s, t \in [0, 1]$, $d_m(\phi_k(s), \phi_k(t)) \rightarrow d_m(\phi(s), \phi(t))$ as $k \rightarrow \infty$, we get

$$\forall s, t \in [0, 1], \quad d_m(\phi(s), \phi(t)) = v(\phi) |s - t| \quad (3.2.24)$$

where $v(\phi) = d_m(\phi(0), \phi(1))$. The problem is that ϕ might be neither non-decreasing nor outer continuous. So we introduce the map

$$\psi : [0, 1] \longrightarrow \begin{cases} \mathcal{T}_n & \\ t \longmapsto \begin{cases} \bigcap_{\substack{s \in \mathbf{Q} \in [0, 1]: \\ s > t}} \phi(s) & \text{if } s < 1, \\ \phi(1) & \text{if } s = 1. \end{cases} \end{cases}$$

By construction, ψ is an outer continuous \mathcal{T}_n -flow and

$$\forall t \in [0, 1], \quad \psi(t) = \lim_{s \in \mathbf{Q} \rightarrow t^+} \phi(s) \quad (3.2.25)$$

which implies that (3.2.24) also holds for ψ , so $\psi \in \Phi(\mathcal{T}_n)$. Moreover, (3.2.25) also implies that

$$\forall t \in [0, 1], \quad d_m(\phi(t), \psi(t)) = \lim_{s \in \mathbf{Q} \rightarrow t^+} d_m(\phi(t), \phi(s)) = 0.$$

Hence ϕ_k also converges to ψ uniformly as $k \rightarrow \infty$, so $\Phi(\mathcal{T}_n)$ is closed. Hence $\Phi(\mathcal{T}_n)$ is compact. It then readily follows from [9, Lemma 3.99] that $C^0(\Phi(\mathcal{T}_n); C^0([0, 1]; \mathbf{R}))$ is Polish. \square

The link between $C^\Phi(\mathcal{A})$ and flows stems from the fact that any map $h \in C^\Phi(\mathcal{A})$ may be characterized by its projections along geodesic \mathcal{A} -flows, *i.e.* the maps given for all $\phi \in \Phi(\mathcal{A})$ by

$$h^\phi : [0, 1] \ni t \longmapsto h(\phi(t)) \in \mathbf{R}.$$

In particular, Proposition 3.2.8 below identifies functions in $C^\Phi(\mathcal{A})$ with a subset of $C^0(\Phi(\mathcal{A}) \times [0, 1]; \mathbf{R}) \simeq C^0(\Phi(\mathcal{A}); C^0([0, 1]; \mathbf{R}))$. Using Lemma 3.2.6, we endow this space with the projective limit topology given by Theorem 3.2.1 in a fashion similar to $C^\Phi(\mathcal{A})$ in Section 3.2.2.

PROPOSITION 3.2.8 (Embedding of $C^\Phi(\mathcal{A})$ by projection along flows). *The map*

$$\begin{aligned} C^\Phi(\mathcal{A}) &\longrightarrow C^0(\Phi(\mathcal{A}); C^0([0, 1]; \mathbf{R})) \\ h &\longmapsto [\phi \mapsto h^\phi] \end{aligned}$$

is linear, continuous and injective. Moreover, it is an isometry for the metrics given by (3.2.1).

Proof. This map is well-defined since the evaluation map

$$\begin{aligned} \Phi(\mathcal{A}) \times [0, 1] &\longrightarrow \mathcal{A} \\ (\phi, t) &\longmapsto \phi(t) \end{aligned}$$

is continuous. Linearity is obvious. Injectivity as well since any $A \in \mathcal{A}$ may be written as $A = \phi(1)$ for some $\phi \in \Phi(\mathcal{A})$ by Proposition 1.3.9. As for the isometry property, it is enough to prove the following:

$$\forall n \in \mathbf{N}, \forall h \in C^\Phi(\mathcal{T}_n), \quad \|h\|_{\infty, \mathcal{T}_n} = \max_{(\phi, t) \in \Phi(\mathcal{T}_n) \times [0, 1]} |h^\phi(t)|.$$

Let $n \in \mathbf{N}$ and $h \in C^\Phi(\mathcal{T}_n)$. By definition, we have $\|h\|_{\infty, \mathcal{T}_n} \geq \max_{(\phi, t) \in \Phi(\mathcal{T}_n) \times [0, 1]} |h^\phi(t)|$. Since \mathcal{T}_n is compact, we may find $A \in \mathcal{A}$ such that $h(A) = \|h\|_{\infty, \mathcal{T}_n}$. The converse inequality then follows from the same ‘trick’ used for injectivity. \square

3.3 The càdlàg space $D^\Phi(\mathcal{A})$

As in the one-dimensional case, the space of continuous functions does not encompass every process of interest. In particular, the space $\mathcal{D}(0, 1)$ of càdlàg — right continuous with left limits — functions accomodates both continuous and point processes, and so Lévy processes usually prefer to live there.

Thus it becomes interesting to turn $\mathcal{D}(0, 1)$ into a nice enough space. With this goal in mind, Skorokhod [92] endowed $\mathcal{D}(0, 1)$ with four different topologies: J_1 , J_2 , M_1 and M_2 . Straf [96] adapted the J_1 -topology to the multiparameter case while for the set-indexed case, Bass and Pyke [17] took care of the M_2 -topology and Slonowsky [93] developed the J_2 -topology.

Here, we propose an alternative to Slonowsky’s space for set-indexed maps where we favor the J_1 - over the M_2 -topology and the metric d_m over the Hausdorff metric d_H .

This actually partially answers a concern raised in [17, §6.4]. Indeed, one could try to apply the program proposed in [96] to endow set-indexed càdlàg maps with J_1 -topology. However, one would then need to build a reasonable class of homeomorphisms serving as ‘time changes’, which proves to be a challenge. Our construction circumvents this issue.

3.3.1 A quick review of $\mathcal{D}(0, 1)$

We recall here the main properties of the space

$$\mathcal{D}(0, 1) = \left\{ h \in \mathbf{R}^{[0,1]} : \forall t \in [0, 1], h(t^-) \text{ and } h(t^+) \text{ are well-defined and } h(t^+) = h(t) \right\}.$$

and its J_1 -topology (the favorite one, or so it seems). For a more thorough account, we refer to [18, Chapter 3] or [54, Chapter 16].

The uniform topology is a bad deal for $\mathcal{D}(0, 1)$ for two reasons:

- ◊ The (uncountable) family of càdlàg functions $\{\mathbb{1}_{[t,1]} : t \in [0, 1]\}$ is such that $s \neq t$ implies $\|\mathbb{1}_{[s,1]} - \mathbb{1}_{[t,1]}\|_\infty = 1$. In particular, $\mathcal{D}(0, 1)$ cannot be separable under the uniform topology, which is a property required to be Polish.
- ◊ $\mathbb{1}_{[t_n,1]}$ converges uniformly to $\mathbb{1}_{[t,1]}$ as $n \rightarrow \infty$ only if $t_n = t$ for all n big enough. Intuitively, we would like càdlàg functions with jumps of similar sizes placed at similar locations to be close to each other for the topology in $\mathcal{D}(0, 1)$.

Hence the need to weaken the rigid uniform topology and allow some ‘wiggling’ time-wise.

DEFINITION 3.3.1 (Class of ‘time changes’ in $[0, 1]$). *Denote by Λ the group made of all the increasing bijections $\lambda : [0, 1] \rightarrow [0, 1]$. Introduce the following ‘group norms’ defined for all $\lambda \in \Lambda$ by*

$$\begin{aligned} \|\lambda\|_\Lambda &= \sup_{t \in [0,1]} |\lambda(t) - t|, \\ \|\lambda\|'_\Lambda &= \sup_{t \in [0,1]} |\lambda(t) - t| + \sup_{s \neq t \in [0,1]} \left| \log \frac{|\lambda(s) - \lambda(t)|}{|s - t|} \right|. \end{aligned}$$

Both of those quantities express ‘how far’ a given $\lambda \in \Lambda$ is to the identity. They both yield metrics on $\mathcal{D}(0, 1)$ given for all $g, h \in \mathcal{D}(0, 1)$ by

$$d_{\mathcal{D}(0,1)}(g, h) = \inf_{\lambda \in \Lambda} \max \{ \|\lambda\|_\Lambda, \|g - h \circ \lambda\|_\infty \}, \quad (3.3.1)$$

$$d'_{\mathcal{D}(0,1)}(g, h) = \inf_{\lambda \in \Lambda} \max \{ \|\lambda\|'_\Lambda, \|g - h \circ \lambda\|_\infty \}. \quad (3.3.2)$$

It is well-known that those metrics induce the same topology on $\mathcal{D}(0, 1)$, called J_1 -topology. Unfortunately, the simpler metric $d_{\mathcal{D}(0,1)}$ does not turn $\mathcal{D}(0, 1)$ into a complete metric space, but $d'_{\mathcal{D}(0,1)}$ does. For an insightful exposition on the subject — formulated in greater generality — we refer to [96, Sections 1-3]. Let us just mention the fact that the issues raised above are solved for this topology. In particular, using a piecewise affine time change, one may show that

$$\forall s, t \in [0, 1], \quad d_{\mathcal{D}(0,1)}(\mathbb{1}_{[s,1]}, \mathbb{1}_{[t,1]}) = |s - t|. \quad (3.3.3)$$

3.3.2 Blind spots of geodesic flows

Mimicking the idea behind Proposition 3.2.8, we want to define the càdlàg space $\mathcal{D}^\Phi(\mathcal{A})$ as isometrically embedded into $\mathcal{C}^0(\Phi(\mathcal{A}); \mathcal{D}(0, 1))$ thanks to the map $\phi \mapsto h^\phi$. We will see that we need a couple of ideas to make that work.

EXAMPLE 3.3.2 (The need for maximal flows). *A first issue is that two geodesic flows may not be able to ‘see’ the same things, even while being close to one another. This is not a problem in $\mathcal{C}^\Phi(\mathcal{A})$ since continuity saves the day, but that is another story when discontinuities appear.*

Suppose that $\mathcal{T} = [0, 1]$ is endowed with its usual indexing collection from Example 1.2.14. Take the set-indexed map $h = \mathbb{1}_{\{[0,1]\}}$: this is a perfectly reasonable map that should belong to our càdlàg space. However, consider the geodesic \mathcal{A} -flows ϕ and ϕ_ε ($\varepsilon > 0$) defined by

$$\forall t \in [0, 1], \quad \phi(t) = [0, t] \quad \text{and} \quad \phi_\varepsilon(t) = [0, (1 - \varepsilon)t].$$

Even though ϕ_ε converges uniformly to ϕ as $\varepsilon \rightarrow 0^+$, there is no way that $h^{\phi_\varepsilon} = 0$ converges to $h^\phi = \mathbb{1}_{\{1\}}$ in $\mathcal{D}(0, 1)$ as $\varepsilon \rightarrow 0^+$.

A way to bypass this issue is to restrict our attention to maximal flows $\bigcup_{n \in \mathbf{N}} \Phi_{\max}(\mathcal{T}_n)$ so that close enough flows will be seeing exactly the same subset of \mathcal{T} , some \mathcal{T}_n namely.

EXAMPLE 3.3.3 (The uniform topology on the flows is not enough). The second and last issue we could think of is that uniform topology put some flows too close to each other for the following reason: for a given càdlàg function, even though two flows may see the same jump, one might see it much sooner than the other.

Suppose that $\mathcal{T} = [0, 1]^2$ is endowed with its usual product indexing collection from Example 1.2.16. Take once more a set-indexed map h that should reasonably be càdlàg: for all $A \in \mathcal{A}$, $h(A) = \mathbb{1}_{(1/2, 1) \in A}$. Consider the maximal geodesic \mathcal{A} -flows ϕ and ϕ_ε ($\varepsilon > 0$) defined by

$$\forall t \in [0, 1], \quad \phi(t) = [0, (t, 1)] \quad \text{and} \quad \phi_\varepsilon(t) = \begin{cases} [0, (1 - \varepsilon, (1 - \varepsilon)^{-1}t)] & \text{if } t \leq (1 - \varepsilon)^2, \\ [0, (t, t)] & \text{if } t > (1 - \varepsilon)^2. \end{cases}$$

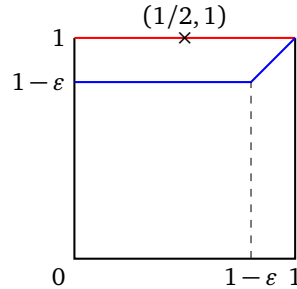


Figure 3.1: Tips of $\phi(t)$ (in red) and $\phi_\varepsilon(t)$ (in blue) for $t \in [0, 1]$.

As illustrated in Figure 3.1, $\phi(t)$ sees the jump of h as soon as $t \geq 1/2$ while $\phi_\varepsilon(t)$ only discovers it at $t = 1$. More precisely, we have $h^\phi = \mathbb{1}_{[1/2, 1]}$ and $h^{\phi_\varepsilon} = \mathbb{1}_{\{1\}}$ while ϕ_ε converges uniformly to ϕ as $\varepsilon \rightarrow 0^+$.

What Example 3.3.3 taught us is that, even for a reasonable càdlàg map h , the projection h^ϕ might not depend continuously on ϕ , at least with respect to the uniform topology. For that to happen, we need a topology such that once a flow ‘discovers’ a jump, all the flows close-by should expect to discover it quite soon, if not already. With that in mind, we introduce the following metric defined for all $\phi, \psi \in \Phi(\mathcal{A})$ by

$$d_\Phi(\phi, \psi) = \begin{cases} \sup_{s \in [0, 1]} \min_{t \in [0, 1]} \{ |s - t| : \phi(s) \subseteq \psi(t) \text{ and } \psi(s) \subseteq \phi(t) \} & \text{if } \phi(1) = \psi(1), \\ 1 & \text{if } \phi(1) \neq \psi(1). \end{cases} \quad (3.3.4)$$

PROPOSITION 3.3.4. For all $n \in \mathbf{N}$, the space $(\Phi_{\max}(\mathcal{T}_n), d_\Phi)$ is compact.

Proof. Let us show that d_Φ is well-defined, i.e. that the ‘min’ in (3.3.4) is reached. Let $\phi, \psi \in \Phi(\mathcal{A})$ such that $\phi(1) = \psi(1)$ and denote

$$\forall s \in [0, 1], \quad \rho(s) = \inf_{t \in [0, 1]} \{ |s - t| : \phi(s) \subseteq \psi(t) \text{ and } \psi(s) \subseteq \phi(t) \}.$$

Let $s \in [0, 1]$. By definition, we know that

$$\forall \varepsilon > 0, \quad \phi(s) \subseteq \psi(s + \rho(s) + \varepsilon)^1 \quad \text{and} \quad \psi(s) \subseteq \phi(s + \rho(s) + \varepsilon).$$

By outer continuity of ϕ and ψ (Definition 1.3.5), we may take $\varepsilon \rightarrow 0^+$ and get that $\phi(s) \subseteq \psi(s + \rho(s))$ and $\psi(s) \subseteq \phi(s + \rho(s))$. Hence $\rho(s)$ is a minimum attained for $t = s + \rho(s)$.

Let us check that d_Φ is a metric (a ‘true one’ for once). For $\phi, \psi \in \Phi(\mathcal{A})$, having $d_\Phi(\phi, \psi) = 0$ means that for all $t \in [0, 1]$, $\phi(t) \subseteq \psi(t)$ and $\psi(t) \subseteq \phi(t)$, hence $\phi = \psi$ and separability follows. Symmetry is obvious. As for triangle inequality, consider $\phi_1, \phi_2, \phi_3 \in \Phi(\mathcal{A})$. If $\phi_i(1) \neq \phi_j(1)$ for some $i, j \in \{1, 2, 3\}$, then at least $\phi_1(1) \neq \phi_2(1)$ or $\phi_2(1) \neq \phi_3(1)$ must hold. Hence

$$d_\Phi(\phi_1, \phi_2) + d_\Phi(\phi_2, \phi_3) \geq 1.$$

Since $d_\Phi(\phi_1, \phi_3) \leq 1$, the triangle inequality follows in this case.

If $\phi_1(1) = \phi_2(1) = \phi_3(1)$, we have for all $s \in [0, 1]$,

$$\forall s \in [0, 1], \quad \phi_1(s) \subseteq \phi_2(s + d_\Phi(\phi_1, \phi_2)) \subseteq \phi_3(s + d_\Phi(\phi_1, \phi_2) + d_\Phi(\phi_2, \phi_3)).$$

Hence $d_\Phi(\phi_1, \phi_2) \leq d_\Phi(\phi_1, \phi_2) + d_\Phi(\phi_2, \phi_3)$, so d_Φ is a metric on $\Phi(\mathcal{A})$.

Fix $n \in \mathbf{N}$. Writing $\mathcal{T}_n = \bigcup_{i=1}^j A_i$ where each A_i is maximal for \subseteq in \mathcal{T}_n , we get $\Phi_{\max}(\mathcal{T}_n) = \bigcup_{i=1}^j \Phi_{\max}([\emptyset, A_i])$. So in order to show that $\Phi_{\max}(\mathcal{T}_n)$ is compact, it is enough to show that for all $A \in \mathcal{A}_n$, $\Phi_{\max}([\emptyset, A])$ is compact. Let $A \in \mathcal{A}_n$.

Completeness is checked as follows: consider a Cauchy sequence $(\phi_k)_{k \in \mathbf{N}}$ in $\Phi_{\max}([\emptyset, A])$. By definition of d_Φ , we have

$$\forall \varepsilon > 0, \exists n_\varepsilon \in \mathbf{N} : \forall j, k \geq n_\varepsilon, \forall t \in [0, 1], \quad \phi_k(t - \varepsilon) \subseteq \phi_j(t) \subseteq \phi_k(t + \varepsilon)$$

which enables us to define the geodesic flow $\phi \in \Phi_{\max}([\emptyset, A])$ given by

$$\forall t \in [0, 1], \quad \phi(t) = \bigcap_{\varepsilon \in \mathbf{Q}_+^*} \bigcap_{j \geq n_\varepsilon} \phi_j(t + \varepsilon).$$

Then it may be shown that $d_\Phi(\phi_k, \phi) \rightarrow 0$ as $k \rightarrow \infty$: completeness follows.

Let us show that $\Phi_{\max}([\emptyset, A])$ is totally bounded. Let $k \geq n$. We wish to use the finite subcollection \mathcal{A}_k (Definition 1.2.1) to create an ε_k -cover of $\Phi_{\max}([\emptyset, A])$ (where $\varepsilon_k \rightarrow 0$ as $k \rightarrow \infty$). Using Proposition 1.3.9, we may fix for all chain $\mathcal{A}' \subseteq \mathcal{A}_k$ a flow $\psi_{\mathcal{A}'}$ such that for all $A \in \mathcal{A}'$, there is $t_A \in [0, 1]$ such that $A \subseteq \psi_{\mathcal{A}'}(t_A)$ and $\mathbf{m}(\psi_{\mathcal{A}'}(t_A) \setminus A) = 0$.

Let $\phi \in \Phi_{\max}([\emptyset, A])$ and consider the chain $\mathcal{A}' = \{g_k(\phi(t)) : t \in [0, 1]\} \subseteq \mathcal{A}_k$. We claim that ϕ and $\psi_{\mathcal{A}'}$ are close. Let $s \in [0, 1]$. Since

$$\max_{\substack{A' \in \mathcal{A}' \\ A' \subset g_k(\phi(s))}} A' \subseteq \phi(s) \subseteq g_k(\phi(s)),$$

¹We commit a slight abuse of notation here: consider that $\psi(t) = \psi(1)$ whenever $t > 1$. The same abuse will be committed throughout the proof.

we have

$$\min_{\substack{t \in [0,1]: \\ \phi(s) \subseteq \psi_{\mathcal{A}'}(t)}} \mathbf{m}(\psi_{\mathcal{A}'}(t) \setminus \phi(s)) \leq \delta_k \quad \text{and} \quad \min_{\substack{t \in [0,1]: \\ \psi_{\mathcal{A}'}(s) \subseteq \phi(t)}} \mathbf{m}(\phi(t) \setminus \psi_{\mathcal{A}'}(s)) \leq \delta_k \quad (3.3.5)$$

where

$$\delta_k = \max_{\substack{A_0, A_1 \in [\emptyset, A] \cap \mathcal{A}_k: \\ A_0 \text{ maximal proper subset of } A_1}} \mathbf{m}(A_1 \setminus A_0)$$

is a quantity that has been introduced back in Definition 1.4.1. Since both ϕ and $\psi_{\mathcal{A}'}$ are geodesic flows with speed $\mathbf{m}(A)$, we have

$$\forall t \in [0, 1], \quad \mathbf{m}(\phi(t)) = \mathbf{m}(\psi_{\mathcal{A}'}(t)) = \mathbf{m}(A)t. \quad (3.3.6)$$

Without loss of generality, we may suppose that $\mathbf{m}(A) > 0$. It follows from (3.3.5) and (3.3.6) that

$$d_{\Phi}(\phi, \psi_{\mathcal{A}'}) \leq \delta_k / \mathbf{m}(A).$$

Hence $\{\psi_{\mathcal{A}'} : \mathcal{A}' \text{ chain in } \mathcal{A}_k\}$ is a $\delta_k / \mathbf{m}(A)$ -cover of $\Phi_{\max}([\emptyset, A])$. According to the shrinking mesh property (Definition 1.4.1), we know that $\delta_k \rightarrow 0$ as $k \rightarrow \infty$. Thus $\Phi_{\max}([\emptyset, A])$ is complete and totally bounded, hence compact. The result follows. \square

PROPOSITION 3.3.5. *The topology induced by d_{Φ} is stronger than the topology of uniform convergence, i.e. if $d_{\Phi}(\phi_k, \phi) \rightarrow 0$ as $k \rightarrow \infty$, then ϕ_k converges uniformly to ϕ as $k \rightarrow \infty$.*

Proof. Consider geodesic flows $\phi, \phi_0, \phi_1, \dots$ such that $d_{\Phi}(\phi_k, \phi) \rightarrow 0$ as $k \rightarrow \infty$. Let us show that the convergence also holds uniformly. Let $\varepsilon > 0$. By definition of d_{Φ} , there exists $n_{\varepsilon} \in \mathbf{N}$ such that

$$\forall k \geq n_{\varepsilon}, \quad \phi_k(1) = \phi(1) \quad \text{and} \quad \forall t \in [0, 1], \quad \phi_k(t) \subseteq \phi(t + \varepsilon). \quad (3.3.7)$$

Thus, we get for all $k \geq n_{\varepsilon}$ and $t \in [0, 1]$,

$$\begin{aligned} d_m(\phi_k(t), \phi(t)) &\leq d_m(\phi_k(t), \phi(t + \varepsilon)) + d_m(\phi(t + \varepsilon), \phi(t)) \\ &= [\mathbf{m}(\phi(t + \varepsilon)) - \mathbf{m}(\phi_k(t))] + [\mathbf{m}(\phi(t + \varepsilon)) - \mathbf{m}(\phi(t))] \quad \text{by (3.3.7),} \\ &= [v(\phi)(t + \varepsilon) - v(\phi_k)t] + [v(\phi)(t + \varepsilon) - v(\phi)t] \\ &= 2v(\phi)\varepsilon \quad \text{since } v(\phi_k) = v(\phi). \end{aligned}$$

This bound being independent from t , the result follows. \square

3.3.3 The space $\mathcal{D}^{\Phi}(\mathcal{A})$ and its topology

For all $n \in \mathbf{N}$, denote

$$\mathcal{D}^{\Phi}(\mathcal{T}_n) = \left\{ h \in \mathbf{R}^{\mathcal{T}_n} : \phi \mapsto h^{\phi} \text{ is continuous from } (\Phi(\mathcal{T}_n), d_{\Phi}) \text{ to } (\mathcal{D}(0, 1), d_{\mathcal{D}(0,1)}) \right\}. \quad (3.3.8)$$

Those spaces are Polish for the topology of uniform convergence thanks to Proposition 3.3.4 and [9, Lemma 3.99]. The càdlàg space $\mathcal{D}^{\Phi}(\mathcal{A})$ may finally be defined as

$$\mathcal{D}^{\Phi}(\mathcal{A}) = \left\{ h \in \mathbf{R}^{\mathcal{A}} : \forall n \in \mathbf{N}, h|_{\mathcal{T}_n} \in \mathcal{D}^{\Phi}(\mathcal{T}_n) \right\} \quad (3.3.9)$$

and endowed with the projective limit topology from Theorem 3.2.1 applied to the sequence $(\mathcal{D}^\Phi(\mathcal{T}_n))_{n \in \mathbb{N}}$, the maps $\pi_m^n = (\cdot)|_{\mathcal{T}_m}$ and $\pi_n = (\cdot)|_{\mathcal{T}_n}$ ($m \leq n$). In particular, $\mathcal{D}^\Phi(\mathcal{A})$ is a Polish space and, as a consequence of Proposition 3.3.5, $\mathcal{C}^\Phi(\mathcal{A})$ continuously embeds itself into $\mathcal{D}^\Phi(\mathcal{A})$.

The comparison of $\mathcal{D}^\Phi(\mathcal{A})$ with Slonowsky's space $\mathcal{D}(\mathcal{A})$ [47, Definition 7.1.1] is not entirely straightforward, in the sense that our topology stems from d_m whereas his is based on the Hausdorff metric d_H given in (1.4.2). What we can say is that his definition appears to be less restrictive than ours since no continuity is imposed when one changes flow. So it appears that a function will generally have an easier time belonging to $\mathcal{D}(\mathcal{A})$ than $\mathcal{D}^\Phi(\mathcal{A})$. However, this heuristic does not always hold as Proposition 3.3.6 will tell.

This construction has two nice aspects for itself that other constructions do not:

- ◊ Bass and Pyke [17, §6.4] raised the concern that endowing a set-indexed càdlàg space with the J_1 -topology would be hard due to the difficulty to define and manipulate set-indexed time changes in the spirit of Definition 3.3.1. Here, we circumvent this issue by having one-dimensional time changes along paths.
- ◊ This construction may be readily adapted to any other Skorokhod topology that one might fancy. One just has to change the topology of $\mathcal{D}(0, 1)$ in (3.3.8) and the deal is done.

3.3.4 A Lévy-Itô decomposition

The following result may be compared with [47, Theorem 7.1.6] where set-indexed point processes are shown to belong to Slonowsky's space $\mathcal{D}(\mathcal{A})$, provided they verify an additional geometric assumption.

PROPOSITION 3.3.6. *Any set-indexed point process $X = \{\mu(A) : A \in \mathcal{A}\}$ has a version in $\mathcal{D}^\Phi(\mathcal{A})$.*

Proof. It is enough to prove the result for a deterministic measure μ and since $\mathcal{D}^\Phi(\mathcal{A})$ is a vector space, we might as well suppose that $\mu = \delta_t$ is the Dirac measure for some $t \in \mathcal{T}$. Let $n \in \mathbb{N}$ be big enough so that $t \in \mathcal{T}_n$ and consider $\phi \in \Phi_{\max}(\mathcal{T}_n)$. We need to show that the map

$$\begin{aligned} \Phi_{\max}(\mathcal{T}_n) &\longrightarrow \mathcal{D}(0, 1) \\ \psi &\longmapsto [s \mapsto \mathbb{1}_{t \in \psi(s)}] \end{aligned}$$

is continuous at ϕ . For all $\psi \in \Phi_{\max}(\mathcal{T}_n)$, define $t^\psi = \min \{s \in [0, 1] : t \in \psi(s)\}$. By (3.3.3), we get

$$\forall \psi \in \Phi_{\max}(\mathcal{T}_n), \quad d_{\mathcal{D}(0,1)}(\mathbb{1}_{t \in \phi(\cdot)}, \mathbb{1}_{t \in \psi(\cdot)}) = d_{\mathcal{D}(0,1)}(\mathbb{1}_{[t^\phi, 1]}, \mathbb{1}_{[t^\psi, 1]}) = |t^\phi - t^\psi|.$$

So we only need to prove that $\psi \mapsto t^\psi$ is continuous at ϕ . Let $\varepsilon \in (0, 1)$ and take $\psi \in \Phi_{\max}(\mathcal{T}_n)$ such that $d_\Phi(\phi, \psi) < \varepsilon$. By definition of d_Φ , we have

$$\phi(t^\phi) \subseteq \psi(t^\phi + \varepsilon) \quad \text{and} \quad \psi(t^\psi) \subseteq \phi(t^\psi + \varepsilon).$$

Thus

$$t^\psi \leq t^\phi + \varepsilon \quad \text{and} \quad t^\phi \leq t^\psi + \varepsilon.$$

Hence $|t^\phi - t^\psi| \leq \varepsilon$ and the continuity follows. \square

DEFINITION 3.3.7 (Primitive process). *Suppose that X extends to a stochastic measure and consider $f \in L_{loc}(X)$. The primitive process of f with respect to X is the set-indexed process $Y = \{Y_A : A \in \mathcal{A}\}$ given by*

$$\forall A \in \mathcal{A}, \quad Y_A = \mathcal{X}(f \mathbb{1}_A) = \int_A f dX.$$

When $\mathcal{T} = \mathbf{R}_+$, \mathcal{A} is its usual indexing collection (Example 1.2.7) and ΔX the Lebesgue measure, the primitive process corresponds to the usual primitive of f that vanishes at 0.

THEOREM 3.3.8 (Lévy-Itô decomposition of the primitive process). *Suppose that Y is the primitive process of $f \in L_{\text{loc}}(X)$ with respect to some siLévy X with triplet (b, σ^2, ν) and representation (2.4.7) from Corollary 2.4.9. If $\dim \mathcal{A} < \infty$, then Y is such that*

$$\forall A \in \mathcal{A}, \quad Y_A = b \int_A f d\mathbf{m} + \sigma \int_A f dW + \lim_{\varepsilon \rightarrow 0^+} \int_{A \times \{x \in \mathbf{R}^*: |x| \geq \varepsilon\}} [x.f(t)] \tilde{N}(dt, dx) \quad (3.3.10)$$

where \tilde{N} is the compensated Poisson process of intensity ν that appears in (2.4.3) and the convergence as $\varepsilon \rightarrow 0^+$ happens a.s. uniformly in $A \in \mathcal{T}_n$ for all $n \in \mathbf{N}$. In particular, if f is bounded on each \mathcal{T}_n , then Y has a version in $\mathcal{D}^\Phi(\mathcal{A})$.

This result seems strictly weaker than [41, Theorem 7.9] when $f = 1$ (so that $X = Y$) because we further require that $\dim \mathcal{A} < \infty$. The reason is because we believe to have a counter-example when $\dim \mathcal{A} = \infty$ (see Example 3.3.9), meaning that the assumptions of [41, Theorem 7.9] should be reinforced.

Sketch of proof. The representation (3.3.10) has already been obtained in (2.4.8) but for the a.s. uniform convergence on each \mathcal{T}_n .

As mentioned in the proof of [41, Theorem 7.9], the proof goes like [3, Theorem 4.6], but for the application of Wichura's inequality which is replaced by the maximal inequality (2.3.29), which is exactly why we required $\dim \mathcal{A} < \infty$.

The fact that the Gaussian (resp. Poissonian) part of (3.3.10) belongs to $\mathcal{C}^\Phi(\mathcal{A})$ (resp. $\mathcal{D}^\Phi(\mathcal{A})$) is a consequence of Proposition 3.2.2 (resp. Proposition 3.3.5). \square

EXAMPLE 3.3.9 (Why Theorem 3.3.8 may fail when $\dim \mathcal{A} = \infty$). *In Example 3.2.4, we proved that for a specific Hilbert's cube $\mathcal{T} = \prod_{j \in \mathbf{N}^*} [0, a_j]$ and measure \mathbf{m} , the siBm W has a.s. discontinuous sample paths at \emptyset . This siBm could be constructed as the sum of iid one-dimensional Brownian motions (3.2.7) due to the particular form (3.2.6) of \mathbf{m} .*

Let us do the same here for a compensated compound Poisson process, i.e. suppose that \mathcal{T} , \mathcal{A} and \mathbf{m} are the ones from Example 3.2.4, consider iid \mathbf{R}_+ -indexed compensated compound Poisson processes $(\tilde{P}^j)_{j \in \mathbf{N}}$ of intensity ν and define the process $\tilde{P} = \{\tilde{P}_A : A \in \mathcal{A}\}$ by

$$\forall A \in \mathcal{A}, \quad \tilde{P}_A = \sum_{j=1}^{\infty} \tilde{P}_{t(A)_j}^j.$$

The last ingredient used was a knowledge of the (pointwise) Hölder regularity of the Brownian motion at 0 to prove (3.2.13). No other important fact about the Brownian motion has been used. In particular, if one carefully chooses ν so that the pointwise Hölder exponent of the \tilde{P}^j 's are the same than the Brownian motion, we would still be able to prove (3.2.13) for \tilde{P} , i.e.

$$\limsup_{A \rightarrow \emptyset} \tilde{P}_A > \kappa \quad \text{a.s.} \quad (3.3.11)$$

for some constant $\kappa > 0$. This is made possible by taking $\nu(dx) = |x|^{-3/2} \mathbb{1}_{x \in [-1, 1] \setminus \{0\}} \mathbf{Leb}(dx)$ and applying [20, Theorem 3.3].

Now, we may see that (3.3.11) contradicts the uniform convergence in (3.3.10) since $\tilde{P}_\emptyset = 0$ a.s. Hence Theorem 3.3.8 may not hold for some infinite-dimensional indexing collections.

A big advantage of Theorem 3.3.8 is that it enables us to talk about the ‘jump structure’ of the primitive process. The notion of point-mass jump given below comes from [41, Definition 7.1] where it has been used to give the Lévy-Itô representation of a siLévy.

DEFINITION 3.3.10 (Point-mass jumps, [41, Definition 7.1]). *The point-mass jump of a function $h : \mathcal{A} \rightarrow \mathbf{R}$ at $t \in \mathcal{T}$ is the quantity $J_t(h) = \lim_{n \rightarrow \infty} \Delta h(C_n(t))$ whenever the limit exists. The function h has well-defined jumps if $J_t(h)$ is well-defined for all $t \in \mathcal{T}$ and the set $\{t \in A : |J_t(h)| \geq \varepsilon\}$ is finite for all $A \in \mathcal{A}$ and $\varepsilon > 0$. In that case, the jump set of h is given by*

$$\Pi(h) = \{t \in \mathcal{T} : J_t(h) \neq 0\}. \quad (3.3.12)$$

PROPOSITION 3.3.11 (Representation of a set-indexed compensated Poisson process). *Let \tilde{N} be a set-indexed compensated Poisson process of intensity ν as considered in (2.4.2). Then, there exists a Poisson random set Π on $\mathcal{T} \times \mathbf{R}^*$ such that*

$$\forall A' \in \mathcal{A} \times \mathcal{A}_{\mathbf{R}^* \cup \{\infty\}}, \quad \tilde{N}_{A'} = \left[\sum_{(s,J) \in \Pi} \delta_{(s,J)}(A') \right] - \mathbf{m} \otimes \nu(\{(s,x) \in A' : |x| \leq 1\}).$$

Proof. For a more details about Poisson random sets, we refer to Remark 2.4.4 and the references mentioned there.

By definition of \tilde{N} , we know that such a representation exists, at least in distribution. Remark that this representation may live in the Borel space $\mathbf{R}^{\mathbf{N}}$ since only countable many random variables are required to define it. Thus, this representation in distribution may be turned into a version of \tilde{N} by virtue of [54, Theorem 6.10]. \square

COROLLARY 3.3.12 (Jump structure of the primitive process). *If $\dim \mathcal{A} < \infty$ and Y is the primitive process of $f \in L_{loc}(X)$ with respect to some siLévy X , then the following holds with probability one: both X and Y have well-defined jumps, $\Pi(Y) = \Pi(X) \cap \{t \in \mathcal{T} : f(t) \neq 0\}$ and for all $t \in \mathcal{T}$, $J_t(Y) = f(t)J_t(X)$.*

Proof. Due to (3.3.10), the primitive process Y has three parts: a drift, a Gaussian part and a Poissonian part. We show that each part has well-defined jumps separately.

Since \mathbf{m} is compatible, we have $\mathbf{m}(\{t\}) = 0$ for all $t \in \mathcal{T}$. Hence $\mathbf{m}(C_n(t)) \rightarrow 0$ as $n \rightarrow \infty$ and, by the dominated convergence theorem,

$$\forall t \in \mathcal{T}, \quad b \int_{C_n(t)} f d\mathbf{m} \longrightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (3.3.13)$$

For the Gaussian part, one may copy the proof of [41, Theorem 7.3]² and prove that the following holds with probability one:

$$\forall t \in \mathcal{T}, \quad \sigma \int_{C_n(t)} f dW \longrightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (3.3.14)$$

As for the Poissonian part — denoted by \tilde{P} — let us use Proposition 3.3.11 on the process \tilde{N} from (3.3.10). This enables us to write

$$\forall A \in \mathcal{A}, \quad \tilde{P}_A = \lim_{\varepsilon \rightarrow 0^+} \left[\sum_{\substack{(s,J) \in \Pi: \\ s \in A, |J| \geq \varepsilon}} f(s)J - \int_A f d\mathbf{m} \int_{\varepsilon \leq |x| \leq 1} x \nu(dx) \right]$$

²Indeed, only the fact that the process is Gaussian and has independent increments has been used in the proof, and not that it is a siBm per say.

where the convergence as $\varepsilon \rightarrow 0^+$ happens a.s. uniformly in $A \in \mathcal{T}_k$ for all $k \in \mathbf{N}$. Let Ω^* be the event where this uniform convergence happens. In particular, we get for all $\omega \in \Omega^*$ and $t \in \mathcal{T}$,

$$\forall n \in \mathbf{N}, \quad \Delta \tilde{P}_{C_n(t)}(\omega) = \lim_{\varepsilon \rightarrow 0^+} \left[\sum_{\substack{(s,J) \in \Pi: \\ s \in C_n(t), |J| \geq \varepsilon}} f(s)J - \int_{C_n(t)} f d\mathbf{m} \int_{\varepsilon \leq |x| \leq 1} x \nu(dx) \right](\omega). \quad (3.3.15)$$

Since $\dim \mathcal{A} < \infty$, we know that the left-neighborhoods \mathcal{C}^ℓ — which the $C_n(t)$'s are part of — are made of at most $\dim \mathcal{A}$ elements of \mathcal{A} . In particular, the convergence as $\varepsilon \rightarrow 0^+$ in (3.3.15) happens uniformly in $t \in \mathcal{T}_k$ and $n \in \mathbf{N}$ for all $k \in \mathbf{N}$. So we may invert limits to get

$$\forall \omega \in \Omega^*, \forall t \in \mathcal{T}, \quad \Delta \tilde{P}_{C_n(t)}(\omega) \longrightarrow f(t)J \mathbb{1}_{(t,J) \in \Pi(\omega)} \quad \text{as } n \rightarrow \infty. \quad (3.3.16)$$

Specializing (3.3.16) to X instead of Y ($f = 1$) proves that

$$\forall \omega \in \Omega^*, \forall t \in \mathcal{T}, \quad \Delta X_{C_n(t)} \longrightarrow J \mathbb{1}_{(t,J) \in \Pi(\omega)} = J_t(X) \quad \text{as } n \rightarrow \infty. \quad (3.3.17)$$

The result follows from (3.3.13), (3.3.14), (3.3.16) and (3.3.17). \square

3.4 Hölder exponents for generalized processes

Hölder regularity is expressed in terms of exponents and may vary depending on the context and the behavior one wishes to capture. In Section 3.4.1, we provide the necessary definitions to this effect.

In Section 3.4.2, we further push ideas from Jaffard [50, 51] to obtain (deterministic) upper bounds for the Hölder regularity of a function $h : \mathcal{A} \rightarrow \mathbf{R}$ based on its pointwise jumps. The notion of *vicinity* developed in Section 1.4.3 is the key concept to improve on the ‘naive’ upper bound.

This is all preparatory work for Section 3.5 where we give bounds on the Hölder regularity on the primitive process Y of some deterministic function f with respect to a siLévy X .

In the sequel, we fix a function $h : \mathcal{A} \rightarrow \mathbf{R}$. Contrary to the previous section, we go back to a more general metric $d_{\mathcal{A}}$ than d_m .

3.4.1 Set-indexed pointwise Hölder exponents

In [42], Herbin and Richard defined a number of Hölder exponents which localize different visions of the continuity property in the set-indexed case. In a general fashion, those Hölder exponents are used to finely study the regularity of h around a given $A \in \mathcal{A}$ or equivalently through the TIP bijection, around a given $t \in \mathcal{T}$. We will be using the convention $\sup \emptyset = 0$ which is usual for regularity exponents.

Hölder exponent

First, the (*pointwise*) Hölder exponent constitutes the natural generalization of its one-dimensional analog to the metric space $(\mathcal{A}, d_{\mathcal{A}})$:

$$\forall A \in \mathcal{A}, \quad \alpha_h(A) = \sup \left\{ \alpha \geq 0 : \limsup_{\rho \rightarrow 0^+} \sup_{A' \in B_{\mathcal{A}}(A, \rho)} \frac{|h(A) - h(A')|}{\rho^\alpha} < \infty \right\}. \quad (3.4.1)$$

If positive, for any $\alpha \in (0, \alpha_h(A))$, the pointwise Hölder exponent yields the following control of h in the neighborhood of t :

$$\exists \rho_\alpha > 0 : \forall A' \in B_{\mathcal{A}}(A, \rho_\alpha), \quad |h(A) - h(A')| \leq d_{\mathcal{A}}(A, A')^\alpha. \quad (3.4.2)$$

Conversely, the estimate (3.4.2) implies $\alpha \leq \alpha_h(A)$.

In modern literature, one usually uses a slight modification of the above definition where one substracts the smooth part of the function — its Taylor expansion — before comparing it to a power of the radius (see [15] for an in-depth comparison in the case $\mathcal{T} = \mathbf{R}_+$). However, the set-indexed setting does not seem to have any natural substitutes for polynomials, hence the definition. Moreover, keeping the polynomial part has even proven to be useful [15] to study stochastic processes when $\mathcal{T} = \mathbf{R}_+$.

Herbin and Richard also introduced the (*pointwise*) Hölder \mathcal{C} -exponent in order to look at the variation of h in terms of the \mathcal{C} -indexed increment map Δh . The authors proved [42, Proposition 3.2] that the following definition does not depend on the choice of $k \in \mathbf{N}^*$:

$$\forall A \in \mathcal{A}, \quad \alpha_{h, \mathcal{C}}(A) = \sup \left\{ \alpha \geq 0 : \limsup_{\rho \rightarrow 0^+} \sup_{C \in \mathcal{C}_{(k)} \cap B_{\mathcal{C}}(A, \rho)} \frac{|\Delta h(C)|}{\rho^\alpha} < \infty \right\} \quad (3.4.3)$$

where the class $\mathcal{C}_{(k)}$ has been given in Definition 1.2.35. Such a definition leads to the corresponding estimate — apparently stronger than (3.4.2) for now — for $\alpha \in (0, \alpha_{h, \mathcal{C}}(A))$:

$$\forall k \in \mathbf{N}^*, \exists \rho_{\alpha, k} > 0 : \forall C \in \mathcal{C}_{(k)} \cap B_{\mathcal{C}}(A, \rho_{\alpha, k}), \quad |\Delta h(C)| \leq d_{\mathcal{C}}(A, C)^\alpha. \quad (3.4.4)$$

A reason why it is preferred over a more natural definition on \mathcal{C} is that \mathcal{C} is usually not a Vapnik-Červonenkis class (since \mathcal{C}^ℓ is a dissecting system), so \mathcal{C} -indexed processes are far from having continuous sample paths in general (see [4, Sections 1.4.3 and 1.4.5] for more details).

Finally, it has not been seen in [42] that the usual Hölder exponent and the \mathcal{C} -exponent actually coincide.

PROPOSITION 3.4.1. *For all $A \in \mathcal{A}$, $\alpha_h(A) = \alpha_{h, \mathcal{C}}(A)$.*

In particular, we will only mention $\alpha_h(A)$ in the following and still use both estimates (3.4.2) and (3.4.4).

Proof. Let $A \in \mathcal{A}$. Then

$$\begin{aligned} \sup_{C \in \mathcal{C}_0 \cap B_{\mathcal{C}}(A, \rho)} |\Delta h(C)| \rho^{-\alpha} &= \sup_{\substack{A_0, A_1 \in B_{\mathcal{A}}(A, \rho): \\ A_0 \subseteq A_1}} |\Delta h(A_1 \setminus A_0)| \rho^{-\alpha} \\ &= \sup_{\substack{A_0, A_1 \in B_{\mathcal{A}}(A, \rho): \\ A_0 \subseteq A_1}} |h(A_1) - h(A_0)| \rho^{-\alpha} \\ &\leq \sup_{\substack{A_0, A_1 \in B_{\mathcal{A}}(A, \rho): \\ A_0 \subseteq A_1}} [|h(A_1) - h(A)| + |h(A) - h(A_0)|] \rho^{-\alpha} \\ &\leq 2 \sup_{A' \in B_{\mathcal{A}}(A, \rho)} |h(A) - h(A')| \rho^{-\alpha}. \end{aligned}$$

Hence taking $k = 1$ in (3.4.3) immediately yields that $\alpha_h(A) \leq \alpha_{h, \mathcal{C}}(A)$.

Conversely, if $\alpha_{h, \mathcal{C}}(A) = 0$, then equality immediately holds. Otherwise, take $\alpha \in (0, \alpha_{h, \mathcal{C}}(A))$. Consider $\rho_{\alpha, 0} > 0$ just as in the estimate (3.4.4) for $k = 1$ and let $A' \in B_{\mathcal{A}}(A, \rho_{\alpha, 0})$. Then,

$$|h(A) - h(A')| \leq |\Delta h(A \setminus A')| + |\Delta h(A' \setminus A)|. \quad (3.4.5)$$

Since the extremal representation of $A \setminus A'$ is $A \setminus (A \cap A')$, it follows by definition of $d_{\mathcal{C}}$ that

$$d_{\mathcal{C}}(A \setminus A', A) = \max \{d_{\mathcal{C}}(A, A'), d_{\mathcal{C}}(A \cap A', A)\}.$$

By contractivity, it follows that $d_{\mathcal{C}}(A \cap A', A) \leq d_{\mathcal{C}}(A, A')$, so $d_{\mathcal{C}}(A \setminus A', A) = d_{\mathcal{C}}(A, A')$. Similarly, $d_{\mathcal{C}}(A' \setminus A, A) = d_{\mathcal{C}}(A, A')$. In particular, we may apply the estimate (3.4.4) to (3.4.5) and get

$$\forall A' \in B_{\mathcal{C}}(A, \rho_{\alpha, 0}), \quad |h(A) - h(A')| \leq 2d_{\mathcal{C}}(A, A')^{\alpha}.$$

Hence $\alpha \leq \alpha_h(A)$, the result follows. \square

REMARK 3.4.2 (Why we do not consider a local exponent). Taking $\alpha \in (0, \tilde{\alpha}_h(A))$ where $\tilde{\alpha}_h(A)$ is the local analog of $\alpha_h(A)$ (see [42] and references therein for a precise definition and context) would yield an estimate similar to (3.4.2): for all $\alpha \in (0, \tilde{\alpha}_h(A))$,

$$\forall A_0, A_1 \in B_{\mathcal{C}}(A, \rho_{\alpha}), \quad |h(A_0) - h(A_1)| \leq d_{\mathcal{C}}(A_0, A_1)^{\alpha}$$

from which one could deduce an estimate similar to (3.4.4) of the form

$$\forall C, C' \in \mathcal{C}_{(k)} \cap B_{\mathcal{C}}(A, \rho_{\alpha, k}), \quad |\Delta h(C) - \Delta h(C')| \leq d_{\mathcal{C}}(C, C')^{\alpha}. \quad (3.4.6)$$

The reason why we do not consider such exponent here is because we study the primitive process Y of some function f with respect to a siLévy X .

If Y were to have a non-trivial local regularity, since it has well-defined jumps by Corollary 3.3.12, it follows from Lemma 1.4.5 and (3.4.6) that $t \mapsto J_t(Y)$ must be Hölder-continuous in a neighborhood of A (i.e. for points $t \in \mathcal{T}$ such that $d_{\mathcal{C}}(A(t), A)$ is small).

This happens only if the Lévy measure ν is null, in which case Y is Gaussian and [42] shows that it generally implies that $\alpha_Y = \tilde{\alpha}_Y$. So local exponents do not constitute the right tools here.

$d_{\mathcal{T}}$ -localized exponent

In Section 1.4.3, we talked about the problem that characterizing the regularity of h through increments of the form $h(A) - h(A')$ — which is the case for $\alpha_h(A)$ — requires to take non-local information into account. We introduced in Definition 1.4.6 the notion of vicinity especially to tackle this issue. Another way to solve the problem is to swap the increment $h(A) - h(A')$ by $\Delta h(C)$ for $C \in \mathcal{C}$ close to A and of small diameter. Not only the shrinking mesh property (Definition 1.4.1) will ensure that this definition is well-posed, but also $\Delta h(C)$ actually constitutes the ‘right’ notion of increment in the set-indexed setting. This fact has already been noted in two-parameter literature.

In this section, we momentarily denote $p = \dim \mathcal{A}$ for notation’s sake.

With that in mind, we define the $d_{\mathcal{T}}$ -localized exponent of $h : \mathcal{A} \rightarrow \mathbf{R}$:

$$\forall A \in \mathcal{A}, \quad \alpha_{h, d_{\mathcal{T}}}(A) = \sup \left\{ \alpha \geq 0 : \limsup_{\rho \rightarrow 0^+} \sup_{\substack{C \in \mathcal{C}_{(p)} \cap B_{\mathcal{C}}(A, \rho) \\ \text{diam}(C) < \rho}} \frac{|\Delta h(C)|}{\rho^{\alpha}} < \infty \right\} \quad (3.4.7)$$

with the convention $\mathcal{C}_{(\infty)} = \mathcal{C}$ and where $\text{diam}(C) = \sup \{d_{\mathcal{T}}(s, s') : s, s' \in C\}$. Remark that, contrary to (3.4.3) which defines the \mathcal{C} -exponent, this definition does depend on p due to the condition on the diameter. Indeed, elements of $\mathcal{C}_{(k+1)}$ with diameter smaller than ρ cannot in general be only expressed using elements of $\mathcal{C}_{(k)}$ with diameter smaller than ρ .

The following proposition gives some properties about the $d_{\mathcal{T}}$ -localized exponent in order to get a better feel for it.

PROPOSITION 3.4.3. *Let $A = A(t) \in \mathcal{A}$. The following properties hold:*

1. (Equivalent definition using the TIP bijection).

$$\alpha_{h,d_{\mathcal{T}}}(A(t)) = \sup \left\{ \alpha \geq 0 : \limsup_{\rho \rightarrow 0^+} \sup_{\substack{C \in \mathcal{C}_{(p)} \cap B_{\mathcal{C}}(A(t), \rho) \\ C \subseteq B_{\mathcal{T}}(t, \rho)}} \frac{|\Delta h(C)|}{\rho^\alpha} < \infty \right\}. \quad (3.4.8)$$

2. (Corresponding estimate). *For all $\alpha \in (0, \alpha_{h,d_{\mathcal{T}}}(A))$, there exists $\rho_\alpha > 0$ such that for all $C \in \mathcal{C}_{(p)} \cap B_{\mathcal{C}}(A, \rho_\alpha)$,*

$$C \subseteq B_{\mathcal{T}}(t, \rho_\alpha) \implies |\Delta h(C)| \leq \max \{d_{\mathcal{C}}(A, C), \text{diam}(C)\}^\alpha. \quad (3.4.9)$$

Conversely, (3.4.9) implies $\alpha \leq \alpha_{h,d_{\mathcal{T}}}(A)$.

3. (Comparison to the Hölder exponent). *If $\dim \mathcal{A} < \infty$, then $\alpha_h(A) \leq \alpha_{h,d_{\mathcal{T}}}(A)$.*

Proof. Let us fix $A = A(t) \in \mathcal{A}$.

1. It is just a consequence of the fact that for all $\rho > 0$ and $C \in B_{\mathcal{C}}(A, \rho)$,

$$\begin{aligned} \text{diam}(C) < \rho &\implies C \subseteq B_{\mathcal{T}}(t, 2\rho) \\ C \subseteq B_{\mathcal{T}}(t, \rho) &\implies \text{diam}(C) \leq 2\rho \end{aligned}$$

2. It is straightforward consequence of (3.4.8).
3. Since $\mathcal{C}^\ell \subseteq \mathcal{C}_{(p)}$, it is but a simple comparison between (3.4.3) for $k = p$, Proposition 3.4.1 and (3.4.7). □

This exponent should also be compared to the *pointwise continuity exponent* introduced in [42, Definition 3.4]:

$$\forall t \in \mathcal{T}, \quad \alpha_h^{pc}(t) = \sup \left\{ \alpha \geq 0 : \limsup_{n \rightarrow \infty} \frac{|\Delta h(C_n(t))|}{m(C_n(t))^\alpha} < \infty \right\}. \quad (3.4.10)$$

However, we found our definition easier to work with since it is more closely linked to a metric, does not directly rely on the countable class \mathcal{C}^ℓ and yields a more powerful estimate at the end while still answering our need to replace $h(A) - h(A')$ with a better, more local, notion of increment.

3.4.2 Regularity of generic OJI functions

In [50], Jaffard used the discontinuities of a càdlàg function $h : \mathbf{R}_+ \rightarrow \mathbf{R}$ to obtain the following upper bound on its Hölder exponent:

$$\forall t \in \mathbf{R}_+ \setminus \Pi(h), \quad \alpha_h(t) \leq \liminf_{s \in \Pi(h) \rightarrow t} \frac{\log |J_s(h)|}{\log |s - t|} \quad (3.4.11)$$

where $J_s(h) = h(s) - h(s^-)$ for all $s \in \mathbf{R}_+$. In this section, we strive to generalize this approach in order to use it in a fashion similar to [51]. We start by treating the Hölder exponent only. We explain how to do a similar (and simpler) study of the $d_{\mathcal{T}}$ -localized exponent in Section 3.5.3.

Jump sets in generic configuration

In order to adapt (3.4.11) to a more general setting, we could consider the point-mass jumps $J_s(h) = \lim_{n \rightarrow \infty} \Delta h(C_n(t))$ of a map $h : \mathcal{A} \rightarrow \mathbf{R}$ with well-defined jumps (Definition 3.3.10) in a neighborhood of t and reproduce Jaffard's proof. This would yield the following bound:

$$\forall t \in \mathcal{T}, \quad \alpha_h(t) \leq \liminf_{\substack{s \in \Pi(h): \\ d_{\mathcal{T}}(s,t) \rightarrow 0^+}} \frac{\log |J_s(h)|}{\log d_{\mathcal{T}}(s,t)} \vee 0. \quad (3.4.12)$$

Recall that $\alpha_h(A)$ and $\alpha_h(t)$ are the same — provided that $A = A(t)$ — due to the correspondence (1.4.1) between $d_{\mathcal{A}}$ and $d_{\mathcal{T}}$. Moreover, the ‘ $\vee 0$ ’ ensures that this inequality holds for all $t \in \mathcal{T}$.

Alas, unless \mathcal{T} is one-dimensional, this upper bound turns out not to be so sharp. The reason is that we failed to consider the majority of the point-mass jumps contributing to lessen the regularity, *i.e.* the ones in the vicinity $\mathcal{V}(A, \rho)$ of $A = A(t)$ (Definition 1.4.6). Continuing our illustration with $(\mathcal{T}, d_{\mathcal{T}}) = (\mathbf{R}_+^2, d_2)$ from Example 1.4.7, the area of $B_{\mathcal{T}}(t, \rho)$ in Figure 1.11 is of order ρ^2 whereas the area of $\mathcal{V}(A, \rho)$ is of order ρ as $\rho \rightarrow 0$, so not taking the jumps in $\mathcal{V}(A, \rho) \setminus B_{\mathcal{T}}(t, \rho)$ into account incurs severe losses in the sharpness of the argument. It means that, in order to obtain a better upper bound on $\alpha_h(A)$, we need to be able to ‘fetch’ the jumps of h in the vicinity of A while only using elements in $B_{\mathcal{A}}(A, \rho)$ for small ρ . We do so for a given class of functions that fit our purposes.

DEFINITION 3.4.4 (OJI function). *A map $h : \mathcal{A} \rightarrow \mathbf{R}$ with well-defined jumps is said to be only jump-irregular (OJI) if it can be written in the form:*

$$\forall A \in \mathcal{A}, \quad h(A) = \lim_{\varepsilon \rightarrow 0^+} \left[\sum_{\substack{s \in \Pi \cap A: \\ |J_s| \geq \varepsilon}} J_s - a(A, \varepsilon) \right] \quad (3.4.13)$$

where Π is a set such that $\Pi \cap A$ is finite for all $A \in \mathcal{A}$, $(J_s)_{s \in \Pi}$ is a family of real numbers, $a : \mathcal{A} \times (0, 1) \rightarrow \mathbf{R}$ is a map such that for all $\varepsilon \in (0, 1)$, $a(\cdot, \varepsilon) \in \mathcal{C}^{\Phi}(\mathcal{A})$ and the limit as $\varepsilon \rightarrow 0^+$ happens uniformly in $A \in \mathcal{T}_n$ for all $n \in \mathbf{N}$.

The term ‘only jump-irregular’ is meant to indicate that OJI functions are only allowed to have discontinuities in the form of point-mass jumps. Although OJI functions belong to $\mathcal{D}^{\Phi}(\mathcal{A})$, we do not know whether the converse holds or not. Such a result would be interesting to prove in order to link it back to a similar study in the multiparameter setting [3].

REMARK 3.4.5. *If $\dim \mathcal{A} < \infty$ and h is an OJI function with representation (3.4.13), then one has $\Pi(h) = \Pi$ and for all $s \in \Pi(h)$, $J_s(h) = J_s$.*

In order to improve on (3.4.12) and take all the jumps of h in the vicinity of $A \in \mathcal{A}$ into account, it may happen that some jumps cannot be picked separately from the viewpoint of a given A . It is the case for instance whenever $J_s(h)$ and $J_{s'}(h)$ are both non-zero in Figure 3.2.

DEFINITION 3.4.6 (Generic configuration). *For a given $A \in \mathcal{A}$, an OJI function $h : \mathcal{A} \rightarrow \mathbf{R}$ with representation (3.4.13) is said to be in generic configuration in the vicinity of A if there exists $\rho_0 > 0$ such that for all $\rho \in (0, \rho_0)$, $s \in \Pi \cap \mathcal{V}(A, \rho)$ and $\varepsilon > 0$,*

$$\exists \underline{A}, \bar{A} \in B_{\mathcal{A}}(A, \rho) : \quad \underline{A} \subseteq \bar{A} \quad \text{and} \quad |\Delta h(\bar{A} \setminus \underline{A}) - J_s| \leq \varepsilon.$$

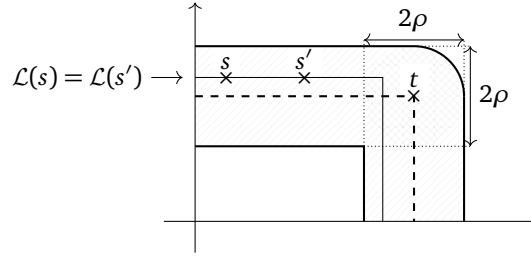


Figure 3.2: $s, s' \in \mathcal{V}(A(t), \rho)$ cannot be isolated from one another using elements in $B_{\mathcal{A}}(A(t), \rho)$.

Remark that there are OJI functions that are not in generic configuration in the vicinity of A . For instance, take the function defined for all $A \in \mathcal{A}$ by $h(A) = \mathbb{1}_{s \in A} + \mathbb{1}_{s' \in A}$ for s and s' as in Figure 3.2.

THEOREM 3.4.7 (Generic configuration for a siLévy's primitive process). *Let $A \in \mathcal{A}$. If $\dim \mathcal{A} < \infty$ and Y is the primitive process of $f \in L_{loc}(X)$ with respect to some siLévy X , then both X and Y are a.s. in generic configuration in the vicinity of A .*

Before proving this theorem, we need to introduce some notations in order to correctly define the following:

- ◊ The 'boundary' — later called $\mathcal{L}(s)$ — on which s (and s') stands in Figure 3.2.
- ◊ Approximating sequences $(\underline{A}_n(s))_{n \in \mathbb{N}}$ and $(\overline{A}_n(s))_{n \in \mathbb{N}}$ such that $\Delta h(\overline{A}_n(s) \setminus \underline{A}_n(s)) \rightarrow J_s(h)$ as $n \rightarrow \infty$. The sets $\overline{A}_n(s) \setminus \underline{A}_n(s) \in \mathcal{C}_0$ should be thought of as 'thick' versions of $\mathcal{L}(s)$ decreasing to $\mathcal{L}(s)$ as $n \rightarrow \infty$.

Fix $A \in \mathcal{A}$, $\rho > 0$ and $s \in \mathcal{V}(A, \rho)$. By Proposition 1.4.8, we may write $\mathcal{V}(A, \rho) = \bigcup_{n \in \mathbb{N}} \mathcal{V}_n(A, \rho)$ where

$$\forall n \in \mathbb{N}, \quad \mathcal{V}_n(A, \rho) = \bigcup_{\substack{\overline{A}, \underline{A} \in \mathcal{A}_n \cap B_{\mathcal{A}}(A, \rho): \\ \underline{A} \text{ maximal proper subset of } \overline{A}}} (\overline{A} \setminus \underline{A}) \quad (3.4.14)$$

and the 'maximal proper subset' condition is not a restriction since it just eliminates redundancy in the definition of $\mathcal{V}_n(A, \rho)$.

From (3.4.14), there exists a non-decreasing sequence $(\overline{A}_n(s) \setminus \underline{A}_n(s))_{n \geq n_0}$ such that for all $n \geq n_0$, $\underline{A}_n(s), \overline{A}_n(s) \in \mathcal{A}_n \cap B_{\mathcal{A}}(A, \rho)$, $\underline{A}_n(s)$ is a maximal proper subset of $\overline{A}_n(s)$ and $s \in \overline{A}_n(s) \setminus \underline{A}_n(s)$. Denote, as intended,

$$\mathcal{L}(s) = \mathcal{L}(A, \rho, s) = \bigcap_{n \in \mathbb{N}} (\overline{A}_n(s) \setminus \underline{A}_n(s)). \quad (3.4.15)$$

By the shrinking mesh property (Definition 1.3.1), we necessarily have $\mathbf{m}(\mathcal{L}(s)) = 0$.

LEMMA 3.4.8 (Sufficient condition for generic configuration). *Let $A \in \mathcal{A}$ and $h : \mathcal{A} \rightarrow \mathbf{R}$ an OJI function with representation (3.4.13). A sufficient condition for h to be in generic configuration in the vicinity of $A \in \mathcal{A}$ is that $\dim \mathcal{A} < \infty$ and there exists $\rho_0 > 0$ such that for all $\rho \in (0, \rho_0) \cap \mathbf{Q}$ and $s \in \Pi \cap \mathcal{V}(A, \rho)$,*

$$\Pi \cap \mathcal{L}(A, \rho, s) = \{s\}.$$

Proof. Let us fix $\rho \in (0, \rho_0) \cap \mathbf{Q}$ and $s \in \Pi$. Since h is OJI, we may write

$$\forall n \geq n_0, \quad \Delta h(\bar{A}_n(s) \setminus \underline{A}_n(s)) = \lim_{\varepsilon \rightarrow 0^+} \left[\sum_{\substack{t \in \Pi \cap (\bar{A}_n(s) \setminus \underline{A}_n(s)): \\ |J_t| \geq \varepsilon}} J_t - \Delta a(\bar{A}_n(s) \setminus \underline{A}_n(s), \varepsilon) \right] \quad (3.4.16)$$

where the convergence as $\varepsilon \rightarrow 0^+$ happens uniformly in $n \geq n_0$.

Since $m(A_{n_0}(s)) < \infty$ and $\bar{A}_n(s)$ decreases to $\bigcap_{k \in \mathbf{N}} \bar{A}_k(s)$ as $n \rightarrow \infty$, we have

$$d_m(\bar{A}_n(s), \bigcap_{k \in \mathbf{N}} \bar{A}_k(s)) \longrightarrow 0 \quad \text{as} \quad n \rightarrow \infty.$$

Moreover, the shrinking mesh property tells that $d_m(\underline{A}_n(s), \bar{A}_n(s)) \rightarrow 0$ as $n \rightarrow \infty$. Thus

$$d_m(\underline{A}_n(s), \bigcap_{k \in \mathbf{N}} \bar{A}_k(s)) \longrightarrow 0 \quad \text{as} \quad n \rightarrow \infty.$$

Hence, due to the continuity of $a(\cdot, \varepsilon)$ with respect to d_m , taking $n \rightarrow \infty$ in (3.4.16) and using $\Pi \cap \mathcal{L}(s) = \{s\}$ yields

$$\lim_{n \rightarrow \infty} \Delta h(\bar{A}_n(s) \setminus \underline{A}_n(s)) = \lim_{\varepsilon \rightarrow 0^+} \sum_{\substack{t \in \Pi \cap \mathcal{L}(s): \\ |J_t| \geq \varepsilon}} J_t = J_s.$$

Hence h is in generic configuration in the vicinity of A . \square

Proof of Theorem 3.4.7. According to Corollary 3.3.12, it is enough to prove the result for the silévy X . Using the Lévy-Itô decomposition (Theorem 3.3.8), we know that X has an OJI version.

Fix $A \in \mathcal{A}$ and $\rho > 0$. Define the event

$$\Omega_\rho = \left\{ \exists s, s' \in \Pi(X) \cap \mathcal{V}(A, \rho) : s \neq s' \text{ and } s' \in \mathcal{L}(s, A, \rho) \right\}.$$

Since $\Pi(X)$ is a Poisson random set, we may write $\Pi(X) = \{s_i : i < n\}$ for some random variable $n : \Omega \rightarrow \mathbf{N} \cup \{\infty\}$ and where a.s. for all $i, j < n$, $s_i \neq s_j$. In particular, we get

$$\mathbf{P}(\Omega_\rho | n) \leq \sum_{\substack{i, j < n: \\ i \neq j}} \mathbf{P}(s_j \in \mathcal{L}(s_i, A, \rho) | n). \quad (3.4.17)$$

Since $m(\mathcal{L}(s, A, \rho)) = 0$ for all $s \in \mathcal{T}$, each term in the sum of (3.4.17) is a.s. equal to 0. Thus for all $\rho > 0$, $\mathbf{P}(\Omega_\rho) = 0$. Hence $\Omega^* = \bigcap_{\rho \in \mathbf{Q}_+^*} \Omega_\rho^c$ is an event of probability one. Lemma 3.4.8 then implies that for all $\omega \in \Omega^*$, $X(\omega)$ is in generic configuration in the vicinity of $A \in \mathcal{A}$. \square

Bounding regularity with point-mass jumps

As promised, the study of generic OJI functions yields a better bound than (3.4.12). We recall that \mathfrak{s} stands for the notion of divergence we introduced in Definition 1.4.6.

THEOREM 3.4.9. *Consider $A \in \mathcal{A}$, $h : \mathcal{A} \rightarrow \mathbf{R}$ an OJI function with representation (3.4.13) in generic configuration in the vicinity of A and a sequence $(s_n)_{n \in \mathbf{N}}$ in Π such that $\mathfrak{s}(s_n, A) \rightarrow 0$ as $n \rightarrow \infty$. Then,*

$$\alpha_h(A) \leq \liminf_{n \rightarrow \infty} \frac{\log |J_{s_n}|}{\log \mathfrak{s}(s_n, A)} \vee 0.$$

Proof. The case $\alpha_h(A) = 0$ is trivial. Otherwise, take $\alpha \in (0, \alpha_h(A))$. Let us consider $\rho_{\alpha,0} > 0$ such that the estimate (3.4.4) holds for $k = 0$. Without loss of generality, we may suppose that for all n , $\mathfrak{J}(s_n, A) < \min\{1, \rho_0, \rho_{\alpha,0}\}$ where ρ_0 is the one of Definition 3.4.6. Let $\varepsilon > 0$ and $n \in \mathbf{N}$. The estimate (3.4.4) yields

$$\forall \underline{A}, \bar{A} \in B_{\mathcal{A}}(\mathfrak{J}(s_n, A) + \varepsilon), \quad |\Delta h(\bar{A} \setminus \underline{A})| \leq (\mathfrak{J}(s_n, A) + \varepsilon)^\alpha. \quad (3.4.18)$$

Since $s_n \in \Pi \cap \mathcal{V}(A, \mathfrak{J}(s_n, A) + \varepsilon)$ and h is in generic configuration, we may find \underline{A} and \bar{A} in $B_{\mathcal{A}}(\mathfrak{J}(s_n, A) + \varepsilon)$ such that

$$|\Delta h(\bar{A} \setminus \underline{A}) - J_{s_n}| \leq \varepsilon. \quad (3.4.19)$$

Combining (3.4.18) and (3.4.19) and taking $\varepsilon \rightarrow 0^+$ yields $|J_{s_n}| \leq \mathfrak{J}(s_n, A)^\alpha$ and thus

$$\alpha \leq \frac{\log |J_{s_n}|}{\log \mathfrak{J}(s_n, A)}.$$

The result follows from taking the lower limit as $n \rightarrow \infty$ and then $\alpha \rightarrow \alpha_h(A)^-$. \square

Following ideas from [51], let us introduce for any given map $h : \mathcal{A} \rightarrow \mathbf{R}$ with well-defined jumps, measurable set $L \subseteq \mathcal{T}$ and $\delta > 0$,

$$\forall j \in \mathbf{N}, \quad E_{j|L}^\delta(h) = \bigcup_{\substack{s \in \Pi(h) \cap L: \\ |J_s(h)| \in \Gamma_j}} \mathcal{V}'(s, |J_s(h)|^\delta) \quad (3.4.20)$$

where \mathcal{V}' is the dual vicinity given in (1.4.9) and

$$\forall j \in \mathbf{N}, \quad \Gamma_j = \{x \in \mathbf{R} : 2^{-j} \leq |x| < 2^{-(j-1)}\}. \quad (3.4.21)$$

Let us also introduce

$$E_{|L}^\delta(h) = \limsup_{j \rightarrow \infty} E_{j|L}^\delta(h) = \bigcap_{k \in \mathbf{N}} \bigcup_{j \geq k} E_{j|L}^\delta(h). \quad (3.4.22)$$

The set L allows to select the jumps of h falling in a specific region. It will prove to be useful when determining an upper bound for the Hölder regularity of the primitive process in Theorem 3.5.2. More precisely, we will jointly use Theorem 3.4.9 with the following result to get an upper bound very much like (11) in [51].

PROPOSITION 3.4.10. *If $h : \mathcal{A} \rightarrow \mathbf{R}$ has well-defined jumps and $A \in \mathcal{A}$, then*

$$A \in E_{|L}^\delta(h) \implies \liminf_{\substack{s \in \Pi(h) \cap L: \\ \mathfrak{J}(s, A) \rightarrow 0}} \frac{\log |J_s(h)|}{\log \mathfrak{J}(s, A)} \leq \frac{1}{\delta}.$$

Proof. Let $A \in E_{|L}^\delta(h)$. Then there exists an increasing sequence $(j_k)_{k \in \mathbf{N}}$ in \mathbf{N} and a sequence $(s_k)_{k \in \mathbf{N}} \in (\Pi(h) \cap L)^{\mathbf{N}}$ such that

$$\forall k \in \mathbf{N}, \quad 2^{-j_k} \leq |J_{s_k}(h)| < 2^{-(j_k-1)} \quad \text{and} \quad \mathfrak{J}(s_k, A) \leq |J_{s_k}(h)|^\delta.$$

So $\mathfrak{J}(s_k, A) \rightarrow 0$ and $|J_{s_k}(h)| \rightarrow 0^+$ as $k \rightarrow \infty$. In particular, there exists k_0 such that

$$\forall k \geq k_0, \quad \frac{\log |J_{s_k}(h)|}{\log \mathfrak{J}(s_k, A)} \leq \frac{1}{\delta}.$$

The result follows from taking the lower limit in the above inequality. \square

3.5 Hölder regularity of the primitive with respect to a siLévy

In the sequel, suppose that $\dim \mathcal{A} < \infty$, X is a siLévy, $f \in L_{\text{loc}}(X)$ and Y is the primitive process of f with respect to X . The goal of this section is to characterize the almost sure regularity of Y .

In Section 3.5.1, we divide the problem into smaller chunks. Section 3.5.2 gives bounds on the Hölder regularity of Y . Section 3.5.3 does the same work for the $d_{\mathcal{T}}$ -localized exponent. Finally, Section 3.5.4 is devoted to some examples and applications of the main results.

3.5.1 Divide and conquer

Let us take a closer look at the Lévy-Itô decomposition of Y (2.4.7). Recall that we have for some triplet (b, σ^2, ν) depending purely on X ,

$$\forall A \in \mathcal{A}, \quad Y_A = b \int_A f dm + \sigma \mathcal{W}(f \mathbb{1}_A) + \tilde{\mathcal{P}}(f \mathbb{1}_A)$$

where \mathcal{W} (resp. $\tilde{\mathcal{P}}$) is the linear process associated with a siBm (resp. a set-indexed compensated compound Poisson process).

Moreover, if $\int_{|x| \leq 1} |x| \nu(dx) < \infty$, we may rewrite (3.3.10) into

$$\forall A \in \mathcal{A}, \quad Y_A = b' \int_A f dm + \sigma \mathcal{W}(f \mathbb{1}_A) + \mathcal{P}(f \mathbb{1}_A)$$

where $b' = b - \int_{|x| \leq 1} x \nu(dx)$ and \mathcal{P} is the ‘non-compensated version’ of $\tilde{\mathcal{P}}$, i.e. the linear process associated to the set-indexed Poisson process of intensity $m \otimes \nu$ considered in Example 2.4.6.

In the case where $\mathcal{T} = \mathbf{R}_+$ and m is the Lebesgue measure, asking the regularity of the drift part of Y — the first term of (3.3.10) — is the same as asking the regularity of a primitive of f . This problem has been entirely dealt with through the use of a tool called the *2-microlocal frontier* which characterizes how the regularity evolves when one takes fractional integrals and/or derivatives of f . The 2-microlocal formalism dates back to [21] and has seen a lot of developments throughout the years (see for instance [49, 68, 74] and references therein). In order to do the same in the set-indexed setting, one would need to develop an analog for the 2-microlocal frontier. Being an entirely deterministic endeavor, we chose to push it aside for this article.

Namely, we are going to cancel the drifts in the expression of Y (i.e. take either $b = 0$ or $b' = 0$ depending on the case), giving the following simpler expressions for $A \in \mathcal{A}$:

$$Y_A = \begin{cases} \sigma \mathcal{W}(f \mathbb{1}_A) + \tilde{\mathcal{P}}(f \mathbb{1}_A) & \text{if } \int_{|x| \leq 1} |x| \nu(dx) = \infty, \\ \sigma \mathcal{W}(f \mathbb{1}_A) + \mathcal{P}(f \mathbb{1}_A) & \text{if } \int_{|x| \leq 1} |x| \nu(dx) < \infty \end{cases} \quad (3.5.1)$$

Hence we are going to express the regularity of Y as given in (3.5.1) in terms of (σ^2, ν) — which characterizes the law of X — and f . The following result explains that we may treat the Gaussian and Poissonian components of Y separately.

PROPOSITION 3.5.1. *Let $A \in \mathcal{A}$. If $m(\mathcal{V}(A, \rho)) \rightarrow 0$ as $\rho \rightarrow 0$, then $\alpha_Y(A)$ is a.s. deterministic and*

$$\alpha_Y(A) = \alpha_{\sigma\mathfrak{W}(f\mathbb{1}_A)}(A) \wedge \alpha_{\mathfrak{P}(f\mathbb{1}_A)}(A) \quad \text{a.s.}$$

The same holds if one replaces \mathfrak{P} by $\tilde{\mathfrak{P}}$ and/or the Hölder exponents $\alpha(A)$ by the corresponding $d_{\mathcal{T}}$ -localized exponents $\alpha_{\cdot, d_{\mathcal{T}}}(A)$.

The condition on the vicinity in Proposition 3.5.1 is not innocuous. The speed at which $m(\mathcal{V}(A, \rho))$ goes to 0 has actually a great influence on the regularity of Y (see Corollaries 3.5.13 and 3.5.14).

Proof of Proposition 3.5.1. The other proofs being similar, we only prove it for \mathfrak{P} and the Hölder exponents $\alpha(A)$.

Let $A \in \mathcal{A}$. Since $\alpha_Y(A)$, $\alpha_{\sigma\mathfrak{W}(f\mathbb{1}_A)}(A)$ and $\alpha_{\mathfrak{P}(f\mathbb{1}_A)}(A)$ are \mathcal{F}_{A^+} -measurable, they are deterministic according to Theorem 2.3.33. Thus, there exist $\alpha_Y, \alpha_W, \alpha_P \in \mathbf{R}_+ \cup \{\infty\}$ such that the event

$$\Omega^* = \{\alpha_Y(A) = \alpha_Y\} \cap \{\alpha_{\sigma\mathfrak{W}(f\mathbb{1}_A)}(A) = \alpha_W\} \cap \{\alpha_{\mathfrak{P}(f\mathbb{1}_A)}(A) = \alpha_P\}$$

happens with probability one.

It is a classical, deterministic result that $\alpha_Y \geq \alpha_W \wedge \alpha_P$ and equality happens as soon as $\alpha_W \neq \alpha_P$. Suppose that $\alpha_Y > \alpha_W = \alpha_P$ and let us show that a contradiction follows. Consider $\alpha = (\alpha_W + \alpha_P)/2$. Since $\alpha > \alpha_P$, there exists (on Ω^*) a random sequence $(A_k)_{k \in \mathbf{N}}$ in \mathcal{A} such that $d_{\mathcal{A}}(A, A_k) \rightarrow 0^+$ as $k \rightarrow \infty$ and

$$\lim_{k \rightarrow \infty} \frac{|\mathfrak{P}(f\mathbb{1}_A) - \mathfrak{P}(f\mathbb{1}_{A_k})|}{d_{\mathcal{A}}(A, A_k)^\alpha} = +\infty \quad \text{a.s.}$$

Instead of extracting a subsequence and potentially replacing Y by $-Y$, we might as well suppose that there exists a \mathfrak{P} -measurable event $\Omega_p \subseteq \Omega^*$ of positive probability such that

$$\lim_{k \rightarrow \infty} \frac{\mathfrak{P}(f\mathbb{1}_A) - \mathfrak{P}(f\mathbb{1}_{A_k})}{d_{\mathcal{A}}(A, A_k)^\alpha} = +\infty \quad \mathbf{P}(\cdot | \Omega_p)\text{-a.s.} \quad (3.5.2)$$

Since $0 < \alpha < \alpha_Y$, (3.5.2) implies that

$$\lim_{k \rightarrow \infty} \frac{\sigma\mathfrak{W}(f\mathbb{1}_A) - \sigma\mathfrak{W}(f\mathbb{1}_{A_k})}{d_{\mathcal{A}}(A, A_k)^\alpha} = -\infty \quad \mathbf{P}(\cdot | \Omega_p)\text{-a.s.} \quad (3.5.3)$$

in order to compensate for the divergence. Since $(\mathfrak{W}, \mathfrak{P}) \stackrel{\text{law}}{\equiv} (-\mathfrak{W}, \mathfrak{P})$ and the sequence $(A_k)_{k \in \mathbf{N}}$ only depends on \mathfrak{P} , we may replace \mathfrak{W} by $-\mathfrak{W}$ in (3.5.3) and obtain

$$\lim_{k \rightarrow \infty} \frac{\sigma\mathfrak{W}(f\mathbb{1}_A) - \sigma\mathfrak{W}(f\mathbb{1}_{A_k})}{d_{\mathcal{A}}(A, A_k)^\alpha} = +\infty \quad \mathbf{P}(\cdot | \Omega_p)\text{-a.s.}$$

which contradicts (3.5.3) since $\mathbf{P}(\Omega_p) > 0$. \square

3.5.2 Hölder regularity

As explained in Proposition 3.5.1, treating separately the two cases $\nu = 0$ and $\sigma^2 = 0$ is enough to obtain a complete characterization of the Hölder regularity of Y (apart from the drift, which was supposed to be zero in Section 3.5.1).

The Gaussian part

The case where $\nu = 0$ has already been treated at great lengths in the literature. For the set-indexed case, [42, Corollary 5.3] ensures that under some entropic condition similar to Dudley's, for all $A \in \mathcal{A}$, the following holds with probability one:

$$\alpha_Y(A) = \frac{1}{2} \alpha_{\sigma \int_{A \Delta A'} f^2 d\mathbf{m}}(A). \quad (3.5.4)$$

As already pointed out in Section 3.5.1, the Hölder exponent for $A' \mapsto \int_{A \Delta A'} f^2 d\mathbf{m}$ cannot be deduced solely from the Hölder exponent of f . The knowledge of some kind of 2-microlocal frontier is required, hence (3.5.4) cannot be readily improved. More precise results for the one-dimensional gaussian case concerning the 2-microlocal frontier especially adapted to our exponent are given in [15].

The Poissonian part

In this part, let us suppose that Y is purely Poissonian, *i.e.* that we take $\sigma^2 = 0$ in the Lévy-Itô decomposition (3.5.1). The first study of Hölder regularity of a purely Poissonian Lévy process happened in [20] where Blumenthal and Gettoor determined the value of the pointwise Hölder exponent $\alpha_X(0)$ as defined by (3.4.1). As explained in Section 3.4.1, there seems not to exist any natural extension of polynomials to the set-indexed setting, hence our choice to subtract by hand the 'polynomial part' in the same way as [20]. The authors also introduced the so-called *Blumenthal-Gettoor exponent*:

$$\beta = \inf \left\{ \delta > 0 : \int_{|x| \leq 1} |x|^\delta \nu(dx) < \infty \right\}. \quad (3.5.5)$$

Remark that since ν is a Lévy measure, $\beta \in [0, 2]$. For $\mathcal{T} = \mathbf{R}_+$, Blumenthal and Gettoor proved in particular that $\alpha_Y(0) = 1/\beta$ almost surely together with the convention $1/0 = +\infty$. This result has been extended in [13, 51, 82] in much greater detail. Our goal is to extend those results in the case of the integral process Y , which to our knowledge has not been done even in \mathbf{R}_+ , and for possibly different spaces than \mathbf{R}_+ .

The coming Theorem 3.5.2 will tell us that the Hölder exponent of Y at some $A \in \mathcal{A}$ is governed by the regularity of both f and X . For X , the exponent β and some information about the vicinity of A will suffice (Corollary 3.5.13). However for Y , we need to know about the behavior of f in the vicinity of A . That is the reason why $L_{f,\alpha}(A)$ and $L_{f,\alpha}(A, \rho)$ are introduced below and correspond to the 'irregular part' of f . As for the sets $\bar{R}_f(A)$ and $\underline{R}_f(A)$, they determine the proportion of the vicinity where f is indeed irregular.

Recall that the divergence \mathfrak{s} and the vicinity $\mathcal{V}(A, \rho)$ have been given in Definition 1.4.6. We then define for all $A \in \mathcal{A}$, $\alpha \geq 0$ and $\rho > 0$,

$$\begin{aligned} L_{f,\alpha}(A) &= \{s \in \mathcal{T} : |f(s)| > \mathfrak{s}(s, A)^\alpha\}, \\ L_{f,\alpha}(A, \rho) &= \mathcal{V}(A, \rho) \cap L_{f,\alpha}(A), \\ L_{f,\alpha}^c(A, \rho) &= \mathcal{V}(A, \rho) \setminus L_{f,\alpha}(A) \end{aligned} \quad (3.5.6)$$

and

$$\begin{aligned} \bar{R}_f(A) &= \left\{ (\alpha, q) \in \mathbf{R}_+^2 : \liminf_{\rho \rightarrow 0^+} \frac{\mathbf{m}(L_{f,\alpha}(A, \rho))}{\rho^q} > 0 \right\}, \\ \underline{R}_f(A) &= \left\{ (\alpha, q, q') \in \mathbf{R}_+^3 : \limsup_{\rho \rightarrow 0^+} \left[\frac{\mathbf{m}(L_{f,\alpha}(A, \rho))}{\rho^q} + \frac{\mathbf{m}(L_{f,\alpha}^c(A, \rho))}{\rho^{q'}} \right] < \infty \right\}. \end{aligned} \quad (3.5.7)$$

It is not too surprising that we take the irregularity of f into account only through the measure \mathbf{m} since the integral with respect to X does not differentiate between \mathbf{m} -a.e. equal functions. Comparing the measure of a set ‘of radius ρ ’ to a power of ρ is not a new idea per say and may remind the reader about *Ahlfors-David regular* measures. Likewise in the stochastic field, measuring the ‘local density’ of a set of interest to get regularity results is not new (see [1, Section 4.6] and references therein).

THEOREM 3.5.2 (Hölder regularity of a Poissonian siLévy’s primitive process). *Recall that we supposed $\dim \mathcal{A} < \infty$, $\sigma^2 = 0$ and let $A \in \mathcal{A}$. Suppose that f is bounded in the vicinity of A (i.e. $\sup_{s \in \mathcal{V}(A, \rho_0)} |f(s)| < \infty$ for some $\rho_0 > 0$). Then the following holds with probability one:*

$$\sup_{(\alpha, q, q') \in \underline{R}_f(A)} \min \left\{ \frac{q}{\beta}, \frac{q'}{\beta} + \alpha \right\} \leq \alpha_Y(A) \leq \inf_{(\alpha, q) \in \bar{R}_f(A)} \left\{ \frac{q}{\beta} + \alpha \right\}$$

with the conventions that $\inf \emptyset = 1/0 = +\infty$.

We will comment on this result in Section 3.6.2 and apply it in Section 3.5.4. The proof is divided into several parts: the upper bound, the lower bound when $\beta \geq 1$ and the lower bound when $\beta < 1$.

Proof of the upper bound in Theorem 3.5.2. The upper bound is very similar in spirit as [51] and the key step is the covering argument given by Proposition 3.5.5. Remark that if $\beta = 0$, the upper bound is trivial. So let us suppose for this part that $\beta > 0$.

For now, let us consider some measurable set $L \subseteq \mathcal{T}$ of finite positive \mathbf{m} -measure and denote for all $j \in \mathbf{N}$,

$$\Pi_{j|L} = L \cap \{s \in \Pi(X) : J_s(X) \in \Gamma_j\}, \quad (3.5.8)$$

$$\nu_j = \nu(\Gamma_j) \quad (3.5.9)$$

where $\Gamma_j = \{x \in \mathbf{R} : 2^{-j} \leq |x| < 2^{-(j-1)}\}$ has been introduced in (3.4.21).

LEMMA 3.5.3. *Fix $\gamma < \beta$. There exists an increasing sequence $(j_k)_{k \in \mathbf{N}}$ in \mathbf{N} such that*

$$\nu_j = O(8^j) \quad \text{as } j \rightarrow \infty \quad (3.5.10)$$

$$2^{jk\gamma} = O(\nu_{j_k}) \quad \text{as } k \rightarrow \infty \quad (3.5.11)$$

Proof. Since $\beta > 0$, the convergence of $\int_{|x| \leq 1} |x|^\gamma \nu(dx)$ is equivalent to the convergence of $\sum_{j=0}^{\infty} \nu_j 2^{-j\gamma}$. The Cauchy-Hadamard formula for the radius of convergence of power series then gives

$$\beta = \limsup_{j \rightarrow \infty} \frac{\log_2 \nu_j}{j} \quad (3.5.12)$$

which is a relation that was already noted (but not explicitly proven) in [51]. The estimate (3.5.10) (resp. (3.5.11)) then follows from this formula and the fact that $\beta < 3$ (resp. $\gamma < \beta$). \square

Recall that the random sets $E_{j|L}^\delta(h)$ and $E_{|L}^\delta(h)$ have been introduced in (3.4.20) and (3.4.22) respectively.

LEMMA 3.5.4. *Let $A \in \mathcal{A}$, then for all $\delta > 0$ and $j \in \mathbf{N}$,*

$$\exp[-\nu_j \mathbf{m}(L \cap \mathcal{V}(A, 2^{-j\delta})] \leq \mathbf{P}(A \notin E_{j|L}^\delta(X)) \leq \exp[-\nu_j \mathbf{m}(L \cap \mathcal{V}(A, 2^{-(j+1)\delta})].$$

Proof. We only prove the upper bound, the lower bound being proven in exactly the same way. Fix $\delta > 0$ and $j \in \mathbf{N}$. Then, by definition of \mathcal{V} and \mathcal{V}' ,

$$\begin{aligned} \mathbf{P}(A \notin E_{j|L}^\delta(X)) &= \mathbf{P}(\forall s \in \Pi_{j|L}, A \notin \mathcal{V}'(s, |J_s(X)|^\delta)) \\ &\leq \mathbf{P}(\forall s \in \Pi_{j|L}, A \notin \mathcal{V}'(s, 2^{-(j+1)\delta})) \\ &= \mathbf{P}(\forall s \in \Pi_{j|L}, s \notin \mathcal{V}(A, 2^{-(j+1)\delta})). \end{aligned}$$

Since $\Pi_{j|L}$ is a Poisson random set of intensity measure $\nu_j \mathbf{m}(L \cap \cdot)$, we may write $\Pi_{j|L} = \{s_1, \dots, s_{M_j}\}$ where M_j is a Poisson random variable of intensity $\nu_j \mathbf{m}(L)$ and the s_i 's are iid variables of distribution $\mathbf{m}(L \cap \cdot) / \mathbf{m}(L)$ independent from M_j . Hence, conditioning with respect to M_j yields

$$\begin{aligned} \mathbf{P}(\forall s \in \Pi_{j|L}, s \notin \mathcal{V}(A, 2^{-(j+1)\delta})) &= \mathbf{E}[\mathbf{P}(\forall i \in \llbracket 1, M_j \rrbracket, s_i \notin \mathcal{V}(A, 2^{-(j+1)\delta}) \mid M_j)] \\ &= \mathbf{E}[\mathbf{P}(s_1 \notin \mathcal{V}(A, 2^{-(j+1)\delta}))^{M_j}] \\ &= \exp[\nu_j \mathbf{m}(L) (\mathbf{P}(s_1 \notin \mathcal{V}(A, 2^{-(j+1)\delta})) - 1)] \\ &= \exp[-\nu_j \mathbf{m}(L \cap \mathcal{V}(A, 2^{-(j+1)\delta})]. \end{aligned}$$

The result follows. \square

PROPOSITION 3.5.5. *Let $A \in \mathcal{A}$ and suppose that there exists $q > 0$ such that*

$$\rho^q = O(\mathbf{m}(L \cap \mathcal{V}(A, \rho))) \quad \text{as } \rho \rightarrow 0. \quad (3.5.13)$$

Then for all $\delta < \beta/q$, A belongs to $E_{|L}^\delta(X)$ with probability one.

Proof. Fix $\delta < \beta/q$ and $A \in \mathcal{A}$. Then, by (3.5.13) and Lemma 3.5.4, there exists $\kappa > 0$ such that

$$\forall j \in \mathbf{N}, \quad \mathbf{P}(A \notin E_{j|L}^\delta(X)) \leq \exp(-\kappa \nu_j 2^{-j\delta q}).$$

Hence, taking $\gamma \in (\delta q, \beta)$ and using the sequence $(j_k)_{k \in \mathbf{N}}$ of Lemma 3.5.3, we get

$$\forall k \in \mathbf{N}, \quad \mathbf{P}(A \notin E_{j_k|L}^\delta(X)) \leq \exp(-\kappa 2^{j_k(\gamma - \delta q)})$$

which is a convergent series. The result follows from Borel-Cantelli lemma. \square

PROPOSITION 3.5.6. *For all $(\alpha, q) \in \bar{R}_f(A)$, we have*

$$\alpha_Y(A) \leq \frac{q}{\beta} + \alpha \quad \text{a.s.}$$

Proof. The following holds with probability one.

$$\begin{aligned}
\alpha_Y(A) &\leq \liminf_{\substack{s \in \Pi(Y): \\ \delta(s,A) \rightarrow 0}} \frac{\log |J_s(Y)|}{\log \delta(s,A)} && \text{by Theorem 3.4.9,} \\
&= \liminf_{\substack{s \in \Pi(Y): \\ \delta(s,A) \rightarrow 0}} \left[\frac{\log |J_s(X)|}{\log \delta(s,A)} + \frac{\log |f(s)|}{\log \delta(s,A)} \right] && \text{by Corollary 3.3.12,} \\
&\leq \underbrace{\liminf_{\substack{s \in \Pi(Y) \cap L_{f,\alpha}(A): \\ \delta(s,A) \rightarrow 0}} \frac{\log |J_s(X)|}{\log \delta(s,A)}}_{\leq 1/\delta} + \underbrace{\limsup_{\substack{s \in L_{f,\alpha}(A): \\ \delta(s,A) \rightarrow 0}} \frac{\log |f(s)|}{\log \delta(s,A)}}_{\leq \alpha}
\end{aligned}$$

where the last two inequalities are due to Propositions 3.4.10 and 3.5.5 and the definition of $L_{f,\alpha}(A)$. \square

The upper bound of Theorem 3.5.2 may be readily deduced from Proposition 3.5.6 by taking a relevant sequence converging to the claimed upper bound.

Proof of the lower bound in Theorem 3.5.2 when $\beta \geq 1$. The main idea of this part relies on an application of Borel-Cantelli lemma and relevant estimates on the corresponding probabilities.

First, recall that we gave the definitions of the space $L^\gamma(b, \sigma^2, \nu)$ and the function ϕ_γ at the beginning of Section 2.4.2. A consequence of the celebrated result of Rajput and Rosiński [83, Theorem 3.3] is the continuity of the linear map $\mathcal{X}|_{L^\gamma(b, \sigma^2, \nu)}$ from $L^\gamma(b, \sigma^2, \nu)$ to $L^\gamma(\Omega)$ for all $\gamma \in (1, 2]$, i.e. there exists a constant $\kappa_{\phi_\gamma} > 0$ such that

$$\forall g \in L^\gamma(b, \sigma^2, \nu), \quad \|\mathcal{X}(g)\|_{L^\gamma(\Omega)} \leq \kappa_{\phi_\gamma} \|g\|_{\phi_\gamma} \quad (3.5.14)$$

where $\|\cdot\|_{\phi_\gamma}$ is the Luxemburg norm defined in (2.2.2). We further simplify the use of the Luxemburg by means of the following lemma.

LEMMA 3.5.7. *For any $\gamma \in (1, 2]$ and measurable map $g : \mathcal{T} \rightarrow \mathbf{R}$,*

$$\|g\|_{\phi_\gamma} \leq \left[\int_{\mathbf{R}} |x|^\gamma \nu(dx) \right]^{1/\gamma} \|g\|_{L^\gamma(m)}.$$

Proof. In the formulation of the L^γ -space as an Orlicz space, we have

$$\|g\|_{L^\gamma(m)} = \inf \left\{ c > 0 : \int_{\mathcal{T}} c^{-\gamma} |g(s)|^\gamma \mathbf{m}(ds) \leq 1 \right\}.$$

Comparing this norm with (2.2.2), it follows that it is enough to prove the following:

$$\forall c > 0, \quad \int_{\mathcal{T}} \phi(c^{-1}g(s)) \mathbf{m}(ds) \leq c^{-\gamma} \int_{\mathcal{T}} |g(s)|^\gamma \mathbf{m}(ds) \int_{\mathbf{R}} |x|^\gamma \nu(dx). \quad (3.5.15)$$

Let $c > 0$ and $s \in \mathcal{T}$. Then,

$$\begin{aligned}
\Phi(c^{-1}g(s)) &= \int_{|xg(s)| \leq c} |c^{-1}xg(s)|^2 \nu(dx) + \int_{|xg(s)| > c} |c^{-1}xg(s)|^\gamma \nu(dx) \\
&= c^{-2} \int_{|xg(s)| \leq c} |xg(s)|^{2-\gamma} |xg(s)|^\gamma \nu(dx) + c^{-\gamma} \int_{|xg(s)| > c} |xg(s)|^\gamma \nu(dx) \\
&\leq c^{-2} c^{2-\gamma} \int_{|xg(s)| \leq c} |xg(s)|^\gamma \nu(dx) + c^{-\gamma} \int_{|xg(s)| > c} |xg(s)|^\gamma \nu(dx) \\
&= c^{-\gamma} |g(s)|^\gamma \int_{\mathbf{R}} |x|^\gamma \nu(dx).
\end{aligned}$$

Integrating with respect to \mathbf{m} yields (3.5.15), from which the result follows. \square

Recall that we supposed $\beta \geq 1$. From (3.5.14) and Lemma 3.5.7, it follows that for all $\gamma \in (\beta, 2]$ (or $\gamma = 2$ if $\beta = 2$), there exists a finite constant $\kappa_\gamma > 0$ such that

$$\forall B \in \mathcal{B}_m, \quad \mathbf{E}[|\Delta Y_B|^\gamma] \leq \kappa_\gamma \|f \mathbb{1}_B\|_{L^\gamma(\mathbf{m})}^\gamma \quad (3.5.16)$$

where $\Delta Y_B = \mathcal{X}(f \mathbb{1}_B)$.

We are now ready to proceed to the lower bound itself.

PROPOSITION 3.5.8. *Suppose that the hypotheses of Theorem 3.5.2 and $\beta \geq 1$ hold. Then, for all $(\alpha, q, q') \in \underline{R}_f(A)$,*

$$\alpha_\gamma(A) \geq \min \left\{ \frac{q}{\beta}, \frac{q'}{\beta} + \alpha \right\} \quad a.s.$$

Proof. Fix $(\alpha, q, q') \in \underline{R}_f(A)$. We will only prove the result in the case $q'/\beta + \alpha \leq q/\beta$. The second case is proven in exactly the same fashion, one just has to replace q' by q and take $\alpha = 0$ in the following.

Let $\delta > \beta/q'$ and $\eta = 1/\delta + \alpha$. By Borel-Cantelli, it is enough to prove that

$$\sum_{j=1}^{\infty} \mathbf{P} \left(\sup_{d_{\mathcal{A}}(A, A') < 2^{-j}} |Y_A - Y_{A'}| > 2^{-j\eta} \right) < \infty. \quad (3.5.17)$$

Fix $\gamma \in (\beta, 2]$ (or $\gamma = 2$ if $\beta = 2$) and denote $p = \dim \mathcal{A}$. By Corollary 2.3.31, we get for all $j \in \mathbf{N}$,

$$\mathbf{P} \left(\sup_{d_{\mathcal{A}}(A, A') < 2^{-j}} |Y_A - Y_{A'}| > 2^{-j\eta} \right) \leq \kappa_{p,\gamma} 2^{j\eta\gamma} \mathbf{E} \left[|\Delta Y_{\mathcal{V}(A, 2^{-j})}|^\gamma \right]. \quad (3.5.18)$$

It follows from (3.5.16) and (3.5.18) that

$$\mathbf{P} \left(\sup_{d_{\mathcal{A}}(A, A') < 2^{-j}} |Y_A - Y_{A'}| > 2^{-j\eta} \right) = O \left(2^{j\eta\gamma} \|f \mathbb{1}_{\mathcal{V}(A, 2^{-j})}\|_{L^\gamma(\mathbf{m})}^\gamma \right) \quad \text{as } j \rightarrow \infty. \quad (3.5.19)$$

We split $\mathcal{V}(A, 2^{-j})$ in (3.5.19) into $L_{f,\alpha}(A, 2^{-j})$ and $L_{f,\alpha}^c(A, 2^{-j})$ to control what is happening.

On the one hand, since $(\alpha, q, q') \in \underline{R}_f(A)$, we have $\mathbf{m}(L_{f,\alpha}(A, \rho)) = O(\rho^q)$ as $\rho \rightarrow 0$. Moreover, f is also bounded in the vicinity of A . Thus

$$\|f \mathbb{1}_{L_{f,\alpha}(A, \rho)}\|_{L^\gamma(\mathbf{m})}^\gamma = O(\rho^q) \quad \text{as } \rho \rightarrow 0. \quad (3.5.20)$$

On the other hand, using $(\alpha, q, q') \in \underline{R}_f(A)$ once more yields $m(L_{f,\alpha}^{\mathbb{c}}(A, \rho)) = O(\rho^{q'})$ as $\rho \rightarrow 0$. Thus

$$\|f \mathbb{1}_{L_{f,\alpha}^{\mathbb{c}}(A, \rho)}\|_{L^{\gamma}(m)}^{\gamma} = O(\rho^{q'+\alpha\gamma}) \quad \text{as } \rho \rightarrow 0. \quad (3.5.21)$$

Since we supposed that $q'/\beta + \alpha \leq q/\beta$ and if we take γ close enough to β , the estimates (3.5.20) and (3.5.21) give together

$$\|f \mathbb{1}_{\mathcal{V}(A, \rho)}\|_{L^{\gamma}(m)}^{\gamma} = O(\rho^{q'+\alpha\gamma}) \quad \text{as } \rho \rightarrow 0. \quad (3.5.22)$$

Combining (3.5.19) and (3.5.22) yields

$$\mathbf{P}\left(\sup_{d_{\mathcal{A}}(A, A') < 2^{-j}} |Y_A - Y_{A'}| > 2^{-j\eta}\right) = O(2^{-j(q'+(\alpha-\eta)\gamma)}) \quad \text{as } j \rightarrow \infty. \quad (3.5.23)$$

Since

$$q' + (\alpha - \eta)\gamma \rightarrow q' - \frac{\beta}{\delta} \quad \text{as } \gamma \rightarrow \beta^+$$

and $q' - \beta/\delta > 0$, we may take γ close enough to β in (3.5.23) showing that (3.5.17) holds. The result follows. \square

Proof of the lower bound in Theorem 3.5.2 when $\beta < 1$. Since we have here $\sigma = 0$ and $\int_{|x| \leq 1} |x| \nu(dx) < \infty$, we get from (3.5.1) the following expression:

$$\forall A \in \mathcal{A}, \quad Y_A = \mathcal{P}(f \mathbb{1}_A) = \lim_{\varepsilon \rightarrow 0^+} \left[\sum_{\substack{t \in \Pi(X) \cap A: \\ |J_t(X)| \geq \varepsilon}} f(t) J_t(X) \right] \quad (3.5.24)$$

where the limit as $\varepsilon \rightarrow 0^+$ happens a.s. uniformly in $A \in \mathcal{T}_n$ for all $n \in \mathbf{N}$.

Two problems arise when trying to copy the proof for $\beta \geq 1$: Y is not a martingale anymore, so the maximal inequality from Corollary 2.3.31 does not hold anymore, and the Luxemburg norm $\|\cdot\|_{\phi_{\gamma}}$ is harder to control. Instead, we introduce the set-indexed process $Z_{|L}$ given by

$$\forall A \in \mathcal{A}, \quad (Z_{|L})_A = \lim_{\varepsilon \rightarrow 0^+} \left[\sum_{\substack{t \in \Pi(X) \cap L \cap A: \\ |J_t(X)| \geq \varepsilon}} |J_t(X)| \right] \quad (3.5.25)$$

where $L \subseteq \mathcal{T}$ is once more a free parameter set that will be chosen later on. Remark that $Z_{|L}$ is a siLévy, but with respect to $m(L \cap \cdot)$ instead of m as stationarity measure.

We first establish a lower bound on the regularity of $Z_{|L}$ and then deduce one for the regularity of Y .

The following method is inspired from [13]. Denote for all $\eta \geq 0$, the ‘truncated’ set-indexed process $(Z_{|L})^{\eta}$ given by

$$\forall A \in \mathcal{A}, \quad (Z_{|L})_A^{\eta} = \lim_{\varepsilon \rightarrow 0^+} \left[\sum_{\substack{t \in \Pi \cap L \cap A: \\ \varepsilon \leq |J_t(X)| < 2^{-\eta}}} |J_t(X)| \right] \quad (3.5.26)$$

LEMMA 3.5.9. Let $\mathcal{A} \in \mathcal{A}$ and suppose that there exists $q > 0$ such that

$$\mathbf{m}(L \cap \mathcal{V}(A, \rho)) = O(\rho^q) \quad \text{as} \quad \rho \rightarrow 0. \quad (3.5.27)$$

Then for all $\delta > \beta/q$, there exists a constant $\kappa_\delta > 0$ such that

$$\forall j \in \mathbf{N}, \quad \mathbf{P} \left(\sup_{A' \in \mathcal{B}_{\mathcal{A}}(A, 2^{-j})} \Delta(Z|_L)_{A\Delta A'}^{j/\delta} \geq j 2^{-j/\delta} \right) \leq \kappa_\delta e^{-j}.$$

Proof. The proof is inspired from [13, Lemma 2.1], but we chose to give the details for the sake of completeness. Fix $j \in \mathbf{N}$. Then, since Z is non-negative,

$$\forall A' \in \mathcal{B}_{\mathcal{A}}(A, 2^{-j}), \quad \Delta(Z|_L)_{A\Delta A'}^{j/\delta} \leq \Delta(\mathbf{Z}|_L)_{\mathcal{V}(A, 2^{-j})}^{j/\delta}.$$

Hence, by Markov's inequality,

$$\mathbf{P} \left(\sup_{A' \in \mathcal{B}_{\mathcal{A}}(A, 2^{-j})} \Delta(Z|_L)_{A\Delta A'}^{j/\delta} \geq j 2^{-j/\delta} \right) \leq e^{-j} \mathbf{E} \left[\exp \left(2^{j/\delta} \Delta(\mathbf{Z}|_L)_{\mathcal{V}(A, 2^{-j})}^{j/\delta} \right) \right].$$

So we just need to prove that $\mathbf{E} \left[\exp \left(2^{j/\delta} \Delta(\mathbf{Z}|_L)_{\mathcal{V}(A, 2^{-j})}^{j/\delta} \right) \right] = O(1)$ as $j \rightarrow \infty$ to conclude. Using (3.5.26), we may compute the Laplace transform of $\Delta(\mathbf{Z}|_L)_{\mathcal{V}(A, 2^{-j})}^{j/\delta}$ and get

$$\mathbf{E} \left[\exp \left(2^{j/\delta} \Delta(\mathbf{Z}|_L)_{\mathcal{V}(A, 2^{-j})}^{j/\delta} \right) \right] = \exp \left(2^{j/\delta} \mathbf{m}(L \cap \mathcal{V}(A, 2^{-j})) \int_{|x| \leq 2^{-j/\delta}} (e^{|x|} - 1) \nu(dx) \right).$$

Hence, due to (3.5.27) and the fact that $e^{|x|} - 1 \leq 2|x|$ for $|x| \leq 1$, there exists a constant $\kappa > 0$ such that

$$\mathbf{E} \left[\exp \left(2^{j/\delta} \Delta(\mathbf{Z}|_L)_{\mathcal{V}(A, 2^{-j})}^{j/\delta} \right) \right] \leq \exp \left(\kappa 2^{-j(q-1/\delta)} \int_{|x| \leq 2^{-j/\delta}} |x| \nu(dx) \right).$$

Taking $\gamma \in (\beta, 1 \wedge \delta q)$ gives

$$\begin{aligned} \mathbf{E} \left[\exp \left(2^{j/\delta} \Delta(\mathbf{Z}|_L)_{\mathcal{V}(A, 2^{-j})}^{j/\delta} \right) \right] &\leq \exp \left(\kappa 2^{-j(q-1/\delta)} \int_{|x| \leq 2^{-j/\delta}} |x|^{1-\gamma} |x|^\gamma \nu(dx) \right) \\ &\leq \exp \left(\kappa 2^{-j(q-1/\delta+(1-\gamma)/\delta)} \int_{|x| \leq 2^{-j/\delta}} |x|^\gamma \nu(dx) \right) \\ &\leq \exp \left(\kappa 2^{-j(q-\gamma/\delta)} \int_{|x| \leq 1} |x|^\gamma \nu(dx) \right). \end{aligned}$$

Since $q - \gamma/\delta \rightarrow q - \beta/\delta$ as $\gamma \rightarrow \beta^+$ and $q - \beta/\delta > 0$, we may find $\gamma > \beta$ showing that the above expression is bounded as $j \rightarrow \infty$. The result follows. \square

LEMMA 3.5.10. Let $A \in \mathcal{A}$ and suppose that (3.5.27) holds. Then, the following holds with probability one: for all $\delta > \beta/q$, there exists $\rho_\delta > 0$ such that

$$\forall \rho \in (0, \rho_\delta), \forall A' \in \mathcal{B}_{\mathcal{A}}(A, \rho), \quad (\Delta Z|_L)_{A\Delta A'} \leq \rho^{1/\delta}.$$

Proof. Let us fix $\delta > \beta/q$. It is enough to prove that the following holds with probability one: there exists $k \in \mathbf{N}$ such that

$$\forall j \geq k, \forall A' \in B_{\mathcal{A}}(A, 2^{-j}), \quad (\Delta Z|_L)_{A\Delta A'} \leq 2^{-j/\delta}.$$

According to Lemma 3.5.9, this is already true if we replace $(\Delta Z|_L)$ by $(\Delta Z|_L)^{j/\delta}$. Thus, it is enough to prove that the following holds with probability one: there exists $k \in \mathbf{N}$ such that for all $j \geq k$ and $A' \in B_{\mathcal{A}}(A, 2^{-j})$, $(\Delta Z|_L)_{A\Delta A'} = (Z|_L)_{A\Delta A'}^{j/\delta}$. In other words, we want to show that $A \notin E_{|L}^\delta(X)$ almost surely. By Lemma 3.5.4, we have for all $j \in \mathbf{N}$,

$$\begin{aligned} \mathbf{P}(A \in E_{|L}^\delta(X)) &\leq 1 - \exp[-\nu_j \mathbf{m}(L \cap \mathcal{V}(A, 2^{-j\delta}))] \\ &\leq \nu_j \mathbf{m}(L \cap \mathcal{V}(A, 2^{-j\delta})) && \text{since } 1 - e^{-x} \leq x, \\ &= O(\nu_j 2^{-j\delta q}) && \text{by (3.5.27).} \end{aligned}$$

Taking $\gamma \in (\beta, \delta q)$ and using (3.5.12) yields

$$\mathbf{P}(A \in E_{|L}^\delta(X)) = O(2^{-j(\delta q - \gamma)}) \quad \text{as } j \rightarrow \infty$$

which then is a convergent series. Thus, by Borel-Cantelli, $A \notin E_{|L}^\delta(X)$ with probability one. The result follows. \square

PROPOSITION 3.5.11. *Suppose that the hypotheses of Theorem 3.5.2 and $\beta < 1$ hold. Then, for all $(\alpha, q, q') \in \underline{R}_f(A)$,*

$$\alpha_\gamma(A) \geq \min \left\{ \frac{q}{\beta}, \frac{q'}{\beta} + \alpha \right\} \quad \text{a.s.}$$

Proof. Fix $(\alpha, q, q') \in \underline{R}_f(A)$. By definition of $L_{f,\alpha}^c(A, \rho)$, we have

$$\forall s \in L_{f,\alpha}^c(A, \rho), \quad |f(s)| \leq \rho^\alpha$$

for all $\rho > 0$ small enough.

Likewise, since f is bounded in the vicinity of A , there exists $\kappa_f > 0$ such that

$$\forall s \in L_{f,\alpha}(A, \rho), \quad |f(s)| \leq \kappa_f$$

for all $\rho > 0$ small enough. Combining those estimates on f with the expression (3.5.24) of Y yields for all $\rho > 0$ small enough

$$\forall A' \in B_{\mathcal{A}}(A, \rho), \quad |Y_A - Y_{A'}| \leq \rho^\alpha (\Delta Z|_{L_{f,\alpha}(A)^\varepsilon})_{A\Delta A'} + \kappa_f (\Delta Z|_{L_{f,\alpha}(A)})_{A\Delta A'}.$$

Since $(\alpha, q, q') \in \underline{R}_f(A)$, we may apply Lemma 3.5.10 to both $L = L_{f,\alpha}$ and $L = L_{f,\alpha}^c$ (for which (3.5.27) holds if we replace q by q'). Thus, the following holds with probability one: for all $\varepsilon > 0$, there exists $\rho_0 > 0$ such that

$$\forall \rho < \rho_0, \forall A' \in B_{\mathcal{A}}(A, \rho), \quad |Y_A - Y_{A'}| \leq \rho^{\alpha+q'/\beta-\varepsilon} + \kappa_f \rho^{q/\beta-\varepsilon}.$$

The result follows immediately. \square

Just as for the upper bound and Proposition 3.5.6, the lower bound of Theorem 3.5.2 is deduced from Proposition 3.5.11 by taking a relevant subsequence in $\underline{R}_f(A)$.

3.5.3 $d_{\mathcal{T}}$ -localized regularity

The Gaussian part

Using the same method as in [42] and under the same entropic conditions, one is able to determine the regularity of Y in the case where $\nu = 0$. Namely, for all $A \in \mathcal{A}$, the following holds with probability one:

$$\alpha_{Y, d_{\mathcal{T}}}(A) = \frac{1}{2} \alpha_{\sigma \int f^2 dm, d_{\mathcal{T}}}(A). \quad (3.5.28)$$

The Poissonian part

Similarly to Section 3.5.2, we introduce for all $A = A(t) \in \mathcal{A}$, $\alpha \geq 0$ and $\rho > 0$,

$$\begin{aligned} L'_{f, \alpha}(A) &= \{s \in \mathcal{T} : |f(s)| > d_{\mathcal{T}}(s, A)^\alpha\}, \\ L'_{f, \alpha}(A, \rho) &= B_{\mathcal{T}}(t, \rho) \cap L'_{f, \alpha}(A), \\ L'^{\complement}_{f, \alpha}(A, \rho) &= B_{\mathcal{T}}(t, \rho) \setminus L'_{f, \alpha}(A) \end{aligned} \quad (3.5.29)$$

and

$$\begin{aligned} \bar{R}'_f(A) &= \left\{ (\alpha, q) \in \mathbf{R}_+^2 : \liminf_{\rho \rightarrow 0^+} \frac{m(L'_{f, \alpha}(A, \rho))}{\rho^q} > 0 \right\}, \\ \underline{R}'_f(A) &= \left\{ (\alpha, q, q') \in \mathbf{R}_+^3 : \limsup_{\rho \rightarrow 0^+} \left[\frac{m(L'_{f, \alpha}(A, \rho))}{\rho^q} + \frac{m(L'^{\complement}_{f, \alpha}(A, \rho))}{\rho^{q'}} \right] < \infty \right\}. \end{aligned} \quad (3.5.30)$$

THEOREM 3.5.12. *Let $A \in \mathcal{A}$. Suppose that $\sigma^2 = 0$ and that f is bounded in the neighborhood of A . Then, with probability one,*

$$\sup_{(\alpha, q, q') \in \underline{R}'_f(A)} \min \left\{ \frac{q}{\beta}, \frac{q'}{\beta} + \alpha \right\} \leq \alpha_{Y, d_{\mathcal{T}}}(A) \leq \inf_{(\alpha, q) \in \bar{R}'_f(A)} \left\{ \frac{q}{\beta} + \alpha \right\}$$

with the conventions that $\inf \emptyset = 1/0 = +\infty$.

Such a result should be compared to Theorem 3.5.2. It constitutes an adequate counterpart to the fact that the former must take non-local information into account. Here, we clearly see that only properties and behaviors around t are considered relevant.

Sketch of proof. Such a result is similar to Theorem 3.5.2 and its proof may be done using similar ideas. We will only focus on highlighting the few differences that arise when applying the same method.

For the upper bound, the key is to prove an estimate using jumps similar to the one in Theorem 3.4.9. It so turns out that the upper bound coming from this approach is the left-hand side of (3.4.12), which should not come as a surprise since the $d_{\mathcal{T}}$ -localized exponent only takes into account what happens in the neighborhood of t . The rest of the computation is the same once one has replaced $\mathcal{V}(A, \rho)$ by $B_{\mathcal{T}}(t, \rho)$.

As for the lower bound, it is somewhat more involved. For all $A = A(t) \in \mathcal{A}$ and $\rho > 0$, we introduce a localized version $Y_{(\cdot, \rho)} = \{Y_{(A', \rho)} : A' \in \mathcal{A}\}$ of Y around A as follows:

$$\forall A' \in \mathcal{A}, \quad Y_{(A', \rho)} = \Delta Y_{A' \cap B_{\mathcal{T}}(t, \rho)}.$$

The set-indexed processes $Y_{(\cdot, \rho)}$ still have independent increments, and so the martingale arguments developed above will still apply. Let us show that.

For the case where $\beta \geq 1$, the key argument lies in Proposition 3.5.8. We claim that, up to some inconsequential constants, we can replace the probability in (3.5.17) by

$$\mathbf{P} \left(\sup_{\substack{C \in \mathcal{C}^\ell \cap B_{\mathcal{V}}(A, 2^{-j}) \\ C \subseteq B_{\mathcal{T}}(t, 2^{-j})}} |\Delta Y_C| > 2^{-j\eta} \right)$$

and the rest of the proof would still follow once one replaces $\mathcal{V}(A, 2^{-j})$ by $B_{\mathcal{T}}(t, 2^{-j})$. Indeed, for all $\rho > 0$,

$$\begin{aligned} \sup_{\substack{C \in \mathcal{C}^\ell \cap B_{\mathcal{V}}(A, \rho) \\ C \subseteq B_{\mathcal{T}}(t, \rho)}} |\Delta Y_C| &\leq \sup_{C \in \mathcal{C}^\ell \cap B_{\mathcal{V}}(A, \rho)} |\Delta Y_{(C, \rho)}| && \text{where } \Delta Y_{(\cdot, \rho)} \text{ is the increment map of } Y_{(\cdot, \rho)}, \\ &\leq \sup_{C \in \mathcal{C}_{(p)}^\ell \cap B_{\mathcal{V}}(A, \rho)} |\Delta Y_{(C, \rho)}| && \text{since } \mathcal{C}^\ell \subseteq \mathcal{C}_{(p)} \text{ where } p = \dim \mathcal{A}, \\ &\leq 2^p \sup_{A_0, \dots, A_p \in B_{\mathcal{A}}(A, \rho)} |Y_{(A, \rho)} - Y_{(A_0 \cap \dots \cap A_p, \rho)}| && \text{by the inclusion-exclusion formula (1.2.7)}. \end{aligned}$$

Moreover, for all $A_0, \dots, A_p \in B_{\mathcal{A}}(A, \rho)$,

$$\begin{aligned} d_{\mathcal{A}}(A, A_0 \cap \dots \cap A_p) &\leq d_{\mathcal{A}}(A, A_0 \cap A) + d_{\mathcal{A}}(A_0 \cap A, A_0 \cap \dots \cap A_p) \\ &\leq d_{\mathcal{A}}(A, A_0) + d_{\mathcal{A}}(A, A_1 \cap \dots \cap A_p) && \text{by contractivity,} \\ &< \rho + d_{\mathcal{A}}(A, A_1 \cap \dots \cap A_p). \end{aligned}$$

Thus, by induction, we deduce that

$$\sup_{\substack{C \in \mathcal{C}^\ell \cap B_{\mathcal{V}}(A, \rho) \\ C \subseteq B_{\mathcal{T}}(t, \rho)}} |\Delta Y_C| \leq 2^p \sup_{A' \in B_{\mathcal{A}}(A, (p+1)\rho)} |Y_{(A, \rho)} - Y_{(A', \rho)}|.$$

Hence for all $j \in \mathbf{N}$,

$$\mathbf{P} \left(\sup_{\substack{C \in \mathcal{C}^\ell \cap B_{\mathcal{V}}(A, 2^{-j}) \\ C \subseteq B_{\mathcal{T}}(t, 2^{-j})}} |\Delta Y_C| > 2^{-j\eta} \right) \leq \mathbf{P} \left(\sup_{A' \in B_{\mathcal{A}}(A, (p+1)2^{-j})} |Y_{(A, 2^{-j})} - Y_{(A', 2^{-j})}| > 2^{-p} 2^{-j\eta} \right)$$

which proves our claim, since Corollary 2.3.31 still applies to $Y_{(\cdot, 2^{-j})}$ and $\Delta Y_{(\mathcal{V}(A, 2^{-j}), 2^{-j})} = \Delta Y_{B_{\mathcal{T}}(t, 2^{-j})}$.

For the case when $\beta < 1$, the trick of introducing the localized process $Y_{(\cdot, \rho)}$ works in a similar fashion. \square

3.5.4 Examples and applications

In the following, we give some simple criteria when applying Theorems 3.5.2 and 3.5.12. We also give an example showing that the inequalities are not always sharp.

Cases of equality

We will see that the local geometry of the vicinity (resp. open ball) around $A \in \mathcal{A}$ plays a crucial role in order to determine the Hölder exponent (resp. the $d_{\mathcal{T}}$ -localized exponent) of Y at A .

COROLLARY 3.5.13. *Let $A \in \mathcal{A}$. Suppose that the following hypotheses hold:*

(i) $\sigma^2 = 0$.

(ii) *There exists $q_{\mathcal{V}} > 0$ such that for all $\varepsilon > 0$, there exists $\rho_{\mathcal{V},\varepsilon} > 0$ such that:*

$$\forall \rho \in (0, \rho_{\mathcal{V},\varepsilon}), \quad \rho^{q_{\mathcal{V}}+\varepsilon} \leq \mathbf{m}(\mathcal{V}(A, \rho)) \leq \rho^{q_{\mathcal{V}}-\varepsilon}. \quad (3.5.31)$$

(iii) *There exists $\alpha \geq 0$ such that for all $\varepsilon > 0$, there exists $\rho_{\alpha,\varepsilon} > 0$ such that:*

$$\forall s \in \mathcal{V}(A, \rho_{\alpha,\varepsilon}), \quad \mathfrak{I}(s, A)^{\alpha+\varepsilon} \leq |f(s)| \leq \mathfrak{I}(s, A)^{\alpha-\varepsilon}. \quad (3.5.32)$$

Then, with probability one,

$$\alpha_X(A) = \frac{q_{\mathcal{V}}}{\beta} \quad \text{and} \quad \alpha_Y(A) = \frac{q_{\mathcal{V}}}{\beta} + \alpha.$$

Proof. The case of X is a particular case of primitive process when $f = 1$, for which the estimate (3.5.32) works with $\alpha = 0$. Thus we only need to prove the result for Y .

Using (3.5.32), we get

$$\forall \varepsilon > 0, \forall \rho \in (0, \rho_{\alpha,\varepsilon}), \quad L_{f,\alpha+\varepsilon}(\rho) = \mathcal{V}(A, \rho) \quad \text{and} \quad L_{f,\alpha-\varepsilon/2}(\rho) = \emptyset.$$

Hence, according to (3.5.31) and Theorem 3.5.2, for all $\varepsilon > 0$, the following holds with probability one:

$$\min \left\{ \frac{\varepsilon^{-1}}{\beta}, \frac{q_{\mathcal{V}} - \varepsilon}{\beta} + \alpha - \frac{\varepsilon}{2} \right\} \leq \alpha_Y(A) \leq \frac{q_{\mathcal{V}} + \varepsilon}{\beta} + \alpha + \varepsilon.$$

The result follows from taking $\varepsilon \rightarrow 0^+$ along a subsequence. \square

Likewise, Theorem 3.5.12 yields the following result. The proof is exactly the same.

COROLLARY 3.5.14. *Let $A = A(t) \in \mathcal{A}$. Suppose that the following hypotheses hold:*

(i) $\sigma^2 = 0$.

(ii) *There exists $q_B > 0$ such that for all $\varepsilon > 0$, there exists $\rho_{\mathcal{V},\varepsilon} > 0$ such that:*

$$\forall \rho \in (0, \rho_{\mathcal{V},\varepsilon}), \quad \rho^{q_B+\varepsilon} \leq \mathbf{m}(B_{\mathcal{T}}(t, \rho)) \leq \rho^{q_B-\varepsilon}. \quad (3.5.33)$$

(iii) *There exists $\alpha \geq 0$ such that for all $\varepsilon > 0$, there exists $\rho_{\alpha,\varepsilon} > 0$ such that:*

$$\forall s \in B_{\mathcal{T}}(t, \rho_{\alpha,\varepsilon}), \quad d_{\mathcal{T}}(s, t)^{\alpha+\varepsilon} \leq |f(s)| \leq d_{\mathcal{T}}(s, t)^{\alpha-\varepsilon}. \quad (3.5.34)$$

Then, with probability one,

$$\alpha_{X,d_{\mathcal{T}}}(A) = \frac{q_B}{\beta} \quad \text{and} \quad \alpha_{Y,d_{\mathcal{T}}}(A) = \frac{q_B}{\beta} + \alpha.$$

The condition on f for Corollary 3.5.14 is actually equivalent to say that $f(t) = 0$ and $\alpha_f(t) = \underline{\alpha}_f(t) = \alpha$ where $\alpha_f(t)$ is the pointwise Hölder exponent as given in (3.4.1) and $\underline{\alpha}_f(t)$ is the *pointwise Hölder subexponent* given by

$$\underline{\alpha}_f(t) = \inf \left\{ \alpha \geq 0 : \liminf_{\rho \rightarrow 0^+} \inf_{s \in B_{\mathcal{T}}(t, \rho)} \frac{|f(s) - f(t)|}{d_{\mathcal{T}}(s, t)^\alpha} > 0 \right\}$$

whenever it is defined. A slightly modified exponent of this kind has already been introduced in [37] to study the local Hausdorff dimension of trajectories of Gaussian processes. We also remark one could also express the estimate on the local behavior of the vicinity (or the ball) with an exponent-like vocabulary.

As a nice consequence to the previous corollaries, we recover a fact that has already been observed in [15, 38] and many others, namely that even in the context of a stochastic integral, integrating still regularizes in some sense.

COROLLARY 3.5.15 (Y is more regular than X). *Let $A = A(t) \in \mathcal{A}$. If the estimate (3.5.31) holds, then with probability one,*

$$\alpha_Y(A) \geq \alpha_X(A).$$

Similarly, if the estimate (3.5.33) holds, then with probability one,

$$\alpha_{Y, d_{\mathcal{T}}}(A) \geq \alpha_{X, d_{\mathcal{T}}}(A) + \alpha_f(t) \mathbb{1}_{f(t)=0}.$$

Stating a better result for $\alpha_Y(A)$ similar to $\alpha_{Y, d_{\mathcal{T}}}(A)$ is quite straightforward, but would require introducing another exponent for f considering δ instead of $d_{\mathcal{T}}$. We chose against it since the $d_{\mathcal{T}}$ -localized exponent already illustrates our point.

Proof. Let $A = A(t) \in \mathcal{A}$. We only prove the result for the $d_{\mathcal{T}}$ -localized exponent, the Hölder exponent being easier. If $\nu = 0$, then Proposition 3.5.1 and (3.5.28) immediately yield the result. So we might as well suppose that $\sigma^2 = 0$. According to Theorem 3.5.12, it is enough to prove that with probability one,

$$\sup_{(\alpha, q, q') \in \underline{R}'_f(A)} \min \left\{ \frac{q}{\beta}, \frac{q'}{\beta} + \alpha \right\} \geq \alpha_{X, d_{\mathcal{T}}}(A) + \alpha_f(t) \mathbb{1}_{f(t)=0}. \tag{3.5.35}$$

Suppose that $f(t) \neq 0$ or $\alpha_f(t) = 0$ so that $\alpha_f(t) \mathbb{1}_{f(t)=0} = 0$. Then, remark that according to (3.5.31), for all $\varepsilon > 0$ and $(\alpha, q, q') \in \underline{R}'_f(A)$, we may always consider that both q and q' are greater than $q_B - \varepsilon$ in the left-hand side of (3.5.35). In particular,

$$\sup_{(\alpha, q, q') \in \underline{R}'_f(A)} \min \left\{ \frac{q}{\beta}, \frac{q'}{\beta} + \alpha \right\} \geq \frac{q_B}{\beta}. \tag{3.5.36}$$

According to Corollary 3.5.14, $q_B/\beta = \alpha_{X, d_{\mathcal{T}}}(A)$ almost surely. Hence (3.5.35) follows from (3.5.36) in this case.

Now, suppose that $f(t) = 0$ and $\alpha_f(t) > 0$. Then, for all $\alpha \in (0, \alpha_f(t))$ and small enough $\rho > 0$, $L'_{f, \alpha}(A, \rho) = \emptyset$. In particular, $(\alpha, q, q_B - \varepsilon) \in \underline{R}'_f(A)$ for all $q, \varepsilon > 0$. Thus

$$\sup_{(\alpha, q, q') \in \underline{R}'_f(A)} \min \left\{ \frac{q}{\beta}, \frac{q'}{\beta} + \alpha \right\} \geq \min \left\{ \frac{q}{\beta}, \frac{q_B - \varepsilon}{\beta} + \alpha \right\}.$$

Taking $q \rightarrow +\infty$, $\varepsilon \rightarrow 0^+$ and $\alpha \rightarrow \alpha_f(t)^+$ yields

$$\sup_{(\alpha, q, q') \in \mathbb{R}'_f(A)} \min \left\{ \frac{q}{\beta}, \frac{q'}{\beta} + \alpha \right\} \geq \frac{q_B}{\beta} + \alpha_f(t). \tag{3.5.37}$$

Hence (3.5.35) follows from (3.5.37) in this case. □

We proceed to apply those results to a multiparameter Lévy process in order to show that various behaviors start to appear when $p > 1$. The coming example should be compared to the case $p = 1$ where it was proven in [20] that almost sure regularity is $1/\beta$.

EXAMPLE 3.5.16 (Set-indexed Lévy process for $\mathcal{T} = \mathbf{R}_+^p$). Suppose that $\mathcal{T} = \mathbf{R}_+^p$ is endowed with its product indexing collection \mathcal{A} from Example 1.2.10, \mathbf{m} is the Lebesgue measure and $d_{\mathcal{T}} = \|\cdot\|$ is any distance induced by a norm on \mathbf{R}^p .

Let also X be a purely Poissonian siLévy and $t \in \mathbf{R}_+^p$. According to Corollaries 3.5.13 and 3.5.14 which hold for $q_V = 1$ and $q_B = p$ respectively, the following holds with probability one:

$$\alpha_X(A(t)) = \begin{cases} 1/\beta & \text{if } t \neq 0, \\ p/\beta & \text{if } t = 0 \end{cases} \quad \text{and} \quad \alpha_{X, \|\cdot\|}(A(t)) = p/\beta.$$

When $d_{\mathcal{A}} = d_{\mathbf{m}}$ is taken instead, the values of $\alpha_X(A(t))$ and $\alpha_{X, d_{\mathcal{T}}}(A(t))$ do not change as long as $t = (t_1, \dots, t_p)$ is such that $t_1 \dots t_p \neq 0$ since $d_{\mathcal{T}}$ stays (Lipschitz-)equivalent to $\|\cdot\|$ on any compact set away from the coordinate hyperplanes (see [37, Lemma 3.1]).

Let us now consider t in such an hyperplane. If $p = 1$, there is nothing much to say and we recover the result of [20], i.e. $\alpha_X(0) = 1/\beta$ almost surely. However, when $p > 1$, we have $\mathbf{m}(B_{\mathcal{T}}(t, \rho)) = \infty$ for all $\rho > 0$, so an argument based on Borel-Cantelli ensures that the following event holds with probability one:

$$\Omega^* = \bigcup_{\varepsilon > 0} \bigcap_{\rho > 0} \{ \exists s \in \Pi \cap B_{\mathcal{T}}(t, \rho) : |J_s(X)| \geq \varepsilon \}.$$

This means that there are sequences of macroscopic jumps converging to 0. Applying the estimate (3.4.12) yields $\alpha_X(t) = \alpha_{X, d_{\mathcal{T}}}(t) = 0$ almost surely.

For related results in the multiparameter setting, we refer to [29].

The one-dimensional case $\mathcal{T} = \mathbf{R}_+$

When $\mathcal{T} = \mathbf{R}_+$, both exponents are reduced to the usual pointwise one, yielding the following result.

COROLLARY 3.5.17 (Hölder regularity when $\mathcal{T} = \mathbf{R}_+$). Suppose that \mathbf{m} is the usual Lebesgue measure on \mathbf{R}_+ and that $\sigma^2 = 0$. Then, for all $t \in \mathbf{R}_+$, the following holds with probability one:

$$\sup_{(\alpha, q, q') \in \mathbb{R}'_f(t)} \min \left\{ \frac{q}{\beta}, \frac{q'}{\beta} + \alpha \right\} \leq \alpha_Y(t) \leq \inf_{(\alpha, q) \in \mathbb{R}'_f(t)} \left\{ \frac{q}{\beta} + \alpha \right\}. \tag{3.5.38}$$

Moreover, if there exists $\alpha \geq 0$ such that for all $\varepsilon > 0$, there exists $\rho_{\alpha, \varepsilon} > 0$ such that

$$\forall s \in (t - \rho_{\alpha, \varepsilon}, t + \rho_{\alpha, \varepsilon}), \quad |s - t|^{\alpha + \varepsilon} \leq |f(s)| \leq |s - t|^{\alpha - \varepsilon}.$$

then, the following holds with probability one:

$$\alpha_Y(A) = \frac{1}{\beta} + \alpha.$$

Remark that in this case, by a similar argument to Corollary 3.5.15, we may always take $q \wedge q' = 1$ in the left-hand side of (3.5.38) and $q \geq 1$ in its right-hand side, simplifying the expression in practical applications.

EXAMPLES 3.5.18. *We finally address the simplest case and another one where the upper and lower bounds do not coincide.*

◇ Suppose that $f(s) = s^\alpha$ for all $s \in \mathbf{R}_+$. Consider $t \in \mathbf{R}_+$. Using Corollary 3.5.17, we get

$$\alpha_Y(t) = \frac{1}{\beta} + \alpha \mathbb{1}_{t=0} \quad \text{a.s.}$$

◇ Let $q \geq 1$. Consider a Borel set $E \subseteq \mathbf{R}_+$ such that

$$0 < \liminf_{\rho \rightarrow 0^+} \frac{m(E \cap [0, \rho])}{\rho^q} \leq \limsup_{\rho \rightarrow 0^+} \frac{m(E \cap [0, \rho])}{\rho^q} < \infty.$$

For instance, the set

$$E = \bigcup_{j \in \mathbf{N}} (2^{-j} - 2^{-jq}, 2^{-j}]$$

works fine. Then consider $0 \leq \alpha < \alpha'$ and define the function $f(s) = s^\alpha \mathbb{1}_{s \in E} + s^{\alpha'} \mathbb{1}_{s \notin E}$ for all $s \in \mathbf{R}_+$. Using Corollary 3.5.17, it follows that

$$\min \left\{ \frac{q}{\beta}, \frac{1}{\beta} + \alpha \right\} \leq \alpha_Y(0) \leq \min \left\{ \frac{q}{\beta} + \alpha, \frac{1}{\beta} + \alpha' \right\} \quad \text{a.s.}$$

which is not an equality for a large choice of q, α, α' and β .

3.6 Perspectives

3.6.1 Improve on the finite-dimensional assumption

Supposing that the indexing collection \mathcal{A} is of finite dimension has been very helpful to us on multiple occasions. Broadly, we used this hypothesis to ease the study of the regularity of sample paths. However, in quite a few instances here, it was only used through the lens of the entropy (Definition 1.5.11). This approach to regularity for set-indexed processes has been undertaken in [42]. However, as mentioned by Ledoux and Talagrand in [66, Chapter 11, §2], entropy conditions are not suited to spaces that may lack ‘homogeneity’. Due to the possibility of gluing together all kinds of indexing collections (Propositions 1.2.23 and 1.2.26), this issue does concern our setting. The authors then go on to expose the theory of *majorizing measures* spearheaded by Talagrand to solve this issue. This kind of condition is particularly interesting here since we already have a measure m that could verify the majorizing measure condition. In particular, Example 3.2.5 could have been approached from the majorizing measure point of view and the continuity of the process would have followed from [66, Theorem 11.14].

With those considerations in mind, our finite-dimensional assumption could very well be improved upon.

3.6.2 Towards a uniform upper bound for the primitive process

A rightful question would be to know whether or not one may extend Theorems 3.5.2 and 3.5.12 to give an almost sure everywhere upper bound, *i.e.* permute ‘ $\forall A$ ’ with ‘a.s.’ in the statement for the upper bound. This is not a groundless speculation since such a result holds for \mathbf{R}_+ -indexed Lévy processes by [51, Theorem 1].

For the siLévy, the method used to prove the aforementioned theorems may still be adapted, but with two caveats:

- ◊ The heart of the argument is to establish that the random set $E_{|\mathcal{T}}^\delta(X)$ defined in (3.4.22) covers \mathcal{T} almost surely. We may do so by adapting Hoffmann-Jørgensen’s [44, Theorem 2] which is a generalization of Shepp’s celebrated covering theorems [90, 91] to metric spaces (we refer to [44] for related results). The ‘cost to pay’ would be to add the *weak finite-dimensional assumption* made by Hoffmann-Jørgensen in [44, §4].
- ◊ Once the covering holds, one needs to check if Theorem 3.4.9 still holds a.s. for all $A \in \mathcal{A}$. This is the case if the siLévy is a.s. in generic configuration in the vicinity of all $A \in \mathcal{A}$. This should hold with the additional assumption that

$$\forall t \in \mathcal{T}, \quad \mathbf{m}(\dagger(t)) = 0$$

where

$$\dagger(t) = \bigcap_{\rho > 0} \{s \in \mathcal{T} : \delta(s, A(t)) < \rho \text{ or } \delta(t, A(s)) < \rho\} = \bigcap_{\rho > 0} \mathcal{V}(A(t), \rho) \cup \mathbf{t}(\mathcal{V}'(t, \rho))$$

represents the ‘cross of center t ’. This ‘cross’ plays a similar role to the ‘half cross’ $\mathcal{L}(t)$ introduced in (3.4.15) to prove generic configuration for a siLévy around some fixed $A \in \mathcal{A}$ (Lemma 3.4.8). Basically, having all crosses of null \mathbf{m} -measure ensures that no two jump’s crosses of the siLévy share a common branch so that they can be picked apart with the same method as Lemma 3.4.8.

Alas, the case for the primitive process is harder because the covering set we use is $E_{|L_{f,\alpha}(A)}^\delta$, and thus becomes dependent on A , which makes it impossible for the method to be applied directly.

3.6.3 Regularity of \mathcal{C} -exchangeable processes

Just like one may go from a ‘weak’ Lévy-Itô decomposition (Corollary 2.4.9) to a ‘strong’ one (Theorem 3.3.8), we should be able to use the ‘weak’ representation of \mathcal{C} -exchangeable processes (Theorem 2.5.30) and get a version in $\mathcal{D}^\Phi(\mathcal{A})$ out of it. In the $[0, 1]$ -indexed case, this has already been carried out by Hagberg [34, Theorem 3].

Once this is done, this is the gateway to establish Hölder regularity for \mathcal{C} -exchangeable processes and primitive processes with respect to those in the spirit of Theorems 3.5.2 and 3.5.12.

3.7 Perspectives (français)

3.7.1 Améliorer l'hypothèse de dimension finie

Supposer que l'*indexing collection* \mathcal{A} de dimension finie nous a été très utile à de nombreuses reprises. De manière générale, nous avons employé cette hypothèse pour faciliter l'étude de la régularité des trajectoires. Cependant, dans un certain nombre de cas, elle n'a été utilisée qu'au travers d'un argument d'entropie métrique (Définition 1.5.11). Cette approche de la régularité pour les processus indexés par des ensembles a déjà été entreprise dans [42]. Néanmoins, comme très justement noté par Ledoux et Talagrand [66, Chapter 11, §2], les conditions entropiques ne sont pas adaptées à des espaces manquant d'une certaine "homogénéité". Comme nous sommes en mesure de recoller ensemble des *indexing collections* de natures très variées (Propositions 1.2.23 et 1.2.26), notre cadre est bel et bien concerné par cette éventualité. Les auteurs continuent alors et introduisent la théorie des mesures majorantes, initiée par Talagrand, pour résoudre ce problème. Ce type de propriété est particulièrement intéressant dans notre cadre car nous disposons déjà d'une mesure m qui pourrait très bien vérifier la condition d'être une mesure majorante. En particulier, l'Exemple 3.2.5 aurait pu être approché de cette manière et la continuité du processus aurait découlé de [66, Theorem 11.14].

L'ensemble de ces considérations pris en compte, notre hypothèse de dimension finie pourrait fort bien être améliorée.

3.7.2 Vers une borne supérieure uniforme de la régularité de la primitive

Un questionnement légitime serait de savoir si oui ou non les Théorèmes 3.5.2 et 3.5.12 peuvent être améliorés pour donner une borne supérieure partout, *i.e.* si le " $\forall A$ " et le "presque sûrement" peuvent être échangés. Ce n'est pas une spéculation sans fondement car un tel résultat est vrai pour les processus de Lévy indexés par \mathbf{R}_+ grâce à [51, Theorem 1].

Pour un processus de Lévy indexés par des ensembles (siLévy), la méthode employée pour démontrer les théorèmes mentionnés à l'instant peut encore s'adapter, à deux amendements près :

- ◊ Le cœur de l'argument est de montrer que l'ensemble aléatoire $E_{|\mathcal{T}}^\delta(X)$ donné en (3.4.22) recouvre \mathcal{T} presque sûrement. Nous pouvons y arriver à l'aide du [44, Theorem 2] de Hoffmann-Jørgensen qui est une généralisation aux espaces métriques des théorèmes de recouvrement de Shepp [90, 91] (se référer à [44] pour d'autres résultats de la littérature). Le "prix à payer" serait d'ajouter l'hypothèse de *faible dimension finie* [44, §4].
- ◊ Une fois que le recouvrement est assuré, on doit également vérifier que le Théorème 3.4.9 a lieu presque sûrement pour tout $A \in \mathcal{A}$. C'est le cas si le siLévy est presque sûrement en configuration générique autour de $A \in \mathcal{A}$ pour tout $A \in \mathcal{A}$. Cela devrait être vrai sous l'hypothèse additionnelle que

$$\forall t \in \mathcal{T}, \quad m(\dagger(t)) = 0$$

où

$$\dagger(t) = \bigcap_{\rho > 0} \{s \in \mathcal{T} : \mathfrak{d}(s, A(t)) < \rho \text{ or } \mathfrak{d}(t, A(s)) < \rho\} = \bigcap_{\rho > 0} \mathcal{V}(A(t), \rho) \cup \mathfrak{t}(\mathcal{V}'(t, \rho))$$

représente la "croix de centre t ". Cette "croix" joue le même rôle que la "demi-croix" $\mathcal{L}(t)$ introduite en (3.4.15) pour démontrer qu'un siLévy est en configuration générique autour de n'importe quel $A \in \mathcal{A}$ fixé (Lemme 3.4.8). Intuitivement, avoir toutes ces croix de m -mesure

nulle assure que deux croix dont les centres sont des sauts du siLévy ne partagent aucune branche en commun, permettant de séparer les sauts avec la même méthode que celle du Lemme 3.4.8.

Malheureusement, le cas de la primitive d'une fonction contre un siLévy est plus dur car l'ensemble aléatoire utilisé pour le recouvrement est $E_{|L_f, \alpha(A)}^\delta$, qui dépend de A , qui rend une adaptation directe de la méthode précédente caduque.

3.7.3 Régularité des processus \mathcal{C} -échangeables

De la même manière qu'on passe d'une représentation de Lévy-Itô "faible" (Corollaire 2.4.9) à une "forte" (Théorème 3.3.8), nous devrions être capable d'utiliser la représentation "faible" des processus \mathcal{C} -échangeables (Théorème 2.5.30) et d'obtenir une version "forte" à trajectoires dans $\mathcal{D}^\Phi(\mathcal{A})$. Dans le cas où $\mathcal{T} = [0, 1]$, cela a déjà été démontré par Hagberg [34, Theorem 3].

Cette étape constituerait le premier pas vers une étude de la régularité ponctuelle des processus \mathcal{C} -échangeables et des primitives définies par rapport à eux. Nous imaginons bien que des analogues aux Théorèmes 3.5.2 et 3.5.12 doivent être vrais.

Bibliography

- [1] Robert J. Adler, *An introduction to continuity, extrema, and related topics for general Gaussian processes*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, vol. 12, Institute of Mathematical Statistics, Hayward, CA, 1990. MR 1088478
- [2] Robert J. Adler, Sunder Ram Krishnan, Jonathan E. Taylor, and Shmuel Weinberger, *Convergence of the reach for a sequence of Gaussian-embedded manifolds*, Probab. Theory Related Fields **171** (2018), no. 3-4, 1045–1091. MR 3827227
- [3] Robert J. Adler, Ditlev Monrad, Richard H. Scissors, and Richard Wilson, *Representations, decompositions and sample function continuity of random fields with independent increments*, Stochastic Process. Appl. **15** (1983), no. 1, 3–30. MR 694534
- [4] Robert J. Adler and Jonathan E. Taylor, *Random fields and geometry*, Springer Monographs in Mathematics, Springer, New York, 2007. MR 2319516
- [5] David J. Aldous, *Representations for partially exchangeable arrays of random variables*, J. Multivariate Anal. **11** (1981), no. 4, 581–598. MR 637937
- [6] ———, *Exchangeability and related topics*, École d’été de probabilités de Saint-Flour, XIII—1983, Lecture Notes in Math., vol. 1117, Springer, Berlin, 1985, pp. 1–198. MR 883646
- [7] ———, *The continuum random tree. I*, Ann. Probab. **19** (1991), no. 1, 1–28. MR 1085326
- [8] ———, *The continuum random tree. III*, Ann. Probab. **21** (1993), no. 1, 248–289. MR 1207226
- [9] Charalambos D. Aliprantis and Kim C. Border, *Infinite dimensional analysis*, third ed., Springer, Berlin, 2006, A hitchhiker’s guide. MR 2378491
- [10] David Applebaum, *Lévy processes and stochastic calculus*, second ed., Cambridge Studies in Advanced Mathematics, vol. 116, Cambridge University Press, Cambridge, 2009. MR 2512800
- [11] Tim Austin and Dmitry Panchenko, *A hierarchical version of the de Finetti and Aldous-Hoover representations*, Probab. Theory Related Fields **159** (2014), no. 3-4, 809–823. MR 3230009
- [12] Antoine Ayache and Murad S. Taqqu, *Multifractional processes with random exponent*, Publ. Mat. **49** (2005), no. 2, 459–486. MR 2177638

- [13] Paul Balança, *Fine regularity of Lévy processes and linear (multi)fractional stable motion*, Electron. J. Probab. **19** (2014), no. 101, 37. MR 3275853
- [14] ———, *Some sample path properties of multifractional Brownian motion*, Stochastic Process. Appl. **125** (2015), no. 10, 3823–3850. MR 3373305
- [15] Paul Balança and Erick Herbin, *2-microlocal analysis of martingales and stochastic integrals*, Stochastic Process. Appl. **122** (2012), no. 6, 2346–2382. MR 2922632
- [16] Richard F. Bass and Ronald Pyke, *The existence of set-indexed Lévy processes*, Z. Wahrsch. Verw. Gebiete **66** (1984), no. 2, 157–172. MR 749219
- [17] ———, *The space $D(A)$ and weak convergence for set-indexed processes*, Ann. Probab. **13** (1985), no. 3, 860–884. MR 799425
- [18] Patrick Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1999, A Wiley-Interscience Publication. MR 1700749
- [19] Garrett Birkhoff, *Lattice theory*, Third edition. American Mathematical Society Colloquium Publications, Vol. XXV, American Mathematical Society, Providence, R.I., 1967. MR 0227053
- [20] Robert M. Blumenthal and Ronald K. Gettoor, *Sample functions of stochastic processes with stationary independent increments*, J. Math. Mech. **10** (1961), 493–516. MR 0123362
- [21] Jean-Michel Bony, *Second microlocalization and propagation of singularities for semilinear hyperbolic equations*, Hyperbolic equations and related topics (Katata/Kyoto, 1984), Academic Press, Boston, MA, 1986, pp. 11–49. MR 925240
- [22] Nicolas Bourbaki, *Éléments de mathématique. Topologie générale. Chapitres 5 à 10*, Hermann, Paris, 1974. MR 3822133
- [23] Renzo Cairoli and John B. Walsh, *Stochastic integrals in the plane*, Acta Math. **134** (1975), 111–183. MR 420845
- [24] Harry Crane and Henry Towsner, *Relative exchangeability with equivalence relations*, Arch. Math. Logic **57** (2018), no. 5-6, 533–556. MR 3828877
- [25] ———, *Relatively exchangeable structures*, J. Symb. Log. **83** (2018), no. 2, 416–442. MR 3835071
- [26] Nicolas Curien and Bénédicte Haas, *Random trees constructed by aggregation*, Ann. Inst. Fourier (Grenoble) **67** (2017), no. 5, 1963–2001. MR 3732681
- [27] Joseph L. Doob, *Stochastic processes*, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1990, Reprint of the 1953 original, A Wiley-Interscience Publication. MR 1038526
- [28] Richard M. Dudley, *Sample functions of the Gaussian process*, Ann. Probability **1** (1973), no. 1, 66–103. MR 346884
- [29] Arnaud Durand and Stéphane Jaffard, *Multifractal analysis of Lévy fields*, Probab. Theory Related Fields **153** (2012), no. 1-2, 45–96. MR 2925570

- [30] Rick Durrett, *Probability: theory and examples*, fourth ed., Cambridge Series in Statistical and Probabilistic Mathematics, vol. 31, Cambridge University Press, Cambridge, 2010. MR 2722836
- [31] Steven N. Evans, *Probability and real trees*, Lecture Notes in Mathematics, vol. 1920, Springer, Berlin, 2008, Lectures from the 35th Summer School on Probability Theory held in Saint-Flour, July 6–23, 2005. MR 2351587
- [32] Zoltan Füredi, Peter Hajnal, Vojtech Rödl, and William T. Trotter, *Interval orders and shift graphs*, Sets, graphs and numbers (Budapest, 1991), Colloq. Math. Soc. János Bolyai, vol. 60, North-Holland, Amsterdam, 1992, pp. 297–313. MR 1218198
- [33] Bénédicte Haas, *Asymptotics of heights in random trees constructed by aggregation*, Electron. J. Probab. **22** (2017), Paper No. 21, 25. MR 3622891
- [34] Jan Hagberg, *Approximation of the summation process obtained by sampling from a finite population*, Teor. Verоятnost. i Primenen. **18** (1973), 790–803. MR 0328985
- [35] Paul R. Halmos, *Measure Theory*, D. Van Nostrand Company, Inc., New York, N. Y., 1950. MR 0033869
- [36] Petteri Harjulehto and Peter Hästö, *Orlicz spaces and generalized Orlicz spaces*, Lecture Notes in Mathematics, vol. 2236, Springer, Cham, 2019. MR 3931352
- [37] Erick Herbin, Benjamin Arras, and Geoffroy Barruel, *From almost sure local regularity to almost sure Hausdorff dimension for Gaussian fields*, ESAIM Probab. Stat. **18** (2014), 418–440. MR 3333997
- [38] Erick Herbin and Jacques Lévy-Véhel, *Stochastic 2-microlocal analysis*, Stochastic Process. Appl. **119** (2009), no. 7, 2277–2311. MR 2531092
- [39] Erick Herbin and Ely Merzbach, *A set-indexed fractional Brownian motion*, J. Theoret. Probab. **19** (2006), no. 2, 337–364. MR 2283380
- [40] ———, *Stationarity and self-similarity characterization of the set-indexed fractional Brownian motion*, J. Theoret. Probab. **22** (2009), no. 4, 1010–1029. MR 2558663
- [41] ———, *The set-indexed Lévy process: stationarity, Markov and sample paths properties*, Stochastic Process. Appl. **123** (2013), no. 5, 1638–1670. MR 3027894
- [42] Erick Herbin and Alexandre Richard, *Local Hölder regularity for set-indexed processes*, Israel J. Math. **215** (2016), no. 1, 397–440. MR 3551904
- [43] Chris Heunen, Ohad Kammar, Sam Staton, and Hongseok Yang, *A convenient category for higher-order probability theory*, 2017 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), IEEE, [Piscataway], NJ, 2017, p. 12. MR 3776963
- [44] Jørgen Hoffmann-Jørgensen, *Coverings of metric spaces with randomly placed balls*, Math. Scand. **32** (1973), 169–186 (1974). MR 341556
- [45] Douglas N. Hoover, *Relations on probability spaces and arrays of random variables*, Preprint, Institute for Advanced Study, Princeton (1979).

- [46] B. Gail Ivanoff, *Set-indexed processes: distributions and weak convergence*, Topics in spatial stochastic processes (Martina Franca, 2001), Lecture Notes in Math., vol. 1802, Springer, Berlin, 2003, pp. 85–125. MR 1975518
- [47] B. Gail Ivanoff and Ely Merzbach, *Set-indexed martingales*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability (1999), Taylor & Francis.
- [48] B. Gail Ivanoff, Ely Merzbach, and Mathieu Plante, *A compensator characterization of point processes on topological lattices*, Electron. J. Probab. **12** (2007), no. 2, 47–74. MR 2280258
- [49] Stéphane Jaffard, *Pointwise smoothness, two-microlocalization and wavelet coefficients*, vol. 35, 1991, Conference on Mathematical Analysis (El Escorial, 1989), pp. 155–168. MR 1103613
- [50] ———, *Old friends revisited: the multifractal nature of some classical functions*, J. Fourier Anal. Appl. **3** (1997), no. 1, 1–22. MR 1428813
- [51] ———, *The multifractal nature of Lévy processes*, Probab. Theory Related Fields **114** (1999), no. 2, 207–227. MR 1701520
- [52] Paul Jung, Jiho Lee, Sam Staton, and Hongseok Yang, *A generalization of hierarchical exchangeability on trees to directed acyclic graphs*, Accepted for Ann. H. Lebesgue **4** (2020).
- [53] Olav Kallenberg, *Canonical representations and convergence criteria for processes with interchangeable increments*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **27** (1973), 23–36. MR 394842
- [54] ———, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR 1876169
- [55] ———, *Probabilistic symmetries and invariance principles*, Probability and its Applications (New York), Springer, New York, 2005. MR 2161313
- [56] Davar Khoshnevisan, *Multiparameter processes*, Springer Monographs in Mathematics, Springer-Verlag, New York, 2002, An introduction to random fields. MR 1914748
- [57] John F. C. Kingman, *Uses of exchangeability*, Ann. Probability **6** (1978), no. 2, 183–197. MR 494344
- [58] ———, *Poisson processes*, Oxford Studies in Probability, vol. 3, The Clarendon Press, Oxford University Press, New York, 1993, Oxford Science Publications. MR 1207584
- [59] Achim Klenke, *Probability theory*, second ed., Universitext, Springer, London, 2014, A comprehensive course. MR 3112259
- [60] Andrei N. Kolmogorov, *Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum*, C. R. (Doklady) Acad. Sci. URSS (N.S.) **26** (1940), 115–118. MR 0003441
- [61] Sunder Ram Krishnan, Jonathan E. Taylor, and Robert J. Adler, *The intrinsic geometry of some random manifolds*, Electron. Commun. Probab. **22** (2017), Paper No. 1, 12. MR 3607796

- [62] Stanisław Kwapien and Wojbor A. Woyczyński, *Random series and stochastic integrals: single and multiple*, Probability and its Applications, Birkhäuser Boston, Inc., Boston, MA, 1992. MR 1167198
- [63] Serge Lang, *Real and functional analysis*, third ed., Graduate Texts in Mathematics, vol. 142, Springer-Verlag, New York, 1993. MR 1216137
- [64] Günter Last and Mathew Penrose, *Lectures on the Poisson process*, Institute of Mathematical Statistics Textbooks, vol. 7, Cambridge University Press, Cambridge, 2018. MR 3791470
- [65] Jean-François Le Gall, *Random trees and applications*, Probab. Surv. **2** (2005), 245–311. MR 2203728
- [66] Michel Ledoux and Michel Talagrand, *Probability in Banach spaces*, Classics in Mathematics, Springer-Verlag, Berlin, 2011, Isoperimetry and processes, Reprint of the 1991 edition. MR 2814399
- [67] Peter M. Lee, *Infinitely divisible stochastic processes*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **7** (1967), 147–160. MR 208668
- [68] Jacques Lévy Véhel and Stéphane Seuret, *The 2-microlocal formalism*, Fractal geometry and applications: a jubilee of Benoît Mandelbrot, Part 2, Proc. Sympos. Pure Math., vol. 72, Amer. Math. Soc., Providence, RI, 2004, pp. 153–215. MR 2112123
- [69] John E. Littlewood, *Lectures on the Theory of Functions*, Oxford University Press, 1944. MR 0012121
- [70] Gisiro Maruyama, *Infinitely divisible processes*, Teor. Verojatnost. i Primenen. **15** (1970), 3–23. MR 0285046
- [71] Ely Merzbach, *An introduction to the general theory of set-indexed martingales*, Topics in spatial stochastic processes (Martina Franca, 2001), Lecture Notes in Math., vol. 1802, Springer, Berlin, 2003, pp. 41–84. MR 1975517
- [72] Ely Merzbach and Diane Saada, *Stochastic integration for set-indexed processes*, Israel J. Math. **118** (2000), 157–181. MR 1776081
- [73] Ely Merzbach and Arthur Yosef, *Set-indexed Brownian motion on increasing paths*, J. Theoret. Probab. **22** (2009), no. 4, 883–890. MR 2558657
- [74] Yves Meyer, *Wavelets, vibrations and scalings*, CRM Monograph Series, vol. 9, American Mathematical Society, Providence, RI, 1998, With a preface in French by the author. MR 1483896
- [75] Ilya Molchanov, *Theory of random sets*, Probability and its Applications (New York), Springer-Verlag London, Ltd., London, 2005. MR 2132405
- [76] Julian Musielak, *Orlicz spaces and modular spaces*, Lecture Notes in Mathematics, vol. 1034, Springer-Verlag, Berlin, 1983. MR 724434
- [77] Jacques Neveu, *Processus ponctuels*, École d'Été de Probabilités de Saint-Flour, VI—1976, 1977, pp. 249–445. Lecture Notes in Math., Vol. 598. MR 0474493
- [78] ———, *Arbres et processus de Galton-Watson*, Ann. Inst. H. Poincaré Probab. Statist. **22** (1986), no. 2, 199–207. MR 850756

- [79] Oystein Ore, *Theory of graphs*, Third printing, with corrections. American Mathematical Society Colloquium Publications, Vol. XXXVIII, American Mathematical Society, Providence, R.I., 1967. MR 0244094
- [80] Gilles Pisier, *Conditions d'entropie assurant la continuité de certains processus et applications à l'analyse harmonique*, Seminar on Functional Analysis, 1979–1980 (French), École Polytech., Palaiseau, 1980, pp. Exp. No. 13–14, 43. MR 604395
- [81] Henri Poincaré, *Science et méthode*, Ernest Flammarion, Paris, 1920, Bibliothèque de Philosophie scientifique.
- [82] William E. Pruitt, *The growth of random walks and Lévy processes*, Ann. Probab. **9** (1981), no. 6, 948–956. MR 632968
- [83] Balram S. Rajput and Jan Rosiński, *Spectral representations of infinitely divisible processes*, Probab. Theory Related Fields **82** (1989), no. 3, 451–487. MR 1001524
- [84] Jan Rosiński, *On path properties of certain infinitely divisible processes*, Stochastic Process. Appl. **33** (1989), no. 1, 73–87. MR 1027109
- [85] ———, *Representations and isomorphism identities for infinitely divisible processes*, Ann. Probab. **46** (2018), no. 6, 3229–3274. MR 3857855
- [86] Diane Saada and Dean Slonowsky, *The set-indexed Ito integral*, J. Anal. Math. **94** (2004), 61–89. MR 2124455
- [87] Ken-iti Sato, *Lévy processes and infinitely divisible distributions*, Cambridge Studies in Advanced Mathematics, vol. 68, Cambridge University Press, Cambridge, 1999.
- [88] ———, *Fractional integrals and extensions of selfdecomposability*, Lévy matters I, Lecture Notes in Math., vol. 2001, Springer, Berlin, 2010, pp. 1–91, 197–198. MR 2731896
- [89] Delphin Sénizergues, *Random gluing of metric spaces*, Ann. Probab. **47** (2019), no. 6, 3812–3865. MR 4038043
- [90] Lawrence A. Shepp, *Covering the circle with random arcs*, Israel J. Math. **11** (1972), 328–345. MR 295402
- [91] ———, *Covering the line with random intervals*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **23** (1972), 163–170. MR 322923
- [92] Anatoli V. Skorokhod, *Limit theorems for stochastic processes*, Teor. Veroyatnost. i Primenen. **1** (1956), 289–319. MR 0084897
- [93] Dean Slonowsky, *Central limit theorems for set-indexed strong martingales*, ProQuest LLC, Ann Arbor, MI, 1998, Thesis (Ph.D.)—University of Ottawa (Canada). MR 2698879
- [94] Joel H. Spencer, *Minimal scrambling sets of simple orders*, Acta Math. Acad. Sci. Hungar. **22** (1971/72), 349–353. MR 292722
- [95] Stilian A. Stoev and Murad S. Taqqu, *How rich is the class of multifractional Brownian motions?*, Stochastic Process. Appl. **116** (2006), no. 2, 200–221. MR 2197974

-
- [96] Miron L. Straf, *Weak convergence of stochastic processes with several parameters*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory, 1972, pp. 187–221. MR 0402847
- [97] Jacques Tits, *Sur le groupe des automorphismes d'un arbre*, Essays on topology and related topics (Mémoires dédiés à Georges de Rham), 1970, pp. 188–211. MR 0299534
- [98] William T. Trotter, *Applications of the probabilistic method to partially ordered sets*, The mathematics of Paul Erdős, II, Algorithms Combin., vol. 14, Springer, Berlin, 1997, pp. 214–228. MR 1425215
- [99] Kazimierz Urbanik and Wojbor A. Woyczyński, *A random integral and Orlicz spaces*, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **15** (1967), 161–169. MR 215329
- [100] Aad W. van der Vaart and Jon A. Wellner, *Weak convergence and empirical processes*, Springer Series in Statistics, Springer-Verlag, New York, 1996, With applications to statistics. MR 1385671
- [101] Michael J. Wichura, *Inequalities with applications to the weak convergence of random processes with multi-dimensional time parameters*, Ann. Math. Statist. **40** (1969), 681–687. MR 246359
- [102] Yimin Xiao, *Uniform modulus of continuity of random fields*, Monatsh. Math. **159** (2010), no. 1-2, 163–184. MR 2564392
- [103] Dhandapani Yogeshwaran and Robert J. Adler, *On the topology of random complexes built over stationary point processes*, Ann. Appl. Probab. **25** (2015), no. 6, 3338–3380. MR 3404638

Title: Regularity of generalized stochastic processes

Keywords: generalized processes, set-indexed processes, stationarity, sample path properties, Hölder regularity

Abstract: More and less recent studies pinpoint a need for the probabilistic community to understand processes indexed by spaces that are more general than \mathbf{N} or \mathbf{R}_+ . This thesis focuses on processes $\mathbf{X} = \{X_t : t \in \mathcal{T}\}$ indexed by a very general set \mathcal{T} endowed with an order relation that represents a kind of time flow. Classes of manifolds and continuous trees of interest are amongst the great variety of examples, without forgetting about more algebraic-flavored ones.

The order structure allows to seamlessly identify each process $\mathbf{X} = \{X_t : t \in \mathcal{T}\}$ to a process $X = \{X_A : A \in \mathcal{A}\}$ indexed by a collection of sets, making a bridge with the set-indexed theory developed by Ivanoff and Merzbach. Under some assumptions, the latter may be extended to a stochastic measure, leading to the construction of a linear map corresponding to the integral with respect to X . The case when X has independent increments is well understood since the work of Rajput and Rosiński. However, the case when the increments are stationary or exchangeable has been mainly limited to \mathbf{R}_+ so far. New notions of stationarity fit to this general setting are developed and corresponding representation theorems for X and its extensions are proven.

At last, those representations are refined to obtain sample path properties of X : in which functional space does it live? what about its Hölder regularity?...

Titre : Régularité de processus stochastiques généralisés

Mots clés : processus généralisés, processus indexés par des ensembles, stationnarité, propriétés trajectoires, régularité hölderienne

Résumé : De plus et moins récentes études révèlent un besoin par la communauté probabiliste de comprendre des processus indexés par des espaces plus généraux que \mathbf{N} ou \mathbf{R}_+ . Sont donc étudiés dans cette thèse les processus $\mathbf{X} = \{X_t : t \in \mathcal{T}\}$ indexés par un ensemble \mathcal{T} très général muni d'une relation d'ordre représentant une forme d'écoulement temporel. Les situations concernées sont très variées et englobent certaines classes de variétés différentielles et d'arbres continus, sans négliger certains espaces ayant des saveurs plus algébriques.

La structure d'ordre permet d'identifier naturellement chaque processus $\mathbf{X} = \{X_t : t \in \mathcal{T}\}$ à un processus $X = \{X_A : A \in \mathcal{A}\}$ indexé par une certaine collection d'ensembles, créant un pont avec la théorie des processus indexés par des ensembles développée par Ivanoff et Merzbach. Sous certaines conditions, ce dernier peut être étendu à une mesure stochastique, menant à la construction d'une application linéaire correspondant à l'intégrale par rapport à X . Si le cas où X a des accroissements indépendants est bien compris depuis les travaux de Rajput et Rosiński, celui des accroissements stationnaires ou échangeables était principalement resté cantonné à \mathbf{R}_+ . On développe ici des notions de stationnarité adaptées à ce cadre général et en déduisons sous ces hypothèses des représentations pour le processus X et ses extensions.

Dans une dernière partie, ces représentations sont peaufinées pour obtenir des propriétés trajectoires sur X : dans quel espace fonctionnel vit-il ? régularité hölderienne ?...